NBER WORKING PAPER SERIES

THINKING, FAST AND SLOW? SOME FIELD EXPERIMENTS TO REDUCE CRIME AND DROPOUT IN CHICAGO

Sara B. Heller
Anuj K. Shah
Jonathan Guryan
Jens Ludwig
Sendhil Mullainathan
Harold A. Pollack

Working Paper 21178
http://www.nber.org/papers/w21178

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 20115

This project was supported by the University of Chicago's Office of the Provost, Center for Health Administration Studies, and School of Social Service Administration, the Illinois Criminal Justice Information Authority, the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health (R21-HD061757 and P01-HD076816), CDC grant 5U01CE001949-02 to the University of Chicago Center for Youth Violence Prevention, Office of Juvenile Justice and Delinquency Prevention of the U.S. Department of Justice (2012-JU-FX-0019), and grants from the Laura and John Arnold Foundation, the Chicago Community Trust, the Edna McConnell Clark Foundation, the Crown Family, the Exelon corporation, the Lloyd A. Fry Foundation, the Joyce Foundation, JPAL – North America, the Reva and David Logan Foundation, the John D. and Catherine T. MacArthur Foundation, the McCormick Foundation, the Polk Bros Foundation, the Smith Richardson Foundation, the Spencer Foundation, the University of Chicago Women's Board, a pre-doctoral fellowship to Heller from the US Department of Education's Institute for Education Sciences, visiting scholar awards to Ludwig from the Russell Sage Foundation and LIEPP at Sciences Po. For making this work possible we are grateful to the staff of

Youth Guidance, World Sport Chicago, and the Cook County Juvenile Temporary Detention Center, and to Ellen Alberding, Roseanna Ander, Hon. Richard M. Daley, Anthony Ramirez-Di Vittorio, Earl Dunlap, Hon. Rahm Emanuel, Wendy Fine, Hon. Curtis Heaston, Michelle Morrison, Dave Roush, and Robert Tracy. For helpful comments we thank Larry Katz, John Rickford and seminar participants at Case Western, Columbia, Duke, Erasmus, Harvard, MDRC, Notre Dame, Northwestern, Sciences Po, Stanford, University of Chicago, University of Miami, University of Michigan, University of Pennsylvania, University of Toronto, University of Virginia, University of Wisconsin, the MacArthur Foundation, NBER, New York City Department of Probation, and the joint New York Federal Reserve / NYU education workshop. For help accessing administrative data we thank the Chicago Public Schools, the Chicago Police Department, and to ICJIA for providing Illinois Criminal History Record Information (CHRI) data through an agreement with the Illinois State Police. For invaluable help with program monitoring, data collection and analysis we thank Sam Canas, Brice Cooke, Stephen Coussens, Gretchen Cusick, Jonathan Davis, Nathan Hess, Anindya Kundu, Heather Sophia Lee, Duff Morton, Julia Quinn, Kelsey Reid, David Showalter, Robert Webber, David Welgus, John Wolf and Sabrina Yusuf. The findings and opinions or interpretations expressed here are those of the authors and do not necessarily reflect those of the Department of Justice, National Institutes of Health, the Centers for Disease Control, the National Bureau of Economic Research or any other funder.

A randomized controlled trials registry entry is available at:
https://www.socialscienceregistry.org/trials/41

Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago
Sara B. Heller, Anuj K. Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A. Pollack
NBER Working Paper No. 21178
May 2015
JEL No. C91,C93,D03,D1,I24,I3,I32,K42

ABSTRACT

This paper describes how automatic behavior can drive disparities in youth outcomes like delinquency and dropout. We suggest that people often respond to situations without conscious deliberation. While generally adaptive, these automatic responses are sometimes deployed in situations where they are ill-suited. Although this is equally true for all youths, disadvantaged youths face greater situational variability. This increases the likelihood that automaticity will lead to negative outcomes. This hypothesis suggests that interventions that reduce automaticity can lead to positive outcomes for disadvantaged youths. We test this hypothesis by presenting the results of three large-scale randomized controlled trials (RCTs) of interventions carried out on the south and west sides of Chicago that seek to improve the outcomes of low-income youth by teaching them to be less automatic. Two of our RCTs test a program called Becoming a Man (BAM) developed by Chicago-area non-profit Youth Guidance; the first, carried out in 2009-10, shows participation improved schooling outcomes and reduced violent-crime arrests by 44%, while the second RCT in 2013-14 showed participation reduced overall arrests by 31%. The third RCT was carried out in the Cook County Juvenile Temporary Detention Center (JTDC) in 2009-11 and shows reductions in return rates of 22%. We also present results from various survey measures suggesting the results do not appear to be due to changes in mechanisms like emotional intelligence or self-control. On the other hand results from some decision-making exercises we carried out seem to support reduced automaticity as a key mechanism.

Sara B. Heller
Department of Criminology
University of Pennsylvania
McNeil Building, Suite 571
3718 Locust Walk Philadelphia, PA 19104
hellersa@sas.upenn.edu

Anuj K. Shah
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
anuj.shah@chicagobooth.edu

Jonathan Guryan
Northwestern University
Institute for Policy Research
2040 Sheridan Road
Evanston, IL 60208
and NBER
j-guryan@northwestern.edu

Jens Ludwig
University of Chicago
1155 East 60th Street
Chicago, IL 60637
and NBER
jludwig@uchicago.edu

Sendhil Mullainathan
Department of Economics
Littauer M-18
Harvard University
Cambridge, MA 02138
and Consumer Financial Protection Bureau
and also NBER
mullain@fas.harvard.edu

Harold A. Pollack
University of Chicago
School of Social Service Administration
969 East 60th Street
Chicago, IL 60637
haroldp@uchicago.edu

A data appendix is available at:
http://www.nber.org/data-appendix/w21178

I. INTRODUCTION

Disparities in youth outcomes in the United States are striking. For example, for 15-24 year olds, the male homicide rate in 2013 was 18 times higher for blacks than whites (71 vs. 4/100,000).[1] Black males lose more years of potential life before age 65 to homicide than to America's leading overall killer – heart disease.[2] A large body of research emphasizes that – beyond institutional factors – choices and behavior contribute to these outcomes, including decisions around dropping out, involvement with drugs or gangs, or how to respond to confrontations that could escalate to serious violence.

In this paper we explain these behavioral differences using the psychology of *automaticity.*[3] Because it is mentally costly to think through every situation in detail, all of us have automatic responses to the situations we encounter. These responses are tuned to the situations we commonly face. To illustrate how this can potentially create problems, consider two kinds of situations that youth face: "school life" and (for lack of a better term) "street life." In both situations, youths have to deal with assertions of authority. Teachers assert authority in school life by asking them to sit down or be quiet. In street life, someone much larger could assert authority by demanding money or their phones.

For a middle-class youth, the adaptive response in both cases is to comply. In school, they should do what the teacher says. On the street, they should hand over the phone, but then go tell an authority figure. For disadvantaged youths, school life also demands compliance. But street life is different. In places where formal social control is weak, it can be adaptive to develop a reputation as someone who will fight back when provoked to deter future victimization

[1] Our calculations compare non-Hispanic blacks to whites, and focus on homicides (excluding fatalities from legal intervention); see http://webappa.cdc.gov/sasweb/ncipc/mortrate10_us.html.

[2] http://webappa.cdc.gov/sasweb/ncipc/ypll10.html.

[3] Within psychology automaticity is most notably associated with dual systems models, which are well summarized in Kahneman (2011). Other economic models of dual systems thinking include Cunningham (2015) and for impulse control Fudenberg and Levine (2006).

(Anderson 1999, Papachristos 2009).[4] Handing over money or a phone would send a signal to others of weakness and a willingness to comply with almost any request. For disadvantaged youths, assertions of authority in street life demand resistance, not compliance.

In this example, automaticity interacts with the social environment. Students from affluent backgrounds face situations that are more homogeneous in the behaviors they demand: authority should be complied with. They can automatically comply. Disadvantaged youth, on the other hand, face heterogeneous situations that demand different responses. If they automatically comply to authority in both situations, they will be terrorized on the street. If they resist automatically in both situations they will do poorly in school.

This example illustrates our automaticity hypothesis.[5] Automatic responses are effortless, but not necessarily fine-tuned to the particular situation if there is variability across similar-looking situations in what response is adaptive. Imagine for a disadvantaged youth how automatic retaliation to assertions of authority and other provocations outside of school shapes the experience of being confronted inside of school. Being told by a teacher to sit down and be quiet so class can start may at first glance *feel* like one's reputation is at stake. This creates problems because in poor areas the response that is adaptive outside of school has negative consequences if deployed in school (e.g., suspension, expulsion). The teenager in a poor area is *not* behaving any less automatically than the teenager in the affluent area. Instead the problem

---

[4] For example Papachristos (2009, p. 79) notes: "One of the street code's most pervasive norms is that of retribution, a perversion of the 'golden rule' stipulating that personal attacks (verbal or physical) should be avenged… Failure to act in – or win – a given contest not only diminishes one's social standing vis-à-vis one's opponent but also makes one appear weak, a potential target for future street interactions."

[5] Though we use the language of automaticity, this hypothesis could just as easily be viewed as one of coarse thinking or categorization as in Mullainathan, Schwartzstein and Shleifer (2008). The essential ingredient of our hypothesis is that low income youths treat diverse circumstances that may require different responses in a homogeneous way.

arises from the variability in contexts—and the fact that some contexts call for retaliation. In this rendition, youth from disadvantaged circumstances face a higher cost of being automatic.[6]

This explanation is intriguing because it suggests a specific intervention. It suggests that we can improve the outcomes of low-income children simply by teaching them to be less automatic. Notice that the intervention is not about uniformly changing *what* the automatic responses are. For example, teaching youths to *always* comply might help outcomes in school, but could lead to adverse (even disastrous) consequences out of school. Instead, the challenge is to help youths learn when they need to not be automatic, to notice the situations where they ought to slow down and consider whether their automatic behaviors are useful for that situation.

As a concrete example, in one of the programs we study—Youth Guidance's "Becoming a Man" (BAM) program—one of the first activities for youths is the "Fist Exercise." Students are divided into pairs and one is given a ball, which the other student is told he has 30 seconds to get from him. Almost all youths attempt to use physical force to try to take the ball. During debrief, the group leader points out that no student simply *asks* for the ball. When prompted on why they did not simply ask, most respond with some version of "he wouldn't have given it," or "he would have thought I was a punk." The leader then asks the other youth, "How would you have reacted if asked nicely for the ball?" The answer inevitably is something like, "I would have given it; it's just a stupid ball." The exercise illustrates how students automatically followed one strategy rather than stepping back and considering the situation and weighing all their options carefully. The second intervention we study constantly extolls youths to "stop, look and listen."

We conduct three randomized controlled trials (RCTs) of interventions carried out with disadvantaged youth on the south and west sides of Chicago. All the interventions have the

[6] A similar idea in the field of linguistics refers to "code-switching" among people who speak more than one language or language variety, where the language they use helps convey group membership in a given setting (see for example Toribio and Bullock, 2012). This requires speakers to devote more conscious attention to identification of what social or linguistic setting they are in relative to what is required of monolingual speakers. A sociological discussion of code-switching with respect to other behaviors is in Anderson (1999).

shared component of teaching youths how to be less automatic. What makes the interventions we study particularly interesting is that they do not attempt to delineate specific behaviors as "good," but rather focus on teaching youths when and how to be less automatic and more contingent in their behavior.[7]

Our first two RCTs test a Becoming a Man (BAM),[8] which was developed by the Chicago-area non-profit Youth Guidance. Our first study randomized 2,740 males in 7th through 10th grade within 18 public schools on the south and west sides of Chicago. For the 2009-10 academic year some youths were offered some combination of BAM once a week during school or after-school sports that included BAM-like components. We find that participation in programming reduces arrests over the program year for violent crimes by 44% of the control complier mean (CCM) and 36% for other (non-violent, non-property, non-drug) crimes. Participation increases an index of school-engagement measures by 0.14 standard deviations (SD) the program year and 0.19 SD the next year, which we forecast could translate into gains in graduation rates of about 7-22%.

Our second RCT randomly assigned 2,064 male 9th and 10th graders within 9 Chicago public high schools to be offered BAM or to a control condition for the 2013-14 academic year. About half these schools had no after-school sports programming at all, which helps us distinguish the effects of BAM from the after-school component. We find a decline in total

[7] These programs all could be lumped together under the heading of what psychologists call cognitive behavioral therapy (CBT). However not all CBT programs will necessarily focus on reducing automaticity, and not all interventions to reduce automaticity will necessarily be called CBT. Since the 1970s, CBT has been used to address mental health disorders such as substance abuse, anxiety, and depression, and can be more effective than anti-depressant drugs (Rush, Beck, Kovacs and Hollon 1977). Since then, there has been growing practitioner interest in using different versions of CBT as a tool of social policy to address socially-costly behaviors. Yet there is very little good evidence currently about effects on youths for those behaviors of greatest policy concern such as delinquency, violence, and dropout (see Appendix A). A recent exception is Blattman, Jamison and Sheridan (2015) who find a CBT program for adults in Liberia is successful, especially when combined with cash grants. That program has a variety of behavioral components, including teaching anger management and self-discipline.

[8] The program is also abbreviated as "B.A.M." but for consistency we use a common style for all acronyms.

arrests equal to 31% of the CCM. The estimated effect on school engagement is a "noisy zero" with a confidence interval that allows for an effect that is about as large as in study 1.

Our third RCT tests a different intervention that was delivered to high-risk juvenile arrestees housed inside the Cook County Juvenile Temporary Detention Center (JTDC) by trained detention staff. While the curriculum in this program differs from that of the BAM program, they both share a focus on reducing automaticity. The intervention delivered inside the JTDC was part of a package of reforms implemented in some residential units within the facility but not others, which also included a token economy for good behavior inside the JTDC and increased educational requirements for staff. We worked with the JTDC to randomize 5,728 male admissions in 2009-11 to units inside the facility that did versus did not implement these reforms. We find that receipt of programming reduces return rates to the JTDC by about 16 percentage points (21% of the CCM).

Given the evidence of sizable behavior change across all three interventions, it is worth noting what these interventions are *not*. They do not involve academic remediation, or vocational education, or job training, or paid temporary jobs or internships, or early childhood education, or cash or in-kind transfers to reduce poverty. The interventions we study look different from the most common strategies tried with disadvantaged youth, and seem to be more effective.

Of course all social programs inevitably include auxiliary components—and it is possible that these and not the "slowing down," automaticity-reducing component drive these results. Having multiple experimental tests of two entirely different curricula (BAM versus the one used within juvenile detention) is helpful in this regard, since the main common element across them is reducing automaticity. Many other auxiliary components were different between the two curricula, or present in very different forms. For example, most programs for youth involve interacting with a pro-social adult, which in principle could change behavior through a generic mentoring effect. One important difference between JTDC and the two BAM studies is that the intervention inside juvenile detention was delivered by front-line detention staff, with whom youth interact in all manner of other ways in the facility. As another example, simply showing up at the BAM sessions might, through a "practice effect," teach self-control. This also is different in the JTDC intervention, since youth in confinement have no choice about participating.

We test the importance of potential alternative mechanisms by drawing on a survey that was carried out by the Chicago Public Schools (CPS) system-wide in 2011, which does not have ideal response rates but does have some measures relevant to why BAM might have worked with our study 1 cohort in 2009-10. We find no statistically significant impacts on candidate mechanisms like youth feeling they had an adult in school to help them, or perceptions of the importance of education for their futures, or survey measures of emotional intelligence or self-control (as measured by persistence or "grit"). Our estimates let us rule out anything other than very modest effects of BAM on outcomes that are mediated by these candidate mechanisms.

Beyond ruling out alternative mechanisms, we also provide some direct evidence of reduced automaticity. We had youths from our study 2 sample (the 2013-14 BAM cohort) play a modified version of an iterated dictator game, which provoked them to retaliate against unfair behavior by someone whom they believed was another youth at their school (but was actually a confederate of our research team). Our automaticity hypothesis predicts that the BAM youths would make slower, more deliberate decisions than control youths. Consistent with this prediction, we find that BAM increases the time youths spend thinking before they act by 79%.

We also randomized youths to different versions of the game that created moments that prompt the youths to do what we believe the BAM program gets youths to do on their own: slow down, and reflect on what situation they are in before they act.[9] Our theory predicts that this prompt as part of the decision-making exercise should attenuate the BAM-control difference in their propensity to slow down. That is what we find in our data.

Importantly, automaticity does *not* predict that BAM youths will unconditionally suppress aggressive or retaliatory responses, because these programs do not tell youth what their responses should be ("fight" or "don't fight"). Instead these programs help youths more deliberately choose what they feel is the appropriate response. Indeed, we find no BAM effect on overall take amount, which also helps rule out candidate explanations for program effects that focus on non-contingent changes in pro-social orientation or "self-control."

The remainder of our paper is organized as follows. Section II discusses our theory about maladaptive automatic behavior and Section III discusses how these interventions try to reduce automaticity. Section IV presents the results of our three RCTs of the effects on crime and school engagement. Section V presents evidence on mechanisms, and Section VI concludes.

II. AUTOMATICITY

One candidate explanation for different rates of dropout or crime between youths from more affluent versus disadvantaged backgrounds is disparities in the capacity for self-control or persistence ("grit"), or disparities in emotional intelligence or social skills.[10] The logic is that just as disadvantaged environments are less productive for developing academic skills like reading or math, they are also less productive for developing self-control or emotional intelligence.

---

[9] As discussed in detail below one of the randomized conditions in our decision-making experiment does the reverse and tries get youth to ruminate about the confederate's behavior in the experiment, which yields the same prediction of attenuating the BAM-control difference.

[10] These arguments stem partly from a large and growing body of research that documents rich-poor differences in different measures of self-control, grit, emotional intelligence, or social skills with different behavioral outcomes of policy interest, and also correlates these measures with future behavioral outcomes of policy interest. See, for example, Duckworth & Seligman (2005), Duckworth, Peterson, Matthews, & Kelly (2007), Evans, Gonnella, Marcynyszyn, Gentile, & Salpekar (2005), Evans & Rosenbaum (2008), Heckman (2006), Moffitt et al. (2011) and, for a widely-read popular summary of these arguments, Tough (2013).

In this section, we outline a different hypothesis that focuses instead on the role of automaticity. We suggest that automatic behavior can give the *appearance* that someone lacks self-control or emotional intelligence in some situations, which essentially results from situational misdiagnosis. The key to our explanation is that these behaviors are adaptive in some situations but not in others. Our hypothesis also helps explain why such misdiagnoses may be more common in disadvantaged areas than in more affluent areas. We do not claim that automaticity is the *only* way to explain disparities in dropout or delinquency or other behaviors that often look like they are driven by low levels of self-control or emotional intelligence. But by providing a micro-foundation for those psychological processes, automaticity suggests a concrete mechanism to target with intervention.

Our hypothesis is that relative to more affluent areas, in poor neighborhoods there is more variability in the type of automatic response that is adaptive to the different situations that people encounter. This can lead to trouble when young people have developed an automatic response to be adaptive in some situations, but then deploy it in other situations when it is actually maladaptive. Note that the problem is not automatic behavior *per se*. A key lesson of cognitive psychology and behavioral economics is that all of us frequently engage in automatic behaviors. This does not lead to problems in environments where the right automatic response is the same across all the situations someone faces. But automaticity can lead to problems in environments where people encounter situations that seem similar on the surface in many ways, but for which the appropriate automatic response is very different. Consider for example:

- What do people do when things aren't working out at the moment? There is often a tension between persisting versus exploring potentially more rewarding strategies.[11] To make this more concrete, imagine a relative promises to pick up a teenager from school,

[11] This is known in reinforcement learning as the "exploration versus exploitation" problem (see e.g., Kaelbling et al., 1996, Cohen et al., 2007). A key point is that in general there is no one optimal response for all situations.

but right now it is not working out – 15 minutes after school lets out, the ride is still not there yet. For a youth from an affluent background the adaptive response is probably to wait (persist), because the adult will eventually show up and the teen will be spared the cost of having to walk or bus home. In school, if that youth is struggling to learn some difficult material the optimal response is also to keep trying (persist); if needed, eventually a teacher or parent or peer will help the youth master the material. Now imagine a youth from a disadvantaged area where things can be more chaotic and unpredictable. When the youth is waiting for a ride after school that is running late, the adaptive response to things not working out may be to start walking home (explore a more promising strategy).[12] When this youth is struggling with schoolwork the automatic response could be either to just move on, or to persist, depending partly on whether they think of this as the same situation as someone being late picking them up from school (for example if they think "this is yet another situation where I can't rely on others").

- Almost every teenager at some point interviews for a job. Most youths from affluent areas walk in and look the interviewer in the eye. These youths probably do the same thing when they pass someone on the street in their neighborhood. While one function of eye contact is to signal dominance (e.g. Kleinke 1986), in affluent areas the need to establish dominance is low (because social control is strong) and so people on the street may not be vigilant for this cue. However in distressed neighborhoods where formal social control is weak, the importance of establishing social dominance to deter future victimization is more important (see for example Papachristos 2009). So the adaptive automatic response while on the street in these areas may be to avoid eye contact to avoid

[12] Imagine some common explanations for why a relative would be late picking someone up from school, such as an automotive problem. A rich family has a number of solution strategies to overcoming this problem, such as calling triple A for on-site assistance or a tow, using a family's second car, or calling a cab or Uber. For a poor family these strategies may not be an option so whether the youth is picked up may be more vulnerable to random events.

instigating a conflict. The fact that so many job programs for youths from distressed neighborhoods emphasize the importance of making eye contact during interviews suggests many youths may be confusing the interview context with the street context.

- The same incentives that make it adaptive to retaliate in disadvantaged areas (the example we develop in the introduction above) yield another prediction that can help distinguish automaticity from other hypotheses. In places with weak formal social control, the fact that it is adaptive to retaliate and respond aggressively (or even disproportionately) to provocation means that any altercation can quickly spiral out of control. There is no ceiling on how far a conflict might escalate. So in high-poverty areas it is adaptive to avoid *instigating* conflicts.[13] Automaticity predicts that youths in poor areas are *more* likely than those in affluent areas to retaliate, but *less* likely to instigate.

To explore this last idea we draw on data from the 2013 wave of the Center for Disease Control's Youth Risk Behavior Survey (YRBS). Since family or neighborhood incomes were not collected in this wave of the YRBS, we use race as a proxy for socio-economic status (SES). The survey includes several questions that capture what are arguably proxies for "retaliation" – frequency of fighting, and of being injured in a fight – since a fight usually implies a back-and-forth. The 2013 YRBS also includes what could be interpreted as a measure of provocation or "instigation" – bullying, which involves unprovoked aggression. While we would ideally like a measure of whether YRBS respondents bully other people (that is, "instigate"), the survey only includes a measure of having experienced bullying. But since America's public schools remain

[13] For example from Anderson (1999, p. 77-8): "Further down the block a woman simply stops her car in the middle of the street, waiting for her husband or boyfriend to emerge from a barbershop. She waits for about ten minutes, holding up traffic. No one complains, no one honks his horn; they simply go around her, for they know that to complain is to risk an altercation, or at least heated words. They prefer not to incur this woman's wrath, which could escalate to warfare."

so racially segregated,[14] black-white differences in having been bullied are presumably also informative about black-white differences in bullying. As with all self-reported survey questions, misreporting is a possibility here and so the results should be interpreted with that in mind.

Table I provides some suggestive evidence that is consistent with the automaticity hypothesis. The first two rows show that youths from on average more disadvantaged circumstances are more likely to retaliate— black male youth are 10 percentage points more likely than whites to say they had been in a fight the past 12 months ($p<.01$) and nearly twice as likely as whites to say they had been injured in a fight ($p<.05$). The data shown in the first two rows are equally consistent with either automaticity or "low capacity for self-control." But the last row shows poor youth were *less* likely to instigate—10.2% of black youth say they had been bullied on school property the past 12 months compared to 16.2% of white youth ($p<.01$).

These data offer some initial support for the notion that disadvantaged youths rely on two competing scripts related to aggression—to retaliate when provoked, but to diligently avoid starting a fight. Automatic responses can make disadvantaged youths seem like they lack self-control and also have tremendous amounts of self-control, because their behavior is more *contingent*. For youth growing up in wealthy, safe environments, their behavior does not *have* to be as contingent. The expected demeanor at home or in the neighborhood matches what is productive and adaptive in school. Poor youth grow up in areas that require *less* automaticity.

III. INTERVENTION STRATEGY

What does the automaticity hypothesis imply for intervention? The potential solution for maladaptive automatic responses is *not* to teach youth to deploy the same "pro-social" automatic response in every situation. A program that taught youth to "never fight" could lead to dire

[14] For example Orfield et al. (2014) note that the average white student attends a school that is nearly three-quarters white, while the average black student attends a school that is nearly three-quarters non-white (see for example Table 4, page 12).

consequences when they encounter the challenges that unfortunately remain far too common in disadvantaged neighborhoods. Instead what is needed is an intervention to help youths recognize when they are in a high-stakes situation where their automatic responses might be maladaptive, slow down and behave less automatically, and think about the situation they are in and what response is called for. That is exactly the goal of several programs that we studied in Chicago. This section describes what this looks like in practice.

For starters these programs help youths figure out when they are in situations in which their automatic responses might get them into trouble. For example a standard strategy is to get people to recognize that a shift towards some more aversive emotion is an important cue (see for example Beck 2011). Given the context on the south and west sides of Chicago, both of the programs that we study – Youth Guidance's Becoming a Man (BAM) program, and the curriculum employed in the Cook County JTDC – focus on anger as a particularly important indicator for youths. Other important affective cues include feeling threatened, disrespected, envy, frustrated, or having a sense of loss.

When youths encounter these high-stakes situations, both programs help them learn techniques to slow down and behave less automatically – such as deep-breathing and other relaxation techniques. In the BAM program, there is continual emphasis on avoiding "out of control" energy. BAM uses engaging activities called "group missions," which are meant to put students into an agitated or "hot" state and to then help them practice slowing down – to pause and reflect on their thinking while in that state. The program within the JTDC continually emphasizes the importance of learning to "stop, look and listen."

Finally, the programs we study all help youths try to objectively assess the type of situation in which they find themselves. For example, youths in both BAM and the JTDC programs are taught to try to examine high-stakes situations from an outside perspective and distinguish between the facts apparent to any neutral outside observer ("What would a *camera* see in this situation?") versus the meanings or intentions they are projecting on the situation themselves. Practice diagnosing maladaptive automatic responses to key situations comes partly in the BAM program from starting off most sessions with a self-analysis ("check in"), a common element of CBT programs (Beck 2011). In the JTDC program, youths are required to carry out "thinking reports" every time their misbehavior causes detention staff to give them a "time out" (a certain amount of time alone in their cell). Part of learning how to be more deliberate and reflective in situations where one's automatic responses may be maladaptive is to also become more empirical in testing what type of situation one is in, and so both programs help youths learn how to carry out "behavioral experiments" that test their perceptions or beliefs about what type of situation they find themselves in.

Both programs are manualized and can be delivered by college-educated people without specialized training in psychology or social work – although for the BAM program, Youth Guidance had a preference for such training in selecting program providers. From observing sessions of both programs (BAM sessions in Chicago public schools and sessions within the JTDC), it also seems clear that another key skill is the ability to keep youths engaged.

From observing both programs it also seems important for the programs themselves to be culturally competent (in particular they need to avoid being "corny") and to be engaging to a target population that may include many youth who do not normally volunteer for social programs. For example, from an engagement perspective the "Fist Exercise" described in the Introduction is slightly (and cleverly) subversive – to be in BAM, youths get out of a regular class, and then the first activity winds up involving sometimes-rowdy horseplay. Whether these interventions have effects on youths' behaviors in practice is taken up in the next section.

IV. RANDOMIZED TRIAL IMPACTS ON BEHAVIORAL OUTCOMES

This section presents the results of three large-scale RCTs of two different programs that include reducing automaticity as a key element. The first two studies are of the BAM program delivered to male youths in schools, while the third is a program delivered to high-risk juveniles inside the juvenile detention center. The exact outcome measures we examine are not quite identical across all three studies, nor are the exact pattern of impacts, but all three RCTs do find sizable youth responses on various measures of criminal behavior or schooling.

A. STUDIES 1 AND 2: BECOMING A MAN

Our first study of Youth Guidance's Becoming a Man (BAM) program was in 2009-10. During the summer of 2009, our team recruited 18 elementary and high schools in the CPS system located on Chicago's low-income, racially segregated south and west sides, where the city's violent crime is disproportionately concentrated (see Figure I). Our study sample is essentially the 2,740 highest-risk male students in grades 7-10 in these schools, after excluding students who rarely attend school (and so would not benefit from a school-based intervention) or have serious disabilities. This sample represents around 75 percent of all male youths in grades 7-10 in the study schools (see Appendix B for details). Our second BAM study was carried out in 2013-14. Our study sample consists of 2,064 male 9th and 10th graders attending 9 CPS high schools that are mostly located on the south and west sides of Chicago (see Figure I).

Both studies were block-randomized experiments, where students are the unit of randomization and are randomly assigned within schools ("blocks"). In study 1 youths were randomized to treatment (in-school BAM, after-school sports programming that incorporated BAM elements,[15] or both) or control groups.[16] Unfortunately in study 1 we do not have adequate

[15] The after-school programming was delivered by World Sports Chicago (WSC) and designed to both enhance program participation rates and provide youth with more opportunities to reflect on their automatic responses and decision-making. The WSC coaches all receive some training in the BAM program. WSC sessions, one-to-two

statistical power to disentangle the separate effects of BAM from the after-school program. The in-school treatment offered the chance to participate in up to 27 one-hour, once-per-week group sessions during the school day over the school year. The intervention is delivered in groups to help control costs, with groups kept small (assigned groups of no more than 15 youths and a realized average youth-to-adult ratio of 8:1) to help develop relationships. Students skip a class in order to participate, which is one of the draws for many youths to attend.

In study 2 youths were randomly assigned to either be offered the chance to participate in BAM during the school day once a week over the 2013-14 academic year, or to a control condition that was eligible to receive whatever status quo services are currently provided to youth in these CPS high schools. This second study helps us isolate the effect of BAM from sports because four of the nine schools had no sports programming (in the other schools, the sports attendance rates were also very low). As we discuss below, there are sizable effects when we look just at the schools without any sports programming at all.

Tables II and III show that both studies enrolled very disadvantaged samples of youths, and that random assignment appears to have been carried out correctly. In both studies youths were about 15 years old at baseline. Reflecting the composition of their neighborhoods, almost all youths in both studies are minorities – around 70 percent black, the remainder Hispanic. The year before randomization the average study youth had a GPA on the order of about a 2.0 on a 4.0 scale (somewhat lower in study 1 than study 2),[17] and missed about 8 weeks of school on

---

hours each, include non-traditional sports (archery, boxing, wrestling, weightlifting, handball, and martial arts) that require focus, self-control, and proper channeling of aggression, and also provide youth with additional opportunities for reflection on their automatic behavior ("so after you got hit in the face during that boxing match, what were you thinking that led you to drop your hands and charge blindly?").

[16] Three of our 18 schools could not accommodate after-school programming because of logistical or space reasons, so they included only in-school and control conditions. Eight schools offered both in- and after-school treatment arms in some combination, but did not offer all three treatment arms in addition to the control group.

[17] The differences in baseline disadvantage between the study 1 and 2 samples could in principle be due to differences in the age range or schools used to draw the sample, rather than to a higher level of baseline disadvantage among the study 1 youth per se. However when we hold age constant and compare youth in 9th and

average for the study 1 sample and about 6 weeks of school in the study 2 sample.[18] A sizable share of youths had been arrested at some point prior to random assignment. Of 19 total baseline covariates we examined in study 1,[19] none of the treatment-control differences is significantly different at the 5 percent level. An F-test for the joint significance of all baseline characteristics shows we cannot reject the null hypothesis that treatment and control groups are equivalent ($F(18,2543)=1.02$, $p=0.431$). Similarly for study 2, an F-test of the null that the differences in all of these baseline variables are jointly zero has a p-value of 0.266 ($F(13,1752)=1.21$).

In both BAM studies the share of youths who were offered the chance to participate in program activities that actually participated was on the order of one-half. This take-up rate is consistent with other large scale social experiments (Bloom, Orr, Bell, Cave, Doolittle, Lin and Bos 1997, Kling, Liebman and Katz 2007) despite the fact that we randomized first (using administrative data) and then tried to consent people for program participation, rather than consenting and then randomizing, as is more common.[20] We suspect participation rates for the after-school programming in study 1 are under-stated because of inadequate record keeping; Appendix C talks about how we handle this issue. Among participants, the average number of sessions is around 13 in both studies 1 and 2. Some controls also wound up getting services.

Our main schooling outcomes come from longitudinal student-level CPS records on the program year, and for study 1, also for the following year. We used the post-randomization data to form a summary index of three schooling outcomes in Z-score form (GPA, days present, and enrollment status at the end of the year). Use of an index reduces the number of hypothesis tests

---

10th grade across the two studies, and focus on youth enrolled in the three schools common to both of these BAM studies, we still see that the study 1 youth are more disadvantaged on average than those in study 2.

[18] At the time of the first BAM study in AY 2009-10, the CPS school year was 170 days; by the time of the second BAM study in AY 2013-14 the CPS school year had been increased to about 180 days, and so the total days present figure of 149 in Table IV implies about 6 weeks of missed school for the study 2 youth.

[19] The baseline variables in the joint test are: age, grade, number of in-school suspensions, number of out-of-school suspensions, number of each type of arrest (violent, property, drug, and other), indicator variables for being black, Hispanic, old-for-grade, and having an IEP, and each of the number of grades earned (A through F).

[20] Consent was for program participation only; outcome data is available for all youth who were randomized.

and so reduces risk of "false positives" (Anderson 2008, Kling, Liebman and Katz 2007, Westfall and Young 1993), and improves the statistical power available to detect effects for outcomes within a given family expected to move in a similar direction. This approach assumes data are missing completely at random, but our results are similar when we relax this assumption and use other approaches such as multiple imputation (see Appendix C).

To measure criminal behavior by program participants, for study 1 we use electronic arrest records (or "rap sheets") from the Illinois State Police (ISP), which were matched to our study sample for research purposes by the Illinois Criminal Justice Information Authority (ICJIA) using probabilistic matching on name and date of birth.[21] For study 2 we use arrest data from the Chicago Police Department. Arrest records avoid the problem of under-reporting of criminal involvement in survey data (Kling, Ludwig and Katz 2005) but require the assumption that the intervention itself does not affect the likelihood that criminal behavior results in arrest. Because intervention impacts (and social costs) can vary by crime type (Deming 2011, Evans and Owens 2007, Kling, Ludwig and Katz 2005, Lochner and Moretti 2004, Weiner, Lutz and Ludwig 2009), in addition to looking at total arrests we separately present results for violent, property, drug, and "other" crimes (excluding motor vehicle violations). There is no issue with missing data with arrests in the sense that we cannot distinguish missing data from no arrests.

Given our randomized experimental design, our analyses of these data are quite straightforward. Let Y_{ist} denote some post-program outcome for individual *i* at school *s* during post-randomization period *t*, which is a function of treatment group assignment (Z_{is}) and observed variables from ISP and CPS records measured at or before baseline ($X_{is(t-1)}$) as in equation (1) below. We condition on baseline characteristics to improve precision by accounting

[21] The ISP records capture arrests in the state going back to 1990 and include arrests of people below the age of majority within the criminal justice system (juvenile arrests), as well as to those who are above the age of majority. Local police departments are required by law to report all juvenile felony arrests to the ISP, and optionally class A and B misdemeanors.

for residual variation in the outcomes (results without baseline covariates are similar; available upon request).[22] We also control for the "blocking variable" with school fixed effects (γ_s). We present robust standard errors but do not cluster by school, partly because the school fixed-effects account for within-school correlations across students in mean outcomes and partly because with the modest number of "clusters" we have (18 in total) the asymptotic theory on which clustering is based is not applicable. As a robustness check we also present p-values from the wild-t bootstrap of Cameron, Gelbach and Miller (2008), which generally yield similar inferences. The intention to treat effect (ITT) captures the effect of being offered the chance to participate in the program, and is given by the estimate of π_1 in equation (1).

(1) $Y_{ist} = Z_{is}\pi_1 + X_{is(t-1)}\beta_1 + \gamma_s + \varepsilon_{ist1}$

We also report the effect of participating in the program for those who actually participate, which we estimate using two-stage least squares with random assignment (Z_{is}) as an instrumental variable (IV) for participation (P_{ist}), as in equations (2) and (3) (Angrist, Imbens and Rubin 1996, Bloom 1984). This assumes treatment-group assignment has no effect on the behavior of youth who do not participate in the intervention.

(2) $P_{ist} = Z_{is}\pi_1 + X_{is(t-1)}\beta_1 + \gamma_s + \varepsilon_{ist2}$

(3) $Y_{ist} = P_{ist}\pi_2 + X_{is(t-1)}\beta_2 + \gamma_s + \varepsilon_{ist3}$

Since a small share of controls get the program, π_2 is technically a local average treatment effect (LATE) but crossover is low and so this should be close to the effect of treatment on the treated (TOT). We benchmark the size of these effects with the control complier

[22] Specifically, we control for the following variables from the 2008-9 academic year: total days present; number of in- and out-of-school suspensions; number of each grade category (A, B, C, D, and F); dummies for ages 14-15, 15-16, and over 17; black and Hispanic dummies; an indicator for having an Individual Educational Program (IEP); a linear grade term; and dummies for having zero, one, two, or three and over arrests of each type. For the one case with missing baseline covariates, we assign a value of zero and include an indicator that the variable is missing.

mean (CCM) (see Katz, Kling and Liebman 2001) estimated using the formula from Heller et al. (2013).[23]

Table IV shows that in study 1 BAM improves schooling outcomes during both the program year (AY 2009-10), as well as the follow-up year (AY 2010-11). The estimated effect of participation on our schooling index during the program year (top panel) equals 0.14SD. The pairwise-comparison p-value is statistically significant using robust standard errors, or using the wild-t bootstrap (p=.034). In addition as discussed in Appendix C the result is significant at the 5% cutoff when we use either a bootstrap re-sampling method to control for the family-wise error rate, or FWER (the chance that at least one of our outcome measures is significant when the null hypothesis of no effect is true) or if we control for the false discover rate, or FDR (the share of statistically significant estimates that would be expected to be false positives). The impact on schooling outcomes in study 1 persists through the follow-up year, and is, if anything, slightly larger than during the program year (0.19 vs. 0.14SD). Our sample is still too young for most of our youths to have reached graduation age. But based on non-experimental correlations between our different school-engagement measures and subsequent graduation outcomes from previous longitudinal studies of CPS students, we estimate our schooling impacts could increase graduation rates by 3-10 percentage points (7-22%).

Table IV also shows that in study 1 the program generates very large reductions in arrests for violent crimes and "other" crimes during the program year (arrests made between September

[23] If C indicates being a "complier" and Z indicates treatment assignment, the CCM equals: CCM = E(Y|C=1, Z=1) – [E(Y|C=1, Z=1) – E(Y|C=1, Z=0)]. The term in brackets is our LATE estimate. However, we must recover the first right-hand-side term, E(Y|C=1, Z=1), since what we observe in the data is the mean outcome for all treatment group participants - a weighted average of the mean outcomes for compliers and always-takers. Let P indicate actual participation and A be an indicator for always-takers. Then:

$$E(Y \mid Z=1, P=1) = E(Y \mid Z=1, C=1)\left(1 - \frac{E(A \mid Z=1)}{E(P \mid Z=1)}\right) + E(Y \mid Z=1, A=1)\left(\frac{E(A \mid Z=1)}{E(P \mid Z=1)}\right)$$

To recover E(Y|Z=1, C=1) for the CCM calculation, we can estimate the left-hand side and E(P|Z=1) directly from the data, and use random assignment to replace E(A|Z=1) with E(A|Z=0) and E(Y|Z=1, A=1) with E(Y|Z=0,A=1). In our case, block randomization means that these equalities should also be conditional on school. In practice, the difference that calculating them conditionally makes is trivial. In other words, we assume treatment- and control-group always-takers are equivalent on average.

2009 through August 2010), which are no longer statistically significant in the follow-up year (arrests that occur from September 2010 through July 2011). The top panel shows that during the program year, program participation reduces violent-crime arrests by about 8 per 100 youths, equal to about 44% of the CCM (18 per 100 youths). The p-value from the pairwise comparison is significant at conventional levels using either the robust standard error or the wild-t bootstrap (p=0.04). The FWER-adjusted p-value is 0.16, which suggests some caution in interpreting these results, but the FWER is somewhat conservative (as discussed in Appendix C). Using the two-stage FDR-control procedure from Benjamini, Krieger and Yekutieli (2006) we can reject the null of no violent-crime arrests at q=0.1 (that is, as long as we are willing to tolerate the risk that 10% of pair-wise significant p-values are false positives).

In Table IV the top panel's last row shows that program participation reduces the number of arrests for "other" (non-violent, non-property, non-drug) crimes by nearly 12 arrests per 100 youths during the program year, about a 38% reduction relative to the CCM (32 per 100). This impact is driven by reductions in weapons offenses, trespassing, and vandalism (each account for about one-quarter of the total effect).[24] The estimate for BAM's effect on the total arrests for any offense (that is, all offenses pooled together) is negative and proportionately large (27% of the CCM), with a t-statistic for the pair-wise comparison equal to t=1.64.

The bottom panel shows that for the follow-up year, none of the impact estimates is statistically significant. The impacts on both total arrests and arrests for "other crimes" are still large in proportional terms, equal to about one-quarter and one-third of the CCM, respectively

[24] Arrests for disorderly conduct or disobeying a police officer, which together account for a third of arrests in this category, and could in principle have declined because youth are now just better able to interact more constructively with law enforcement, appear to be basically unaffected.

(with t-statistics for the pair-wise comparison of 1.53 and 1.61). Arrest rates decline for our sample from year 1 to 2, consistent with declines in overall Chicago crime over this period.[25]

Table V presents the results of BAM on outcomes for the study 2 sample, which show that all of the individual outcomes we examined in study 1 are in the direction of better outcomes for youth but not quite statistically significant. We do not see statistically significant reductions in violent-crime arrests or school engagement, but our 95% confidence intervals do not let us rule out effects about as large as what we saw in study 1 – so these are imprecisely estimated "zeros." It is also the case that CPS practices seem to be changing over time as well (recorded attendance rates from 2009 to present have increased by 8% system-wide and by 12% in our study schools, recorded out-of-school suspensions have declined by 25% since 2012, and there have also been substantial changes in the CPS early-warning indicator for being "on track" for high school graduation[26]), which is a candidate explanation for why the results for our school engagement index were statistically significant in study 1 but not in study 2.

We see a sizable reduction in total arrests for all offenses combined, equal to 16 fewer arrests per 100 youths over the study year (31 percent of the CCM), which has a p-value for the pair-wise comparison equal to 0.07. As shown in Appendix C, the formal FDR calculation that adjusts for multiple comparisons suggests this has about a one-in-four chance of being a false positive. Yet the fact that we see statistically significant impacts with this same study sample on a measure of "automaticity," together with the fact that we see significant impacts from similar interventions on behavioral outcomes in studies 1 and 3, would presumably shift downward

[25] For example, the overall rate of Chicago homicides declined by 10% from 2008 to 2009, by 6% from 2009 to 2010, and by another 1% from 2010 to 2011. Declines in youth homicide arrests are even more pronounced. There were 15 homicide arrests to 14-16 year olds in 2008 and just 11 in 2011, and 164 among 17-25 year olds in 2008 and just 91 in 2011 (CPD 2011b). From 2009 to 2010 robbery rates declined by 10% and aggravated assault rates declined by 8.5%, while property-crime arrests were basically unchanged (0.4% drop) (CPD 2011a).

[26] The "on track" indicator was developed by the Chicago Consortium on School Research (CCSR) for high school freshmen. Students are "on track" if they accumulate five full-year course credits (in any credit-bearing class), and accumulate not more than one semester F grade in a core class (Allensworth and Easton 2005). From 2010 to 2013, the share of freshman CPS reports is "on track" increased from 65% to 82% system-wide.

one's beliefs about this false-positive risk. For the crime-type-specific arrest categories that make up total arrests we see declines that range from one-fifth to two-fifths of the CCM.

Study 2 unlike study 1 also has the advantage of helping us separate out the effects of the in-school BAM programming from the BAM-infused after-school sports programming. When we analyze data from the subset of schools that had no sports programming (just BAM) the estimated effect on total arrests is not so different from what we see for the full sample, just somewhat less precisely estimated.[27]

B. STUDY 3: COOK COUNTY JUVENILE TEMPORARY DETENTION CENTER

The setting for our third RCT is the Cook County Juvenile Temporary Detention Center (JTDC), which is where the highest-risk juvenile arrestees in Cook County are taken after they are arrested and held until their cases are adjudicated by the juvenile courts (on average 3 to 4 weeks). In May 2007, the JTDC was taken over by a federal judge as the result of an ACLU lawsuit (*Doe v. Cook County*). One of the first acts of the temporary administrator that began to run the 500-bed facility (Earl Dunlap) was to divide the facility into 10 essentially separate residential centers of around 50 beds each, and begin to enact some reforms one-by-one in each of these centers intended to make them more developmentally productive for youth residents.

The reforms included the use of a "token economy" system to help maintain order (staff physical abuse of youth had been a problem prior to the lawsuit), and in the afternoons after youth attended school, rather than have youth watch television, provide them with group sessions in the program described above which has an emphasis on reducing automaticity. The program used a manualized curriculum[28] and was delivered by trained JTDC staff. Partly to help

[27] The ITT is -.08 (0.06) in the schools with no after-school sports programming versus -0.11 (0.06) in the schools with some sports programming (though few youth participated in after-school sports even in those schools).
[28] The specific intervention studied here was developed by Dr. Bernie Glos and his associates from the DuPage County, IL Juvenile Detention Center. The curriculum is adapted from the best material from several prior CBT models that had been used in detention and is based in part on the cognitive behavioral training ideas from Maultsby (1975, 1990) (see also Ellis 1957, Ellis & Harper 1975).

implement both changes, the JTDC also required increased educational requirements for staff working in the newly reformed centers.

The expansion of Dunlap's reforms was halted halfway through due to litigation initiated by the union representing the JTDC staff. The result was that, for an extended period, half the JTDC operated as what we call "treatment centers" while the other half of residential centers used the previous standard operating procedures. Our research team worked with the JTDC staff to implement a randomization algorithm that assigned all incoming male youth to either treatment or control units from November 10, 2009 through March 2, 2011; then the litigation was resolved and the entire facility switched over to new treatment operations. (Girls were not randomized because there are so few girls they are all housed in a single residential center).

We focus on the 2,693 male admissions to the JTDC during our study period for which we have ≥18 months of follow-up data, so that we have a balanced panel for the full duration of the 18 month follow-up period. (Results for the full sample of 5,728 male admissions in the JTDC during our study period are presented in the appendix.) While random assignment was not binding for some youths because of safety or operational reasons, or because they had been assigned to a treatment unit inside the JTDC previously (see Appendix B), randomization greatly increases the likelihood of placement in a treatment unit. The ITT effect of random assignment on placement is about 25 percentage points (40% take-up rate among spells in which youth are assigned to treatment, versus 14% for controls); the first stage F-statistic is 241. We thus have an "encouragement design," where randomization is a valid instrument for estimating the unbiased causal effect of participation on compliers as in equations (2) and (3) above.

The data we have on these youths include intake forms that provide basic demographics and addresses, admissions logs, which the admissions staff uses to record who enters the facility each day, and the JTDC's housing roster, which captures the residential unit in which a youth is located on each day and so lets us measure receipt of treatment. We have these data through December 2011. Our main results focus on a common measure of recidivism – re-entry into the JTDC facility itself. We have also linked these youths to the CPD and ISP arrest databases.

Table VI shows descriptive statistics for the study sample. Consistent with the pattern of incarceration nationwide, the large majority of male admissions is African-American, despite the fact that only one-third of Cook County's youth population is black. These youths come disproportionately from Chicago's economically and racially segregated south and west sides. Being in detention is rarely a one-time event; the average spell in our sample is a youth's third entry into detention.[29] Consistent with successful randomization, none of the treatment-control differences are statistically different; we also fail to reject the null hypothesis that the differences in all available baseline variables are jointly zero ($F(25, 1835) = 0.79$, $p = 0.76$).[30]

Figure II shows the effects of being in a reformed treatment unit on rates of re-admission to the JTDC measured at different points in time since release from the facility. The first panel shows that 2 months after JTDC exit the intention to treat effect (ITT) is a decline in re-admission rates of about 3 percentage points; by 18 months the ITT effect is about 4 percentage points. (The figure also shows the result is not very sensitive to excluding day-of-admission fixed effects in the estimating equation, which is structured like equation 1 above, controlling for the baseline characteristics in Table VI – see the appendix for more details).

[29] Among the 1,862 individuals who make up these 2,693 spells, each visits the JTDC an average of 4.4 times before the end of our data (the maximum total spells per individual over the 7 years of our housing roster data is 23).

[30] Baseline covariates in the joint test are: spell number, age, number of each type of baseline arrest (violent, property, drug, motor vehicle, other), indicators for race (white, Hispanic, other), type of admitting offense (violent, property, drug, or other arrest, or direct admission with no arrest), and neighborhood characteristics from the ACS (unemployment, median income, percent below poverty, percent white, percent black, percent Hispanic, and percent with at least a high school degree). Only non-missing covariates are used in joint test. For outcome regressions, we impute zeros for missing values and include indicator variables for missing-ness. Outcome baseline covariates are: age at entry and age squared, neighborhood characteristics (percent with at least a high school degree, percent black, percent unemployed), and indicator variables for spell number (2, 3, 4, and 5 or over), admission reason (violent, property, drug, or other arrest), and number of each type of baseline arrest (1, 2, 3, or 4 and up, except motor vehicle arrests which are 1 or 2 and up).

The second panel shows that the effect of actually being in a treatment unit on the compliers (the local average treatment effect) is about 13 percentage points 2 months after release and grows slightly to 16 percentage points by 18 months following exit from the JTDC (equal to 41% and 21% of the CCMs, respectively). A different way to gauge the size of this effect is to note that only around one-quarter of the control compliers had *not* been re-admitted to the JTDC within 18 months; being in a treatment unit increases that rate by about two-thirds. In Appendix C we also show sizable effects on the number of re-admissions and arrests to youths.

V. MECHANISMS

In this section we present data on candidate mechanisms of action through which these interventions change behavioral outcomes. We first show that the data that we have, while imperfect in some ways, is not consistent with several obvious candidate explanations for the observed impacts we see from our interventions such as incapacitation of youth in after-school programming, generic mentoring effects, or changes in "grit." We then present findings that provide some support for our automaticity hypothesis.

A. RULING OUT LARGE ROLES FOR AUXILIARY PROGRAM COMPONENTS

As with all social-policy interventions, there are some auxiliary components to these programs aside from reducing automaticity that could in principle account for the behavioral impacts that we report in Section IV. In study 1 some youths receive after-school programming; while serious youth crime is disproportionately concentrated on weekends and summer months, in principle there could have been some mechanical effect in reducing arrests on some school day afternoons by incapacitating youth. In studies 1 and 2, BAM took youths on a field trip to a local college, which might change perceptions of the returns to schooling. And in all three studies youths there is also the issue of program attendance; if self-control is a muscle and attending anything helps develop self-control, then attendance itself could increase self-control.

All three studies involve an adult interacting with a group of youth, which could create some "mentoring effect."

To examine the possible "incapacitation" effects on crime from the after-school programming that was included as part of study 1 for some youths, we exploit the fact that the provider records include the dates of each after-school session and that our ISP rap sheets have the date of each arrest. We find that the estimated effect of BAM on arrests is *not* concentrated on days when after-school programming is held.[31]

Table VII explores the potential influence of the other candidate mechanisms of action on behavioral outcomes. Our data on these mechanisms come from ongoing, bi-annual surveys carried out over the web in all Chicago Public Schools by the Consortium on Chicago School Research (CCSR), designed to measure student perceptions of themselves and their school environments.[32] We use CCSR survey data from the spring 2011, the end of the year *after* the 2009-10 BAM intervention that we examined in study 1 (CCSR did not do a spring 2010 survey). It should be noted that the response rate on this survey among our sample is not ideal, and is a few percentage points higher for the treatment versus control groups (42 versus 38%, $p<.05$). The different candidate measures we examine (reported separately in the rows of Table VII) include a measure of social capital or mentoring ("I have at least one teacher or adult in the school I can talk to if I have a problem"), perceptions about the returns to or importance of schooling for their futures, social skills (e.g., "I can always find a way to help people end arguments"), and persistence or grit (e.g., "I finish whatever I begin," "I don't give up easily").

[31] The ITT effect on an indicator for any violent-crime arrest during days when after-school programming is not offered is $\beta = -0.0217$ (se=0.0103), p = 0.035, CM 0.046, vs. days after-school programming is offered, $\beta = -0.0061$ (0.0076), p = 0.420, CM = 0.094). These estimates do not adjust for the larger number of non-programming days.

[32] The 30-minute survey is designed to address a number of questions regarding school culture and climate. CCSR has administered surveys to CPS teachers, students, and principals for two decades. In Spring 2011, surveys were received from ~146,000 students in more than 600 schools. All students in grades 6-12 and all teachers were asked to participate. The student web survey was administered within each school during school hours with each response registered on a Likert scale. CCSR used Rasch analysis on individual survey items. We standardized these measures into SD units based on the observed distribution within the control group.

The first column of results in Table VII presents the results of estimating BAM participation's effects using equations (2) and (3) for the youths in study 1 who completed the CCSR surveys, with each candidate mechanism as the second-stage outcome in turn. The second column reports the coefficients from using data just from the control group to run a non-experimental regression of one of our outcomes (the school engagement index measured at the end of AY 2009-10) against each candidate mediator in turn, controlling for the same set of baseline covariates and school fixed effects that are in our main ITT and IV estimating equations (1)-(3) above. The third column reports the share of the total BAM participation effect on the schooling outcome that could be explained by each candidate mechanism, which comes from multiplying the (experimentally-estimated) BAM → mechanism link reported in column 1 by the (non-experimentally-estimated) mechanism → outcome link in column 2, and then dividing by the estimated BAM participation effect on the schooling outcome reported in Table III above. We also obtain upper and lower bounds for this effect by multiplying the upper and lower bounds of the 95 percent confidence intervals in columns 1 and 2 together, respectively. The last two columns of the table repeat this exercise for our measure of violent-crime arrests.

The results presented in Table VII suggest that these different counter-explanations to automaticity are unlikely to explain much of the BAM impact on behavior reported in study 1. None of the effects of BAM on any of these candidate mechanisms are statistically significant. Moreover the estimates are not very large. Our results imply that at most a very small share of the BAM → outcome relationship reported in Section IV could be explained by the BAM → mechanism → outcome chain for each candidate mechanism reported in Table VII. Partly this is due also to the fact that these candidate mechanisms, which are common to programs for youths, do not seem to be very strongly related to the key behavioral outcomes of interest.[33,34]

B. EVIDENCE REGARDING AUTOMATICITY AS A MECHANISM

Finally, we present some evidence that is consistent with the idea that automaticity may be an important mechanism through which BAM changes behavioral outcomes.[35] From the sample of youths randomized to BAM versus control in 2013-14 (study 2 in Section IV), we recruited 493 participants (266 who had been assigned to BAM, 227 assigned to control) from 9 schools in which 1547 youths (774 treatment, 773 control) were eligible to participate. One reason for non-participation is many youths randomized in study 2 never showed up at the school CPS thought they would attend; the response rate for youths attending study schools was 44%.

To examine how BAM changes youths' decision-making in confrontational situations in which youths are provoked and retaliation is a possibility, and specifically whether BAM causes youths to "slow down," we had participants play a modified version of a real-stakes iterated dictator game.[36] Youths were told that they would be playing with a "partner" who was another

[33] The size of the relationships between these candidate mechanisms and outcomes are not so different from those reported in other papers. For example the raw correlation in our study 1 sample between GPA in 2010-11 and the grit scale we use, which was constructed by CCSR, equals 0.20; the correlation reported by Duckworth and Quinn (2009) using a more socio-economically and racially/ethnically diverse sample of youth was 0.30 (Table 7, p. 170).

[34] The candidate mechanism for which there is probably the best empirical evidence is for mentoring, from two randomized experimental studies of the Big Brothers / Big Sisters (BB/BS) mentoring program, which yield mixed results. An initial RCT of BB/BS community-based mentoring for children ages 10-16 found evidence of beneficial effects on schooling outcomes like GPA and attendance, and some behavioral measures such as drug and alcohol use or hitting someone else (but not theft or property damage) . The BB/BS study finds few effects when examining our study sample (minority males). The study relied on self-reported outcomes, so there is always the risk that mentoring changed the willingness of youth to report bad behavior rather than the actual behavioral outcomes themselves (social desirability bias). A more recent study of BB/BS mentoring done within schools with children 10-16 found effects on some schooling outcomes in the program year that did not persist to the follow-up year, and found no statistically significant effects on out-of-school behaviors.

[35] This study was adapted from VanderMeer et al. (2015), where they found that participants in the reflection condition described below were significantly less likely to retaliate than were participants in the distraction and no-instruction conditions. This was taken as evidence that retaliation is often an automatic response which can be curbed by reflecting on whether the initial event is actually severe enough to warrant retaliation.

[36] Students were informed during lunch periods that a brief study would be conducted giving them the chance to earn about $10. Because parental consent was required, consent forms were handed out and made available in the school several days before we began conducting the studies. Students could return the consent forms anytime during the duration of the study (approximately three weeks in each school). Students who returned consent forms could

student in their school for multiple rounds (they were not told how many rounds). A research assistant (RA) led each participant through the study. RAs told participants that they would be communicating over walkie-talkie with another RA who was standing with their "partner." However, there was no partner; the other RA was actually a confederate who followed a script.[37]

In the first round, participants were given $10 in one-dollar bills in an envelope. Their "partner" was given the chance to take some money away from the participant. The participants heard the confederate say over the walkie-talkie that the "partner" was taking $6 from the participant. The participant was then asked how much money they would like to take from their partner. (So for participating in the decision-making exercise each participant received $4 from the first round plus whatever they took in the second round).

We expected that participants who had previously been assigned to BAM would make slower, more deliberate decisions than participants who had been assigned to the control conditions. We were also interested in testing whether actively trying to reduce automaticity during the decision-making exercise itself could attenuate the BAM-control difference in decision-making. So we randomized participants to four different versions of the task:

- A "no delay" condition, in which youths could say how much they wanted to take from their partner as soon as they wished after the partner's take amount was announced.
- A "distraction" condition, which is intended to get all youths to do part of what we believe BAM gets youth to initiate on their own – which is to slow down. In this condition, after round 1 the participants were told to first spend 30 seconds completing a word-search puzzle and to then state how much money they wished to take.

---

participate during their lunch period. Studies were conducted in available quiet spaces in the schools, such as hallways and empty classrooms.

[37] Because our study design mimics the psychology study in VanderMeer et al. (2015) and was run by a psychologist on our team it includes deception, which experimental economists would normally avoid. (All study subjects were de-briefed at the end of the experiment.)

- A "reflection" condition, where they were told to first take 30 seconds to rate their partner's action on a scale from -5 (extremely selfish) to +5 (extremely generous) before deciding how much to take from the partner.
- A "rumination" condition that gets youths to slow down but then, instead of reflecting and taking a different perspective on the event, they are given an exercise intended to promote unhelpful thinking (rumination). Specifically they were told to take 30 seconds to read over a list of adjectives and to circle the ones that represented their feelings in the moment, where the word list included terms like rude, unfriendly, mean, and unkind.

Unfortunately random assignment *across* conditions did not seem to work quite as well as we had hoped; the youths who wound up assigned to the distraction, reflection and rumination condition have somewhat higher levels of absences and disciplinary actions during the pre-program year relative to those assigned to the no-delay condition. However *within* conditions baseline characteristics are balanced for BAM vs. control youths. So while comparing across conditions within the control or BAM youths is challenging, we can still compare the difference-in-difference (how the BAM-control contrast differs across conditions). The basic idea as outlined in Table VIII is that our automaticity hypothesis implies that under all conditions BAM should get youths all on their own to slow down and reflect about what their optimal response would be. Conditions 2, 3 and 4 are intended to prompt different combinations of automaticity reduction from youths in the control group as well.

Table IX shows that BAM does indeed seem to get youths to slow down before they make a decision. We had the RAs working with participants subtly time how long it took participants to respond.[38] The variable is quite skewed and so we report results for the log of the time it takes youth to respond (see Appendix C for additional results). The second row shows

[38] We have time data for 303 of the 493 total youth who participated in our decision-making exercise because during the first phase of our field work the RAs were only timing youth who were randomized to condition 1.

that in condition 1 (where our automaticity theory makes a clean prediction that BAM should generate more "slowing down" compared to controls) the average control complier takes 1.1 seconds to decide, while the coefficient on BAM in our log-linear specification is 0.58, which implies a (statistically significant) 79% gain in the time youths take to decide. While our theory does not yield a clear prediction about whether we should see a BAM effect on slowing youths down in the other conditions, in practice we do see that BAM seems to introduce a moment of reflection across the other situational manipulations; the average effect pooled across all conditions is 0.32 (38%).

One potential concern one might have is the possibility that these results are somehow an artifact of response rates that are substantially less than 100%. To examine this possibility we take advantage of the fact that random assignment was carried out within schools, so each school is essentially its own experiment. In Figure III we plot the school-specific BAM impact on the log of how long it takes each youth to decide in our decision-making experiment against the response rate for our decision-making exercise in that school. That is, if (s) indexes schools, we estimate equations (2) and (3) above separately for each school to get π_{2s} and also calculate school-specific response rates R_s to our decision-making task. Figure III plots π_{2s} against R_s for schools in our study sample; the line in Figure III shows the slope from running the regression:

(4) $\pi_{2s} = \rho_0 + \rho_1 R_s + v_s.$

The figure is calculated pooling data from all 4 conditions in our decision-making experiment to maximize sample. The figure suggests the BAM effect on this candidate mechanism is if anything *larger* in schools with higher response rates.

Additional suggestive evidence to support the automaticity hypothesis comes from the fact that the sub-groups for which we see the largest impacts of BAM on outcomes are also where we see more pronounced BAM effects on slowing down. In the very first BAM RCT (study 1) there was only one detectable sub-group difference in effects on criminal behavior – youths who had not been arrested for a violent crime prior to randomization showed larger BAM effects on violent-crime arrests compared to those youth who had a prior violence arrest. We see the same pattern in our second BAM RCT (study 2) – the negative impact of BAM on arrests and positive impact on decision time are limited to youths with no prior violent-crime arrests. The same mirrored dose-response patterns hold if we use the approach from Kling, Liebman and Katz (2007) and use interactions of treatment assignment and school indicators as instrumental variables to estimate the effects of BAM sessions attended ("dose") on total arrests, as seen in Panel A of Figure IV, and automaticity, as in Panel B.

How much of the total BAM effect might be explained by automaticity? The coefficient on log decision time in a non-experimental regression against total arrests (using data just from our control group, controlling for block fixed effects and baseline covariates) is -0.07 (with a 95 percent confidence interval of -0.24 to 0.09). With an overall BAM effect on total arrests of -0.16, our point estimates taken at face value suggest that reduced automaticity could account for a decline in arrests of about $(-0.07\times0.58) = -0.044$ or one-quarter of the total effect (much larger than any of the other candidate mechanisms in Table VII). The 95% confidence intervals around our estimates imply automaticity could explain up to the entire BAM effect.

The other rows of the table show that by prompting all youths to do what we believe BAM gets youths to do on their own (slow down and reflect), the distraction, reflection, and even rumination conditions succeed in narrowing the BAM-control difference in the tendency to slow down and be less automatic when deciding by how much to retaliate.

Our automaticity theory does not make any clear prediction about whether BAM youths should actually retaliate less than controls in this iterated dictator game. BAM never tells youths not to fight or retaliate when provoked, since the program recognizes that in the neighborhoods in which these youths are growing up there are indeed circumstances in which fighting and an aggressive response may be (unfortunately) necessary and adaptive. The focus of the program instead is to get youth to slow down and reflect on what sort of response is most adaptive for the circumstance they are facing. Since our sample thought they were playing with other youths in their school and that they would play multiple rounds of the iterated dictator game, it may well be that they thought retaliation was rational.

Consistent with this focus of the program Table IX shows that we find no statistically significant BAM effect on the retaliation amount. Moreover the point estimates are much smaller as a share of the CCMs compared to what we see for the estimated BAM effect on the degree to which youth slow-down in their decision-making. This finding would also seem to argue against an explanation for why BAM works that emphasizes a more general or non-contingent shift towards more pro-social or "self-controlled" behavior.

VI. CONCLUSION

This paper presents results from three large-scale randomized experimental studies carried out in Chicago with economically disadvantaged male youth. While the exact sets of outcomes and patterns of results are not totally identical across the studies, all three experiments show sizable behavioral responses to fairly short-duration automaticity reducing interventions that get youths to slow down and behave less automatically in high-stakes situations. We also present some evidence suggesting that reduced automaticity may be a key mechanism.

These results as a whole suggest that automaticity can be an important explanation for disparities in outcomes. They also suggest—indirectly—two things. First, it suggests that non-adaptive automatic responses are one source of the problem that contributes to dropout and delinquency. Second, it suggests that lower-income youths do not respond to the need for reduced automaticity demanded in their neighborhoods by (sufficiently) reducing automaticity on their own. It is possible that young people would eventually develop reduced automaticity as a natural byproduct of aging, so that the interventions we study simply accelerate this process, or it could be that youths would never develop reduced automaticity absent the intervention. Our study cannot answer that question, but it is an important one for future research.

As with all randomized experiments, there is always some question about the degree to which these impacts generalize to other samples and settings. Because each of our three studies was carried out with large numbers of disadvantaged male youths from distressed areas of Chicago, they are closer to what medical researchers call "effectiveness trials" (testing interventions at scale) than to "efficacy trials" of a model (or "hothouse") program. Each intervention is manualized and so can in principle be scaled up further. The cost-effectiveness of these interventions would also seem to be favorable for scaling-up. Our best estimates on the cost side equal $1,178 and $2,000 per participant in study 1 and 2, respectively, and about $60 per juvenile detention spell in study 3.[39] The results imply a level of cost-effectiveness that is at least as favorable as almost any other crime-prevention intervention that has been studied seriously.[40]

The sizable impacts we observe across all three interventions stand in stark contrast to the generally dismal record of efforts to improve the long-term life outcomes of disadvantaged youths. As one juvenile detention staff member told us: "20 percent of our residents are criminals, they just need to be locked up. But the other 80 percent, I always tell them – if I could

---

[39] The difference in the cost per youth for BAM in study 1 and study 2 is driven by Youth Guidance's efforts to provide additional training and supervision of counselors to best accommodate implementation with fidelity to the model at large scale in Chicago. Namely, delivery of BAM in AY 2013-14 included for the first time the hiring of BAM supervisors (with a staff to supervisor ratio of 5 to 1), the development of infrastructure support and capacity building roles, a fidelity monitoring dashboard, and additional efforts to develop and manualize the curriculum.

[40] Costs reported in 2015 dollars. A common way to report impacts is in percentage terms. Under that metric study 1 yields a 27% reduction in arrests at a cost of $1,178 per participant, study 2 yields a 33% reduction in arrests at a cost of $2,000 per participant, while some of the most cost-effective interventions reported on in the review by Greenwood (2008) are multidimensional treatment foster care, which costs $7,570 per participant for a 22% reduction in crime, functional family therapy that costs $2,534 for a 16% reduction in crime, teen courts that cost $1,020 for a 11% reduction, adolescent diversion that costs $2,085 for 20% reduction, nurse family partnership that costs $799 for a 16% reduction, and pre-K that costs $646 for a 14% reduction. (We update costs from Greenwood's Table 2 to 2015 dollars and focus just on interventions that are listed as "proven" or "preferred" and for which cost figures are available; of the interventions on this list only pre-K is listed as "proven").

give them back just ten minutes of their lives, most of them wouldn't be here."[41] Our results suggest that it is possible to generate sizable changes in outcomes by helping disadvantaged youths recognize their automatic responses and make better decisions during those crucial ten-minute windows.

UNIVERSITY OF PENNSYLVANIA
UNIVERSITY OF CHICAGO
NORTHWESTERN UNIVERSITY AND NBER
UNIVERSITY OF CHICAGO AND NBER
HARVARD UNIVERSITY AND NBER
UNIVERSITY OF CHICAGO

[41] Personal communication, Darrien McKinney to Jens Ludwig, Sendhil Mullainathan, and Anuj Shah, 10/18/2012.

REFERENCES

Allensworth, Elaine Marie, and John Q Easton, *The on-track indicator as a predictor of high school graduation* (Consortium on Chicago School Research, University of Chicago, 2005).

Anderson, Elijah, *Code of the Street* (New York: Norton, 1999).

Anderson, ML, "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," Journal of the American Statistical Association, 103 (2008), 1481-1495.

Angrist, J.D., G.W. Imbens, and D.B. Rubin, "Identification of causal effects using instrumental variables," Journal of the American Statistical Association, 91 (1996), 444-455.

Beck, Judith S, *Cognitive Therapy: Basics and Beyond* (The Guilford Press, 2011).

Benjamini, Yoav, Abba M Krieger, and Daniel Yekutieli, "Adaptive linear step-up procedures that control the false discovery rate," Biometrika, 93 (2006), 491-507.

Bloom, H.S., L.L. Orr, S.H. Bell, G. Cave, F. Doolittle, W. Lin, and J.M. Bos, "The benefits and costs of JTPA Title II-A programs: Key findings from the National Job Training Partnership Act study," Journal of Human Resources, (1997), 549-576.

Bloom, Howard S., "Accounting for No-shows in Experimental Evaluation Designs," Evaluation review, 8 (1984), 225-246.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller, "Bootstrap-based improvements for inference with clustered errors," Review of Economics and Statistics, 90 (2008), 414-427.

Survey Methodology (http://help.ccsrsurvey.uchicago.edu/customer/portal/articles/94362-survey-methodology,

Cohen, Jonathan D, Samuel M McClure, and J Yu Angela, "Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration," Philosophical Transactions of the Royal Society B: Biological Sciences, 362 (2007), 933-942.

CPD, "Annual Report," (Chicago, 2011).

---, "Chicago Murder Analysis," (Chicago, 2011).

Cunningham, Tom, "Hierarchical Aggregation of Information and Decision-Making," (2015).

Deming, David, "Better Schools, Less Crime?," Quarterly Journal of Economics, 126 (2011), 2063-2115.

Duckworth, Angela L, Christopher Peterson, Michael D Matthews, and Dennis R Kelly, "Grit: perseverance and passion for long-term goals," Journal of personality and social psychology, 92 (2007), 1087.

Duckworth, Angela L, and Martin EP Seligman, "Self-discipline outdoes IQ in predicting academic performance of adolescents," Psychological science, 16 (2005), 939-944.

Duckworth, Angela Lee, and Patrick D Quinn, "Development and validation of the Short Grit Scale (GRIT–S)," Journal of personality assessment, 91 (2009), 166-174.

Ellis, Albert, "Outcome of employing three techniques of psychotherapy," Journal of Clinical Psychology, (1957).

Ellis, Albert, and Robert A Harper, *A new guide to rational living* (Prentice-Hall, 1975).

Evans, Gary W, Carrie Gonnella, Lyscha A Marcynyszyn, Lauren Gentile, and Nicholas Salpekar, "The role of chaos in poverty and children's socioemotional adjustment," Psychological science, 16 (2005), 560-565.

Evans, Gary W, and Jennifer Rosenbaum, "Self-regulation and the income-achievement gap," Early Childhood Research Quarterly, 23 (2008), 504-514.

Evans, W.N., and E.G. Owens, "COPS and Crime," Journal of Public Economics, 91 (2007), 181-201.

Greenwood, Peter, "Prevention and intervention programs for juvenile offenders," The future of Children, 18 (2008), 185-210.

Grossman, Jean Baldwin, and Joseph P Tierney, "Does mentoring work? An impact study of the Big Brothers Big Sisters program," Evaluation review, 22 (1998), 403-426.

Heckman, James J, "Skill formation and the economics of investing in disadvantaged children," Science, 312 (2006), 1900-1902.

Heller, Sara, Harold A Pollack, Roseanna Ander, and Jens Ludwig, "Preventing youth violence and dropout: a randomized field experiment," (Cambridge, MA: National Bureau of Economic Research, 2013).

Herrera, Carla, Jean Baldwin Grossman, Tina J Kauh, and Jennifer McMaken, "Mentoring in Schools: An Impact Study of Big Brothers Big Sisters School‚ÄêBased Mentoring," Child Development, 82 (2011), 346-361.

Kaelbling, Leslie Pack, Michael L Littman, and Andrew W Moore, "Reinforcement learning: A survey," Journal of artificial intelligence research, (1996), 237-285.

Kahneman, Daniel, *Thinking, fast and slow* (Macmillan, 2011).

Katz, L.F., J.R. Kling, and J.B. Liebman, "Moving to opportunity in Boston: Early results of a randomized mobility experiment," Quarterly Journal of Economics, 116 (2001), 607-654.

Kleinke, Chris L, "Gaze and eye contact: a research review," Psychological bulletin, 100 (1986), 78.

Kling, J. R., J. Ludwig, and L. F. Katz, "Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment," Quarterly Journal of Economics, 120 (2005), 87-130.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental analysis of neighborhood effects," Econometrica, 75 (2007), 83-119.

Lochner, Lance, and Enrico Moretti, "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," The American Economic Review, 94 (2004), 155-189.

Maultsby, Maxie C., *Help yourself to happiness through rational self counseling* (New York: Institute for Rational Emotive Behavior Therapy, 1975).

---, *Rational Behavior Therapy* (Appleton, WI: Rational Self-Help/FACT, 1990).

Moffitt, Terrie E, Louise Arseneault, Daniel Belsky, Nigel Dickson, Robert J Hancox, HonaLee Harrington, Renate Houts, Richie Poulton, Brent W Roberts, and Stephen Ross, "A gradient of childhood self-control predicts health, wealth, and public safety," Proceedings of the National Academy of Sciences, (2011), 201010076.

Mullainathan, Sendhil, Joshua Schwartzstein, and Andrei Shleifer, "Coarse Thinking and Persuasion*," The Quarterly Journal of Economics, 123 (2008), 577-619.

Orfield, Gary, Erica Frankenberg, Jongyeon Ee, and John Kuscera, "Brown at 60: Great progress, a long retreat and an uncertain future," (The Civil Rights Project, 2014).

Papachristos, Andrew V, "Murder by Structure: Dominance Relations and the Social Structure of Gang Homicide1," American Journal of Sociology, 115 (2009), 74-128.

Rush, Augustus J, Aaron T Beck, Maria Kovacs, and Steven Hollon, "Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed outpatients," Cognitive therapy and research, 1 (1977), 17-37.

Toribio, Almeida Jacqueline, and Barbara E Bullock, *The Cambridge handbook of linguistic code-switching* (Cambridge University Press, 2012).

Tough, Paul, *How children succeed* (Random House, 2013).

VanderMeer, James, Christine Hosey, Nicholas Epley, and Boaz Keysar, "Title," University of Chicago, Unpublished Manuscript.

Weiner, D.A., B. Lutz, and J. Ludwig, "The effects of school desegregation on crime," (National Bureau of Economic Research Cambridge, Mass., USA, 2009).

Westfall, P.H., and S.S. Young, *Resampling-based multiple testing: Examples and methods for p-value adjustment* (Wiley-Interscience, 1993).

Figure I. Study 1 and 2 Schools by Neighborhood Homicide Rate

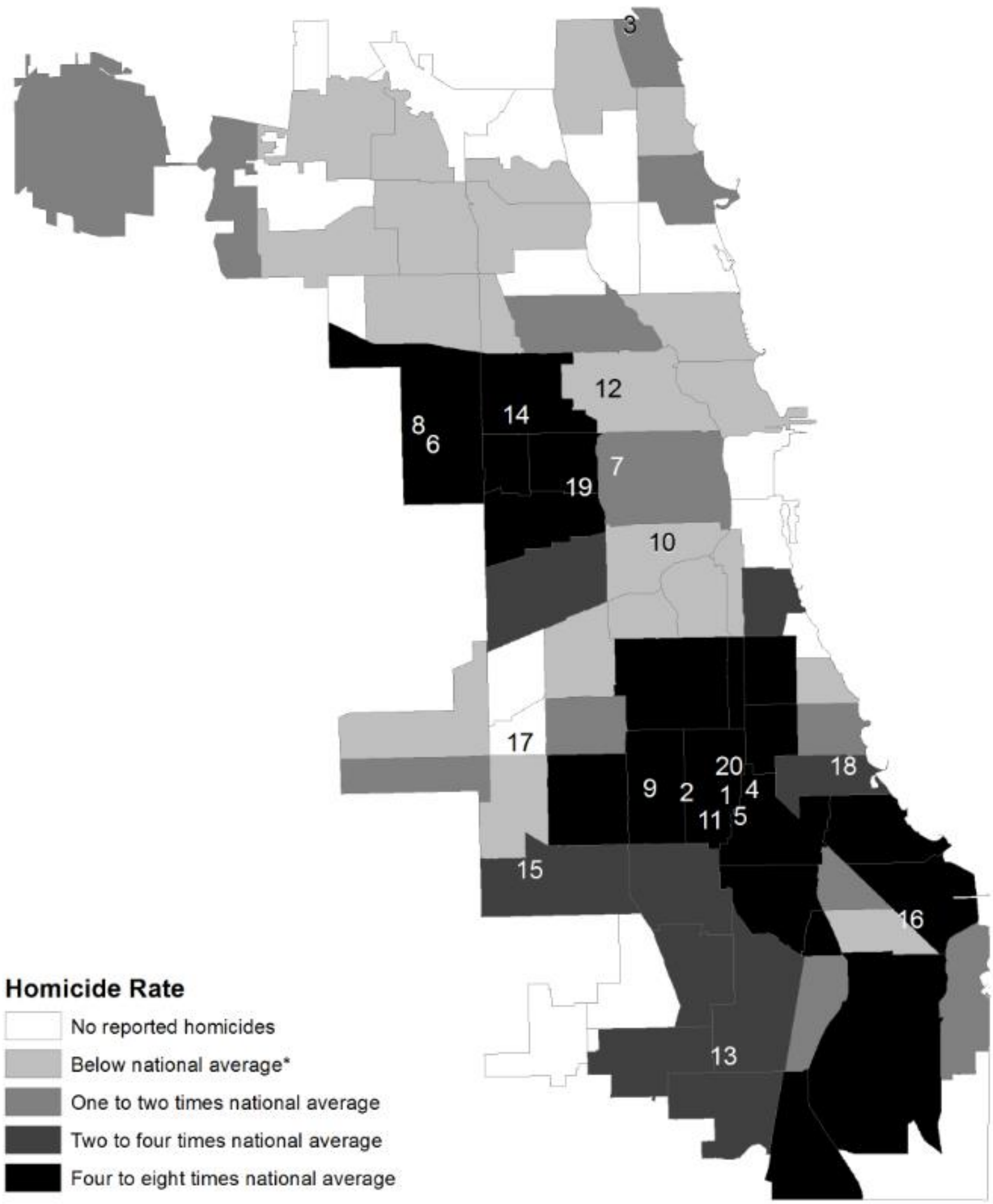


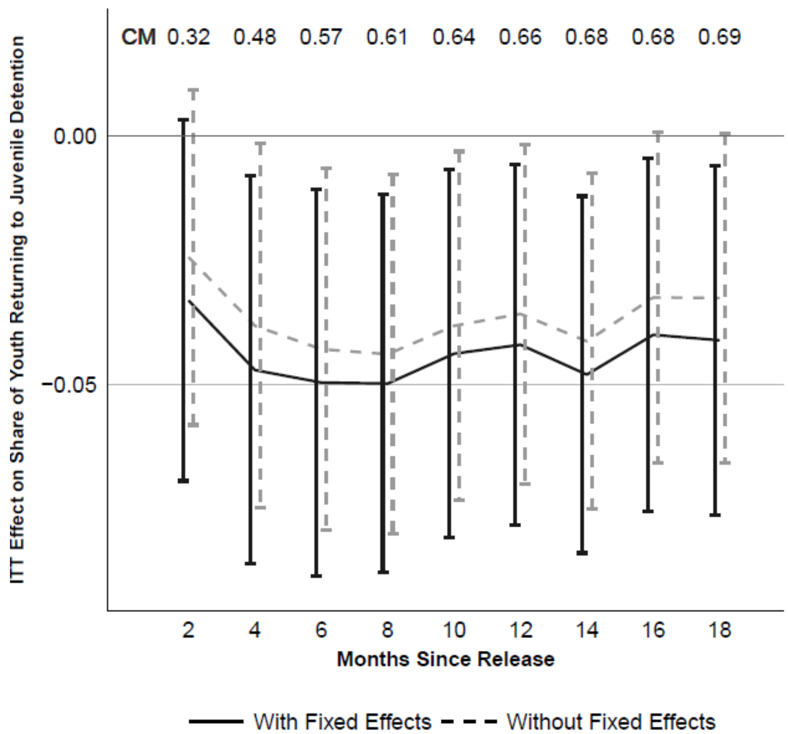
Figure I. Legend

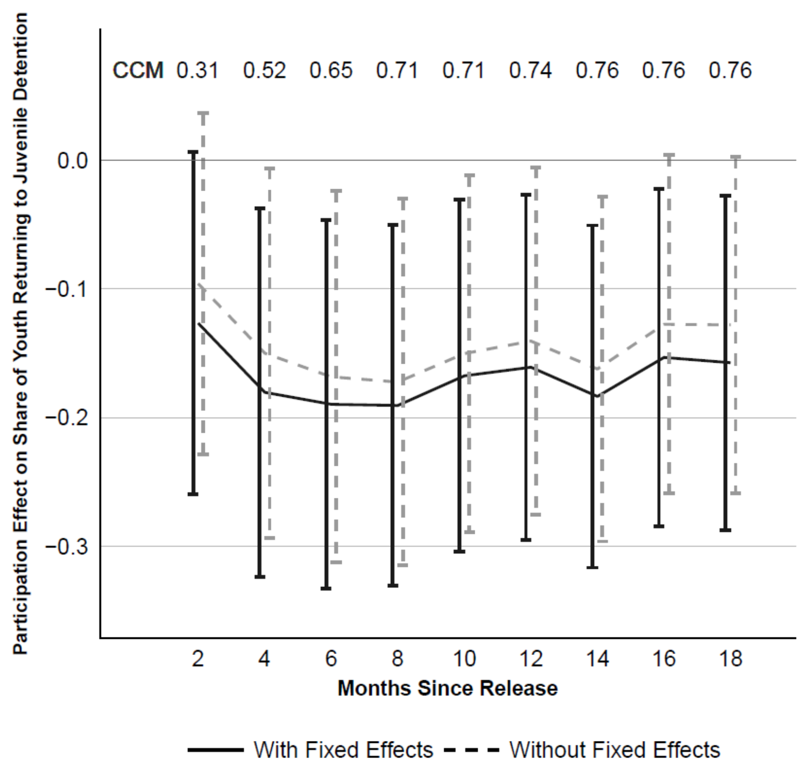
1 Banneker ES†
2 Bass ES†
3 Jordan ES†
4 Parker Community Academy ES†
5 Yale ES†
6 Austin Polytechnic HS†
7 Crane HS†
8 Douglass HS†
9 Harper HS†
10 Juarez HS†
11 Robeson HS†
12 Clemente HS†§
13 Fenger HS†§
14 Orr HS†§
15 Bogan HS§
16 Bowen HS§
17 Hancock HS§
18 Hyde Park HS§
19 Manley HS§
20 Noble Street Charter HS – Johnson§

† Study 1 School
§ Study 2 School

*Rate for cities with population over 1,000,000 is 7.5 per 100,000 citizens
Sources: City of Chicago; US Census Bureau; FBI Uniform Crime Report 2013

Figure II. Effect of Treatment on Re-Admission, Study 3 (Juvenile Detention)





Notes: n = 2693. CM is control mean and CCM indicates control complier mean. Error bars represent 95% confidence intervals.

Figure III. School-Specific Treatment Effect on Decision-Making with Subsample of Study 2 Youth (BAM 2013-14 Cohort)

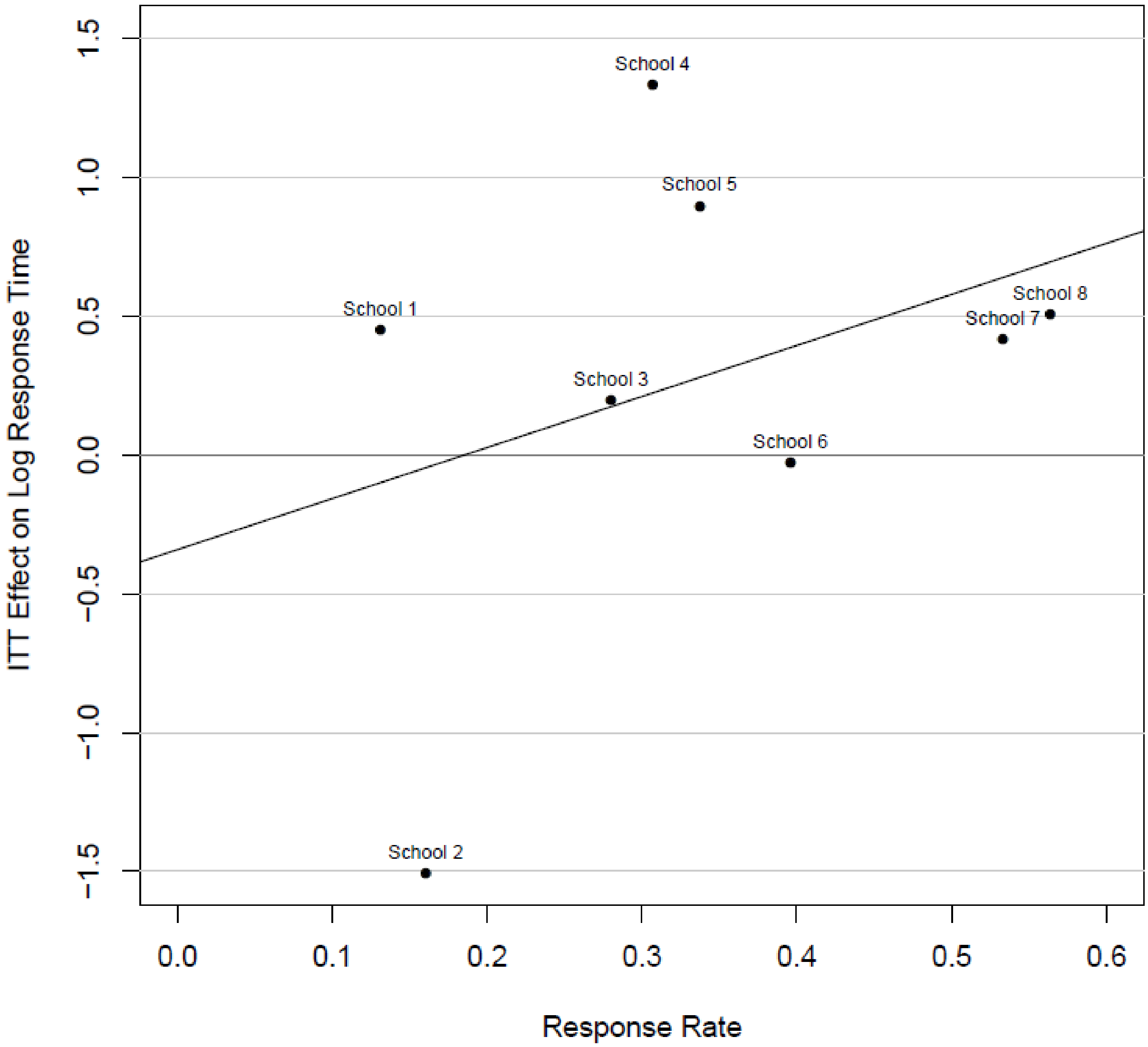


Figure IV. Treatment Effect on Arrests and Automaticity (Decision-Making Time) with Subsample of Study 2 Youth (BAM 2013-14 Cohort)

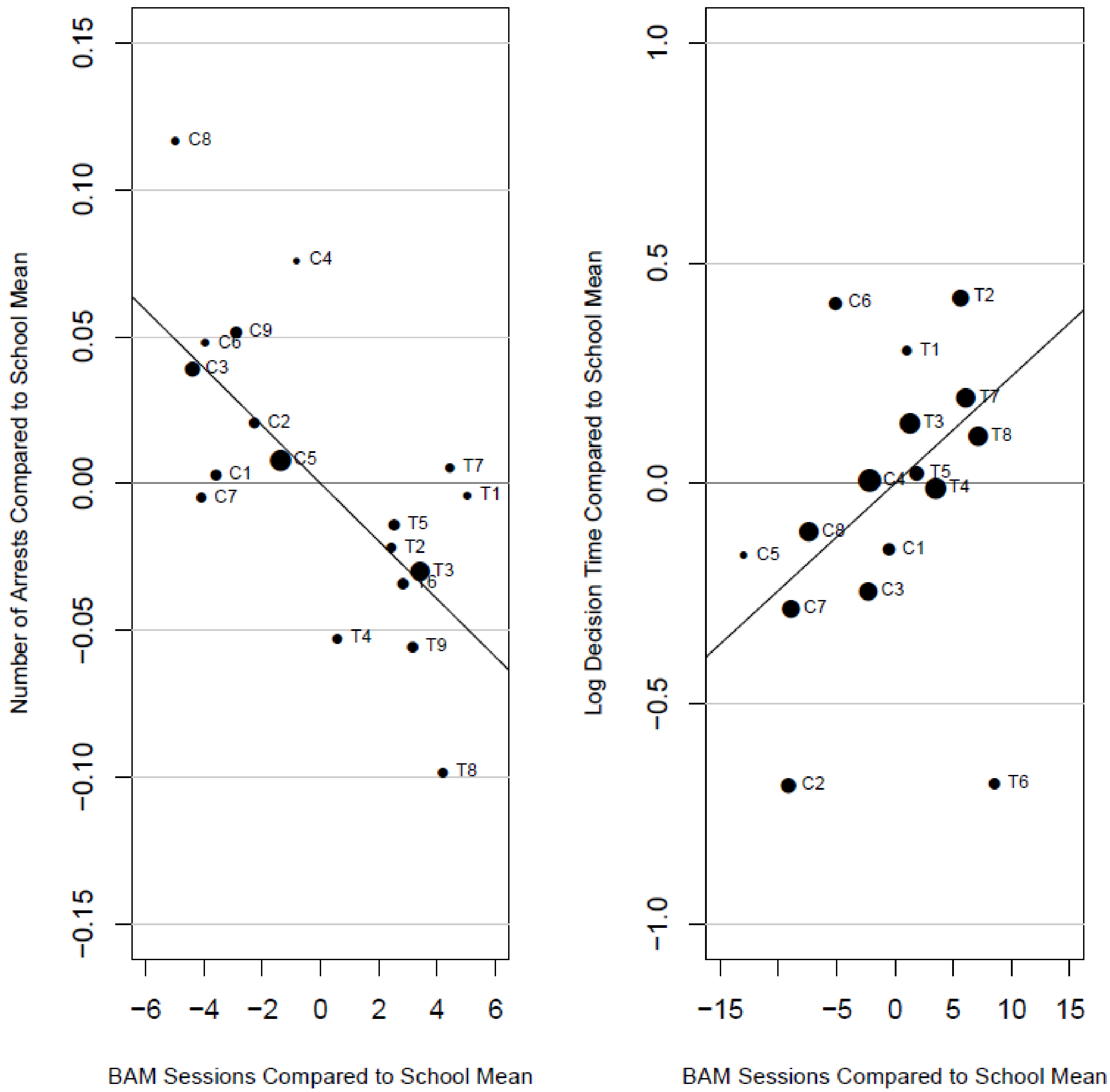


Table I. Socio-Economic Gradient in Retaliation (Fighting) versus Instigation (Bullying)

	White Males (n = 2844)	Black Males (n = 1494)	Racial Difference (p-Value)
Physical fight in the last 12 months?	27.11%	37.54%	< 0.01
	[0.012]	[0.023]	
Injured in fight?	2.74%	4.67%	0.047
	[0.003]	[0.008]	
Bullied on school property in the last 12 months?	16.20%	10.16%	< 0.01
	[0.011]	[0.009]	

Notes: Data source is the CDC Youth Risk Behavior Survey from 2013. Analysis sample is limited to male youth; n = 6950. No questions about family income or parent education were asked during 2013 wave, so race is the closest proxy for socio-economic status in the dataset.

Table II. Descriptive Statistics for Baseline Characteristics of Study 1 Sample (BAM 2009-10 Cohort)

	All Randomized Youth	Assigned to Control	Assigned to Treatment	Treatment-Control Difference (p-Value)
n Students	2,740	1,267	1,473	
Demographics				
Age	15.60	15.70	15.51	0.55
Black	70.28%	71.96%	68.84%	0.32
Hispanic	29.28%	27.65%	30.69%	0.38
Schooling				
Grade	9.35	9.42	9.29	0.24
Old for Grade	53.10%	55.00%	51.00%	0.43
GPA	1.71	1.68	1.73	0.88
Total Days Present	131.89	129.86	133.60	0.34
IEP	20.80%	21.23%	20.43%	0.86
Arrests				
Ever Arrested	35.69%	36.94%	34.62%	0.93
Number of Arrests for:				
Violent Crime	0.35	0.35	0.35	0.56
Property Crime	0.20	0.21	0.19	0.96
Drug Crime	0.17	0.17	0.18	0.35
Other Crime	0.46	0.45	0.47	0.41

Notes: Data source is Chicago Public Schools administrative data for AY 2008-9, as well as Illinois State Police arrest records for pre-random assignment period. The CPS school year during this period was 170 days; GPA is measured on a 0-4 scale. IEP indicates the presence of an individualized education plan. Final column reports p-values from comparing treatment versus control mean, controlling for block fixed effects. Joint significance test for equality of all baseline characteristics using only non-missing data (n=2579), F(18,2543)=1.02, p=0.431. * p<0.10, ** p<0.05, *** p<0.01.

Table III. Descriptive Statistics for Baseline Characteristics of Study 2 Sample (BAM 2013-14 Cohort)

	All Randomized Youth	Assigned to Control	Assigned to Treatment	Treatment-Control Difference (p-Value)
N Students	2,064	1,048	1,016	
Demographics				
Age	14.78	14.75	14.81	0.642
Black	68.90%	69.56%	68.21%	0.709
Hispanic	28.73%	27.48%	30.02%	0.466
Schooling				
Grade	9.43	9.41	9.46	0.715
Old for Grade	34.74%	34.64%	34.84%	0.654
GPA	2.13	2.11	2.15	0.293
Total Days Present	148.92	148.12	149.75	0.032**
Learning Disabled	16.57%	16.70%	16.44%	0.487
Arrests				
Ever Arrested	22.77%	22.61%	22.93%	0.699
Number of Arrests for:				
Violent Crime	0.19	0.19	0.19	0.287
Property Crime	0.15	0.15	0.15	0.963
Drug Crime	0.13	0.11	0.14	0.930
Other Crime	0.31	0.29	0.32	0.973

Notes: Data sources are CPS data for AY 2012-13 and Chicago Police Department arrest data. The CPS school year during this period was 180 days; GPA is measured on a 0-4 scale. Final column reports p-values from comparing treatment versus control mean, controlling for block fixed effects. Joint significance test for equality of all baseline characteristics using only non-missing data (n = 1756); $F(13,1752) = 1.21$, $p > F = 0.2657$. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table IV. Effects of Program Participation on Youth Outcomes, Study 1 (BAM 2009-10 Cohort)

	Control Mean	Intention to Treat (ITT)	Effect of Participation (IV)	Control Complier Mean (CCM)
		Year 1		
School Engagement Index	0	0.0585***	0.1403***	0.218
		[0.0216]	[0.0511]	
Arrests Per Youth Per Year:				
All Offenses	0.699	-0.0740	-0.1775	0.663
		[0.0456]	[0.1085]	
Violent Offenses	0.167	-0.0336**	-0.0806**	0.184
		[0.0165]	[0.0394]	
Property Offenses	0.077	0.0050	0.0120	0.066
		[0.0127]	[0.0303]	
Drug Offenses	0.151	0.0026	0.0062	0.094
		[0.0177]	[0.0420]	
Other Offenses	0.305	-0.0480*	-0.1151*	0.320
		[0.0271]	[0.0646]	
		Year 2		
School Engagement Index	0	0.0786***	0.1887***	0.039
		[0.0216]	[0.0515]	
Arrests Per Youth Per Year:				
All Offenses	0.595	-0.0635	-0.1524	0.604
		[0.0418]	[0.0995]	
Violent Offenses	0.110	-0.0005	-0.0013	0.095
		[0.0142]	[0.0337]	
Property Offenses	0.057	-0.0032	-0.0076	0.052
		[0.0101]	[0.0241]	
Drug Offenses	0.164	-0.0181	-0.0435	0.170
		[0.0194]	[0.0462]	
Other Offenses	0.264	-0.0417	-0.0999	0.288
		[0.0258]	[0.0614]	

Notes: Baseline covariates and randomization block fixed effects included in all model specifications (see text). Heteroskedasticity-robust standard errors in brackets. School engagement index is equal to an (unweighted) average of days present, GPA, and enrollment status at end of school year, all normalized to Z-score form using control group's distribution. * p<0.10, ** p<0.05, *** p<0.01.

Table V. Effects of Program Participation on Youth Outcomes, Study 2 (BAM 2013-14 Cohort)

	Control Mean	Intention to Treat (ITT)	Effect of Participation (IV)	Control Complier Mean (CCM)
School Engagement Index	0.000	0.0074	0.0147	0.204
		[0.0258]	[0.0508]	
Arrests Per Youth Per Year:				
All Offenses	0.499	-0.0793*	-0.1587*	0.519
		[0.0439]	[0.0865]	
Violent Offenses	0.106	-0.0141	-0.0282	0.099
		[0.0147]	[0.0290]	
Property Offenses	0.062	-0.0071	-0.0142	0.063
		[0.0121]	[0.0238]	
Drug Offenses	0.110	-0.0181	-0.0362	0.142
		[0.0205]	[0.0404]	
Other Offenses	0.221	-0.0401	-0.0801	0.214
		[0.0252]	[0.0497]	

Notes: Baseline covariates and randomization block fixed effects included in all model specifications (see text). Heteroskedasticity-robust standard errors in brackets. School engagement index is equal to an (unweighted) average of days present, GPA, and enrollment status at end of school year, all normalized to Z-score form using control group's distribution. * p<0.10, ** p<0.05, *** p<0.01.

Table VI. Descriptive Statistics for Baseline Characteristics of Balanced Study 3 Sample (Juvenile Detention) With Complete 18-Month Follow-Up Data

	All Randomized Youth	Assigned to Control	Assigned to Treatment Group	p-Value
n Students	2,693	1,322	1,371	
Demographics				
Age	16.14	16.13	16.15	0.69
Black	0.84	0.83	0.84	0.52
White	0.13	0.14	0.13	0.39
Hispanic	0.03	0.03	0.03	0.84
Other Race	0.01	0.01	0.01	0.84
Spell Number	3.27	3.35	3.19	0.11
Reason for Juvenile Detention Admission				
Violent Crime Arrest	0.18	0.17	0.18	0.22
Property Crime Arrest	0.10	0.10	0.09	0.31
Drug Crime Arrest	0.08	0.08	0.07	0.38
Other Crime Arrest	0.35	0.35	0.35	0.99
Non-Arrest Violation	0.30	0.30	0.30	0.89
Arrest History				
Number of Total Arrests	8.05	8.25	7.85	0.09
Violent	2.05	2.11	2.00	0.17
Property	1.52	1.55	1.49	0.39
Drug	1.30	1.34	1.25	0.29
Motor Vehicle	0.03	0.03	0.03	0.54
Other	3.15	3.22	3.08	0.27
Neighborhood Characteristics				
Percent With At Least High School Degree	72.77%	72.80%	72.75%	0.95
Percent Black	68.62%	69.13%	68.12%	0.51
Percent Hispanic	18.08%	17.71%	18.44%	0.53
Percent Below Poverty Line	34.56%	34.54%	34.58%	0.95
Percent Unemployed	18.58%	18.59%	18.58%	0.99

Notes: Spell number counts how many juvenile detention admissions the youth had up to and including the focal spell. Neighborhood characteristics come from geocoding address information and linking to tract level data from the American Community Survey. Final column reports p-values from comparing treatment versus control mean, controlling for tract level. Joint significance test for equality of all baseline characteristics using only non-missing data (n=2643), F(25,1835)=0.79, p > F = 0.7631. * p<0.10, ** p<0.05, *** p<0.01.

Table VII. Test of Candidate Mediating Mechanisms, Study 1 Sample (BAM 2009-10 Cohort)

	Participation	Outcomes			
		School Engagement 2009/2010		Violent Crimes 2009/2010	
Candidate mediating measure (Z-score form, normalized to control group distribution)	Effect of BAM participation on candidate mediator	Non-experimental association of mediator with school engagment (n = 428)	Share of total BAM participation effect on school engagement explained by this mechanism	Non-experimental association of mediator with violent crimes (n = 446)	Share of total BAM participation effect on violent crimes explained by this mechanism
Social Capital /Mentoring (n = 999)	0.0186 [0.1237]	-0.0360* [0.0217]	-0.48% [-14.61%, 12.53%]	-0.0136 [0.0132]	0.31% [-10.96%, 12.78%]
Perceived Returns to Schooling (n = 794)	-0.0219 [0.1521]	0.0148 [0.0224]	-0.23% [-13.39%, 11.56%]	-0.0232 [0.0211]	-0.63% [-25.63%, 22.12%]
Social skills (n = 1081)	0.1305 [0.1192]	0.0022 [0.0168]	0.20% [-7.98%, 9.12%]	0.002 [0.0148]	-0.32% [-14.01%, 12.20%]
Grit (n = 1074)	0.1278 [0.1214]	0.0522*** [0.0192]	4.75% [-7.05%, 23.42%]	-0.0045 [0.0170]	0.71% [-13.08%, 17.16%]

Notes: The following are the specific questions for each category. 1. Social Capital/Mentoring: have at least one teacher or adult in school I can talk to if I have a problem. 2. Schooling: classes are useful preparation for future, high school teaches valuable skills, working hard in school matters for future work force, what we learn in class is useful for future. 3. Social skills: I can always find a way to help end arguments, I listen carefully to what other people say about me, I'm very good at working with other students, I'm good at helping people. 4. Grit: I finish whatever I begin, I am a hard worker, I continue steadily towards my goals, I don't give up easily.

First column of results presents coefficient from IV analysis of BAM participation effect on the candidate mediating mechanism measure listed in the row label at left, which comes from CPS survey of youth carried out in 2011 (see text). Second column presents the results of a non-experimental regression of the candidate mediator against the school engagement index (outcome), using just data from the control group, controlling for the same baseline covariates and block fixed effects as in the main analyses. Third column multiplies point estimate from column 1 by point estimate in column 2 and then divides by the estimated IV effect of BAM participation on that outcome taken from Table 3. Confidence intervals come from multiplying 95% CI in column 1 by 95% CI in column 2. Remaining columns of the table are constructed analogously. * p<0.10, ** p<0.05, *** p<0.01.

Table VIII. Design of Decision-Making Experiment Carried Out With Sub-Sample of Study 2 Youth (BAM 2013-14 Cohort)

	Randomized to BAM		Randomized to control	
	Slow down?	Reflect?	Slow down?	Reflect?
Condition 1 no delay	Yes	Yes		
Condition 2 delay (distraction) - "partial CBT" manipulation	Yes	Yes	Yes	
Condition 3 delay plus reflection - "CBT" manipulation	Yes	Yes	Yes	Yes
Condition 4 delay plus rumination - "anti-CBT" manipulation	Yes		Yes	

Table IX. Effect of BAM Participation on Automaticity (Decision-Making Time) and Retaliation with Subsample of Study 2 Youth (BAM 2013-14 Cohort)

	Log time to make decisions (seconds)		Take amount ($)	
	Control Complier Mean	Effect of BAM participation	Control Complier Mean	Effect of BAM participation
All Conditions Pooled	0.974	0.3213**	7.093	0.2060
		[0.1313]		[0.2205]
Condition 1	1.114	0.5841**	7.141	-0.3596
no delay (n=117)		[0.2619]		[0.4354]
Condition 2	0.861	0.1063	6.706	0.9185**
delay (distraction) -		[0.2250]		[0.4115]
"partial CBT" manipulation (n=60)				
Condition 3	0.986	0.2194	7.055	0.2694
delay plus reflection -		[0.2351]		[0.4800]
"CBT" manipulation (n=63)				
Condition 4	0.691	0.2901*	7.515	-0.1427
delay plus rumination -		[0.2336]		[0.4288]
"anti-CBT" manipulation (n=62)				

Notes: Sample of n=302 (log time) and n= 409 (take amount) youth that participated in decision-making experiment. Analysis presents results of IV analysis of BAM participation effect on log of time (in seconds) to make decision about retaliation and take amount in iterated dictator game. Regression specification includes baseline covariates and block fixed effects as in main analyses.
* p<0.10, ** p<0.05, *** p<0.01.