Title Training Police for Procedural Justice

About the authors

Wesley G. Skogan   [corresponding author]
Institute for Policy Research, Northwestern University
2040 Sheridan Road, Evanston IL USA
+1.847.491.3395
e-mail: skogan@northwestern.edu

Maarten Van Craen
Leuven Institute of Criminology (LINC)
University of Leuven
Hooverplein 10, 3000 Leuven, Belgium
0032 16 32 93 93
e-mail: maarten.vancraen@law.kuleuven.be

Cari Hennessy
Institute for Policy Research, Northwestern University
2040 Sheridan Road, Evanston IL USA
+1.847.491.3395
email: cari.hennessy@northwestern.edu

Training Police for Procedural Justice

Abstract

Objectives: This paper reports the findings of an evaluation of a police training program on the principles of procedural justice. This training was part of a larger organizational change strategy aimed at improving the relationship between the police and the public in Chicago.
Methods: The paper reports on the findings of two studies. The short-term effects study was a quasi-experimental test of the immediate effectiveness of the training conducted at the police academy. A longer-term effects study examined the subsequent views of trainees and a comparison group, officers who had not yet been to training. Statistical controls were used to increase confidence in the findings of the second study, which was based on observational data.
Results: In the short term, training increased officer support for all of the procedural justice dimensions included in the experiment. Post-training, officers were more likely to endorse the importance of giving citizens a voice, granting them dignity and respect, demonstrating neutrality, and (with the least enthusiasm) trusting them to do the right thing. All of the effects of training were strong, with standardized effect sizes ranging from 1.2 to 1.6. Longer-term, officers who had attended the procedural justice workshop continued to be more supportive of three of the four procedural justice principles introduced in training; the effect of training on trust was not statistically significant.
Conclusions: There has been little systematic research on police training. This paper concludes that it can play a role in improving police-community relations. It also presents a discussion of some of the limitations of a training-based organizational change strategy.
Keywords: quasi-experiment, survey, police training, procedural justice, neutrality, respect, voice, trust

Training Police for Procedural Justice

A very large body of research has demonstrated the importance of procedural justice in shaping the legitimacy of police in the eyes of the public. We know that the public is concerned that police decisions are made fairly and evenhandedly, that citizens are treated respectfully and given a chance to voice their views, and that officers are thought to abide by the rules that govern their behavior (for reports of research and reviews see Jackson et al. 2013; Mazerolle et al. 2013a; Murphy et al. 2008; Tyler and Jackson 2013). However, there has been almost no research at all regarding how the police can be encouraged to incorporate the principles of procedural justice in their routine interactions with the public. This project examined an initiative by the Chicago Police Department to reshape its relationship with the community, by training officers in the principles of procedural justice. Training began in mid-2012 and concluded in September 2013. In total, about 8,700 serving officers, 230 new recruits, many of the department's civilian employees, and some community members involved in policing issues were trained.

The research team conducted two studies evaluating this effort. The shorter-term effects study was a quasi-experimental test of the immediate effectiveness of the training conducted at the academy. A longer-term effects study followed, with interviews of trainees and a comparison group, officers who had not yet been to training. Statistical controls were used to increase confidence in the findings of the long-term effects study. In addition, we monitored training sessions and interviewed trainers, trainees and program developers.

In the shorter term, training increased officer support for all of the procedural justice dimensions included in the experiment. Post-training they were more likely to endorse the importance of giving citizens a voice, granting them dignity and respect, demonstrating neutrality, and (with the least enthusiasm) trusting them to do the right thing. All of the effects of training were strong, based on standardized effect size measures. Longer-term, officers who had attended the procedural justice workshop continued to be more supportive of three of the four procedural justice principles introduced in training, compared to officers who had not yet been sent to training.

Background on police training

We know virtually nothing about the short or long-term effects associated with police training of any type. A committee established by the National Research Council to evaluate the state of policing in the United States found that there were "scarcely more than a handful of studies" on the effects of training, and that police training and education were being offered without scientific evidence of their likely effects. The panel concluded "[T]he committee cannot overstate the importance of developing a comprehensive and scientifically rigorous program to learn what is and is not effective in the education and training of police officers" (Skogan and Frydl 2004: 154). There has not been much progress in the decade since that report appeared. Police departments themselves often distribute post-training questionnaires that evaluate practical aspects of the class and the capabilities of the trainers, and they occasionally report "before-and-after" studies of a training class. External researchers have done a bit better, but the volume, quality and generality of evaluation work on police training has been low. Wheller and

Morris' (2010) review of research on training for the UK National Police Improvement Agency casts a very wide net, incorporating systematic studies of training for "professionals" of all kinds, but did not find much outside the domain of clinical training for health professionals. They located no published experimental studies of police training; however, they missed Rosenbaum's (1987) experimental study of training police recruits in how to better handle crime victims. He randomized new recruits into training or an alternative control activity and later assessed their attitudes and behavioral intentions.

Virtually no research of any flavor has been done on procedural justice training, despite this being a necessary precursor to turning the theory into practice. The Mazerolle et al. (2013a) systematic review of procedural justice and police legitimacy alludes to the important role that training could play, but did not report a single training study. However, since that time Wheller et al. (2013) released the findings of an experimental study of procedural justice training. Officers from the Greater Manchester Police Service were randomly assigned to treatment or control groups in order to determine the impact of training on the perceived quality of interactions between the police and crime victims. The training program incorporated elements of procedural justice theory. The evaluation identified positive shifts on four of eight police attitudinal outcomes, and positive effects on trained officers' scores in role-playing exercises. The perceptions of crime victims who later were served by trained and control-group officers also differed on some measures. In addition, a randomized trial involving police breath-testing roadblocks conducted by Mazerolle et al. (2013b) can be viewed as a test of the ability of officers to follow scripts that activate perceptions of procedural justice during encounters with the public, when they are briefed on how to do so.

Chicago's training program

The training program was developed internally, by sworn members of the department's training staff. The impetus for this effort came from the top of the organization, when a new chief of police arrived with a commitment to imbue his department with the principles of procedural justice. At his initiative, members of the training staff worked with academics (at another university) to develop an outline for a training program. Based on this preliminary plan, the staff identified concepts and findings from the research literature that they could translate into training units. They used the Internet to locate video clips and images that would illustrate key points. Their goal was to present procedural justice principles to officers as tactics that would encourage the public to recognize the police as a legitimate source of authority, resulting in improved officer safety, more compliance with their instructions, and greater cooperation from the public.

The training involved officers from every unit of the department. As it progressed, the units sent just a few officers to training each day, so participants trickled into the academy rather than arriving in related batches. Classes of at most 25 officers met with rotating teams of three trainers for day-long sessions. Unusually for this organization, the classes mixed employees of various ranks, although they attended in civilian clothing to soften the impact of this heterogeneity. Multiple classes were conducted on two shifts; morning watch officers began at 7:00am, and afternoon and evening watch officers at 3:30pm, Mondays through Thursdays. The trainers came from diverse backgrounds, and many were working in the field when they were

recruited for the job. They had the credible street experience that they needed to convince officers that they knew what they were talking about, and to overcome resistance from those suspecting that the trainers had grown out of touch while working at the academy.

The training modules

The training day was divided into five modules. Each featured lectures accompanied by PowerPoint © slides, video clips, and group exercises. Module 1 introduced the ideas of procedural justice and legitimacy, and how they intersect with the goals of policing. In the ensuing slides and discussion (which the trainers noted were based on research), procedural justice was defined as treating citizens fairly and with respect as human beings. Legitimacy was described as what the police receive in return: a public that views them as entitled to exercise authority in order to maintain order, manage conflicts, and resolve problems in the community. Module 2 featured a discussion of cynicism. This was added to the curriculum in response to issues that arose during the first month of training. A discussion topic was whether cynicism is actually "realism," or if it exaggerates the actual situation facing officers and serves to drive a wedge between them and the public. During this module, half of the table groups discussed and recorded their views on the question, "What does the community expect from police officers?" The other half of the table groups considered the question, "What do police officers expect from the community?" Module 3 included a more in-depth discussion of procedural justice, and featured research findings from Chicago as well as other cities that supported trainers' claims about the importance of the components of procedural justice.

Most slides also stressed one of the major selling points of the training day, which was that following these principles would ultimately benefit police officers by increasing citizen cooperation, encouraging the public to comply with police instructions, and maximizing officer safety. Module 4 began with a discussion of race and policing in historical context, both in Chicago and around the country. In interviews, the trainers indicated that discussing the role of race in policing proved to be the second hardest "sell" in the curriculum, following only having to convince officers that "legitimacy" should be one of their personal concerns. One trainer described this experience:

> *On the [course] evaluation, somebody said the only thing that I did was show white officers beating up on black people. And I thought that I failed, because if that's the only thing you remember in 8 hours, then I did something wrong.*

Near the end of Module 4 the trainers introduced the idea that there can be a "balance of trade" between police and the community. They argued that good encounters with the police are "deposits" and bad encounters are "withdrawals" from the account balance that they have developed with the community. The final module was a wrapping-up exercise. Participants viewed and discussed a video of a drug arrest in which the officer calmly, politely and successfully took both a street dealer and his customer into custody. Then, going around the room, each table made a statement about what they had learned in the class.

The short term effects of training

The short-term training study introduced a modest quasi-experiment into the routine of the academy. It contrasted the views of officers who took the survey before they began their training with the attitudes of officers who completed the survey at the conclusion of their session. The calendar date was used to assign classes to treatment (after) or control (before) status. Questionnaires were distributed before the class began on odd-numbered days (for example, December 3), and at the end of the class on even-numbered days. The odd-day and even-day questionnaires were identical except for their colors (blue or white), which helped the instructors track their proper administration. As an implementation check, a question on the survey asked trainees if they had completed the questionnaire before or after the session. To enhance compliance with procedures, we ensured that, from the beginning, trainers understood that this was an evaluation of training and not trainers. The study materials did not identify which trainers had conducted a class. One class was miss-assigned to take the survey after training instead of before, but it was retained in its improper category for this intention-to-treat analysis.

Because assignment was by classroom, we examined differences between "before" classes and "after" classes at the classroom level. Over the study period December 2012 to May 2013 the odd-even approach to allocating classrooms created groups that were very similar in terms of their measured characteristics. Table 1 describes some features of the data. In total, 2,681 officers completed the survey (another 0.6 percent chose not to participate); they participated in 133 classes, with 67 surveyed before training and 66 after training. The participants split 51-49 percent between the two groups. The survey questioned respondents about just two personal factors, their age and the unit in the department to which they were assigned; these questions were included in order to confirm that the treatment and control groups were similar. Table 1 presents the mean and median ages of the treatment and control groups, and the percentage of each group that was assigned to the agency's patrol division, which is by far the largest. Differences between them proved to be very small and they were not statistically significant.

*** Table 1 goes about here ***

Measuring procedural justice

The survey included questions regarding officers' views of their relationship with the public. Because of time constraints the questionnaires had to be very short; we promised the training academy that it would be one extended page in length. A section at the beginning described the survey and its purpose, and reminded trainees of their rights as human subjects. Officers indicated their assent by continuing on to complete the questionnaire, but they were not required to do so (this section was approved by the University's human subjects committee).

The survey next presented officers with statements to which they could respond, using four point response scales that ranged from "agree strongly" and "agree" to "disagree" and "disagree strongly." The statements were designed to reflect four procedural justice concepts that were emphasized at training. (For an extensive discussion of these dimensions see Hinds and Murphy, 2007, and Bradford, 2011.) How to address them in this study was an open question. Most procedural justice research has focused on the public's view of the police, and almost all procedural justice research involving the police has asked how they are being treated by their

own organizations and leaders (see, for example, Tyler et al. 2007). By contrast, measuring officers' views of how they should be treating members of the public, using questions reflecting the dimensions of procedural justice theory, was unexplored territory. The lack of a research tradition in this area required the evaluation team to craft new questions that promised to address the components of the theory. The questions had to be phrased in ways that generated variance in officers' responses. Some of the assertions were stated in such a way that agreement indicated support for procedural justice principles, while others called for disagreement by those who supported this view. This strategy was used to keep them alert to the content of the questions. The questions assigned to each dimension were also scattered throughout the survey, to discourage inattentive responses. The study team had to wait until the survey was complete to set the item responses that were measuring the attitudinal dimension. Responses to multiple questions were combined to create scale scores when they were single factored, and these factors explained a large proportion of the variance in the combined items.

Neutrality calls for consistency and even-handedness in decision-making across persons (equal treatment for all) and across time (the same procedures are followed every time). Neutral decisions are reasoned, objective and factually driven, and they respect rules and legal principles. In training, officers were encouraged to make it clear to those they encountered that they were acting in this way. Neutrality was measured by responses to four statements: "It is important to give everyone a good reason why we are stopping them, even if there is no need," "If people ask why we are treating them as we are, we should stop and explain," "When dealing with citizens' concerns, officers need to explain what will happen next, when they are done at the scene," and "It is very important that officers appear neutral in their application of legal rules." Responses to these statements were correlated an average of +.34. The Alpha reliability for this scale was .68.

Voice (sometimes this is referred to as "participation") involves giving citizens an opportunity to describe their situation and express their opinions about a problem (to "tell their side of the story") while officers are deciding what to do. Voice was measured by responses to two statements: "Listening and talking to people is a good way to take charge of situations," and "Officers need to show an honest interest in what people have to say, even if it is not going to change anything." The items in the summary voice scale were correlated an average of +.36, and the alpha reliability for the resulting scale was .53.

Respect encompasses treating citizens with dignity, acting politely, and granting them other routine interactional signs of respect. Because they are police officers, it is also important that they evidence concern about respecting people's rights, Respect was measured by responses to three statements: "People should be treated with respect regardless of their attitude," "Officers should at all times treat people they encounter with dignity and respect," and "It is important that we remind people they have rights and that we appear to follow them." The items in the summary respect scale were correlated an average of +.40. The Alpha reliability for this scale was .69.

Trust is evidenced when officers treat citizens as if they can be trusted to do the right thing and demonstrate that they are acting on behalf of the best interests of the people they are dealing with. Trust was measured by responses to: "Police have enough trust in the public for them to work together effectively," and "Officers should treat citizens as if they can be trusted to

do the right thing." Responses to these statements were correlated +.30. The Alpha reliability for this scale was .46.

<div align="center">*** Table 2 goes about here ***</div>

Results of Study 1

The statistical findings are summarized in Table 2. The results were positive; there were substantial differences between the treatment and control classrooms for all four dimensions of procedural justice. Table 1 presents mean scale scores and their standard deviations for all four measures, for both treatment conditions. All of the differences described there were statistically significant. The Table presents standardized effect sizes (Cohen's d) that ranged from 1.4 to 1.6. Effect sizes in this range were classified as "strong" by Cohen (1988: 287 and Table 8.2.2). Lipsey (1990) seconded this classification, based on his meta-analysis of 186 meta-analyses of intervention studies. The Table also reports correlations between treatment and outcome (Eta Squared); these ranged from .34 to .45, which Cohen translated as signaling a "large" effect. Finally, Table 2 also summarizes the results of a multi-level analysis—the percentage of the variance in the outcome measures that was explained by classroom-level treatment or control status. These ranged from 6 to 10 percent, and all were statistically significant.

<div align="center">*** Figure 1 goes about here ***</div>

The effects of training can also be seen in the data. Figure 1 compares treatment and control classrooms on the scores created by combining responses to the questions measuring each procedural justice concept. Each symbol in Figure 1 represents the data for a classroom in the study, divided into treatment and control groups. The "before" dots illustrate average scores for the 67 classrooms surveyed before training, while the "after" dots identify averages for the 66 post-training classes. The dots have been randomly "jittered" to the left and right, in order to visually expose any overlapping class scores. The regression lines presented in Figure 1 pass through the means of each cluster, highlighting differences between the groups.

Officers were most supportive of citizens' right to exercise voice. They gave the highest endorsement to the view that "Listening and talking to people is a good way to take charge of situations"; 48 percent agreed strongly with this statement. This position reflects very practical police wisdom regarding how they can get through their day safely, and the trainers frequently reminded officers of the safety benefits of adopting procedural justice approaches to dealing with the public. Classroom scores on the composite measure of participation are presented at the right of Figure 1, and they were the highest both before and after training. As Figure 1 illustrates, support for citizen participation went up, on average, following training.

In class, officers expressed the least enthusiasm for the procedural justice concept of "trust." The classroom-level trust scores presented in Figure 2 are visibly lower than support for the other procedural justice concepts in this group. Officers were most supportive of the view that "Officers should treat citizens as if they can be trusted to do the right thing"; 23 percent agreed strongly with this statement. On the other hand, 57 percent disagreed with the view that "Police have enough trust in the public for them to work together effectively." However, before-after differences in trust resembled those for the other dimensions of procedural justice, and the

average difference between the two groups was statistically significant. Notably, the largest gain scores (the differences between the treatment and control groups) were for trust; this can also be seen in the more sharply sloping regression line linking the before and after means for trust. Reflecting this, the standardized measure of treatment effect was the largest for trust. Officers' views of trust may have increased the most because they were relatively low to start with, lending more room for improvement.

Between trust and voice lay measures of neutrality and respect. For neutrality, officers gave the most support to the view that "It is very important that officers appear neutral in their application of legal rules"; 33 percent agreed with this statement. On the other hand, only 13 percent agreed strongly with the view that "When dealing with citizens' concerns, officers need to explain what will happen next, when they are done at the scene." However, the correspondence between officers' reactions to all four statements measuring neutrality was strong; the correlation between the most and least supported views was +.47. The effect of training on officers' views of neutrality was the second strongest in this set, based on the Cohen's d reported in Table 2.

When it came to support for the concept of treating citizens with respect, only 15 percent agreed strongly that "People should be treated with respect regardless of their attitude." In contrast, 40 percent endorsed strongly the idea that "Officers should at all times treat people they encounter with dignity and respect." Although these were the most divergent items, these responses were correlated +.49. The composite respect measure also shifted significantly between experimental and control classrooms, in response to training.

In summary, we found that training increased officer support for all of the procedural justice dimensions measured in the experiment. Post-training they were more likely to endorse the importance of giving citizens a voice, granting them dignity and respect, demonstrating their neutrality and placing some trust in the good intentions of the public.

The longer-term effects of training

The next question was, did the effects of training persist post-classroom? Studies conducted in police training academies have tracked officers' views of community policing, and they do not portend well in this regard. In a typical tracking study, Haarr (2001) distributed questionnaires to police recruits over the course of their early career. She found they grew more positive about community policing and problem solving during their stay at the training academy, but that this support began to dissipate during their field training. Their views grew less positive still during their subsequent probationary period.

To probe the persistence question, the longer-term effects of training were examined in a survey of randomly-selected officers. This survey began well into the training period, so that some responding officers would have attended training but others would have not yet been sent. A question in the survey identified those who had already participated in the procedural justice workshop and those who had not yet attended. This measure was used, along with statistical adjustments, to examine training-related differences in officers' views of procedural justice.

The survey

The survey was conducted in each of the city's 22 police district stations. At each station we randomly selected Police Officers (the bottom rank) and Sergeants in fixed proportions from the current duty roster. They were also spread proportionally across shifts. Sampled employees were notified of their opportunity to participate in the survey through an administrative sergeant. Roll call presentations, flyers and wall posters, and an offer of coffee and donuts were used to promote participation by those selected, who remained anonymous to the research team until (and if) they appeared. The survey was presented to respondents on laptop computers that the research team set up in stationhouse conference rooms. This ensured that no one could hear their responses to the questions, and that they could proceed at their own pace. The survey included 120 questions, and took at average of 25 minutes to complete. The survey team made repeat visits to each district, around the clock, until the number of interviews targeted there was completed. The overall response rate was about 28 percent. A total of 714 Police Officers and Sergeants were interviewed. The results could be weighted to ensure that the distribution of respondents matched that of the universe of officers serving in the districts, but doing so did not change any of the results presented here. Therefore, the original, unweighted results are presented.

Assessing the impact of training

In the district survey, 67 percent of those interviewed indicated that they had attended "the all-day training workshop on procedural justice and legitimacy at the Academy." Because of the timing of the district survey, officers could have participated in the workshop as long as ten months in the past, or as recently as the week before; the timing of their involvement is unknown. This raises the possibility that factors having to do with how early in the queue officers were sent to training could be confounded with any effects of the training itself on their views. This is a real possibility in any non-randomized research on training, if trainees are volunteers, chosen for their special characteristics, or even selected for their presumed readiness for training. If this is the case, factors related to their selection, rather than the training they received, could account for their later enthusiasm for the topic. However, there are reasons not to fear this kind of contamination here. Officers in the control group identified in the district survey differed because they had not yet been trained; instead, they were further back in the queue. They were not volunteers; the goal of the process was to train everyone, especially in the patrol division. The bureaucratic process that sent daily quotas of officers to training was quite indifferent to their personal characteristics. By design, each district was instructed to send just a few officers to training each day, so participants trickled into the Academy rather than arriving in related batches. Officers whose turn was to come later, after the follow-up study was completed, were serving in the same districts and working the same shifts as those somewhat closer to the head of the line.

As an additional precaution, we also used propensity score adjustment to further control for possible measured sources of selection bias. They all were factors that could not have been affected by training. The list of potential selection factors included personal background factors such as race, gender, education, age, prior military service, rank, and watch assignment. The selection model also included crime and workload measures and community characteristics for

the areas where they worked (we knew this from the survey). Interactions between these factors were also explored, in order to enhance the fit of the selection model. Any influence of dummy variables representing the 22 police districts were included in calculating the propensity score as well. A Wald test and examination of the residuals indicated that a complementary log-log link logistic model provided the best fit to the data. The model fitting was done in R, using GLM. The result was a propensity score reflecting the likelihood that each respondent had attended training because of their personal characteristics and features of where they worked. It summarizes the apparent effects of all of the likely measured correlates of attending training in one control variable. This score was included in the regression analysis presented below, as one of the control factors accounting for measured differences among the officers that potentially were related to their views of procedural justice and participation in training.

Measuring procedural justice

The district officer survey included questions that were similar to those employed in the training academy questionnaires. Trust was measured by questions about whether citizens could be trusted to do the right thing, if they have good intentions, and (in a reversed item) if it is naïve to trust citizens. Responses to these questions were correlated at average of +.50, and the scale had an Alpha reliability of .75. Voice was measured by responses to the statements "Officers need to show an honest interest in what people have to say, even if it is not going to change anything," and "Listening and talking to people is a good way to take charge of situations." Responses to these questions were correlated +.41, and as a combined scale had an Alpha reliability of .58. The importance of treating people with respect was measured by responses to the statements "People should be treated with respect regardless of their respect for the police," and "People who break the law do not deserve to be treated with respect." Responses to these questions were correlated +.69, and had a scale reliability of .82. Efforts to demonstrate neutrality were assessed by responses to the statements "It is necessary to give everyone a good reason why they are being stopped, even if it is not required," and "If people ask why they are being treated like they are, it is necessary to stop and explain." Responses to these were correlated +.47, and the scale had an Alpha reliability of .64. These positively and negatively worded items were mingled among questions on many other topics, to encourage respondents to pay attention to each.

*** Table 3 goes about here ***

Results of Study 2

The results of the analysis of the long-term impact of training on officers' views of procedural justice are summarized in Table 3. The unstandardized regression coefficient associated with participation in training is presented at the top of each column. As noted earlier, it could be important to control for measured selection variables using the propensity score, and the effects of this are taken into account in Table 3 as well; the measure is labeled "selection." In addition, Table 3 controls for a list of factors that potentially were related to support for procedural justice directly, and not just through any influence they may have had on participation in training. The most important of them was race. Regardless of training, African American officers were more supportive of procedural justice principles than were their white peers.

Chicago's procedural justice training appears to have had a relatively enduring impact on three of the four dimensions of procedural justice included in the short-term experiment. The coefficients associated with training in the first row of Table 3 can be compared in magnitude, and the biggest effect of training was on respect. The propensity score used to model the selection process was not significantly related to any dimension of procedural justice. This reflects the fact that there was not a strong correlation between having attended training and having a high propensity score, because there was no strong measured selection process. In addition, African American officers and officers over age 35 were significantly more likely than others to be supportive of the principle of respect they heard about in training. Training also had significant, positive effects on support for neutrality and voice. Compared to white and Hispanic officers, African Americans were also more likely to endorse evidencing neutrality, and the principle of voice was also endorsed distinctively by female officers. The coefficient associated with having a college degree was negative in every case, but significantly so in only one comparison.

However, in the long run, training does not appear to have had an impact on the willingness of Chicago officers to trust the public. The sign associated with the coefficient estimating the impact of training on trust was in a positive direction, but it was not statistically different from zero. Recall that trust was also the dimension of procedural justice that received the lowest level of support in the short-term training study. The trainers reported that trust was a difficult topic for them to address. As one trainer put it,

> We've been taught to trust nobody, to show less emotion. . . .We come to see
> everything as bullshit – going to another domestic, going to a beat meeting, going
> to training. We come to see people as assholes. But in reality, there's a big
> population we never come in contact with. It's the rest of the world.

In training, trust was discussed around a slide labeled "Us versus Them." It proclaimed that "94-97% of the public is law abiding," and the lecture and discussion revolved around why police too often believe that most members of the community are troublemakers who do not support the police. In this study, however, the influence of training proved least enduring when it came to trust.

The analysis examined other hypotheses that were recommended by our collaborators at the training academy. One common conjecture was that officers with higher-risk assignments might be less interested in extending the civilities of procedural justice to people they encounter. However, neither a two-question measure of perceived job risk from the survey, nor recorded crime or calls-for-service data for the beats and districts in which our respondents worked, influenced their views of procedural justice or interacted with the effects of training. We also heard that officers in high-workload areas might not have the time to listen to what people have to say or to explain things to them, perhaps even if they thought it would be a good idea. But to the contrary, three workload questions on the survey and workload statistics from the city's emergency call center were also unrelated to officers' views of procedural justice or to any effect of training.

A key finding was the importance of the credibility of training. The trainers were concerned about this. A trainer reported, "One of the biggest criticisms [of training] is, 'This won't work in <u>my</u> beat. This won't work in my district.'" On the other hand, trainers reported that many officers found that the class reflected their experience on the street. A trainer told us, "I'll get a lot of participants coming to me after class and saying, 'This is great. This reinforces what I do every day. It's a refresher. It reminds me what I do on the job.'" The survey captured this concern with questions asking trainees "How practical was the training for working officers?" and "How realistic was the training in reflecting the realities of life on the street?" Overall, about 70 percent of trained officers were to some extent positive about these aspects of their training experience. A separate statistical analysis (not detailed here) found that the more credible the training was on either rating dimension, the more impact training had on their views.

Limits of the Study

There were, of course, many limitations to the studies reported here. Both the experiment and the survey assessed only attitudes or opinions. There was no possibility of matching the survey data to the personnel records of participating respondents, so we did not have the ability to track the influence of training on other measures of their on-the-job behavior. This was one of the great contributions of the Wheller et al. (2013) experiment in Greater Manchester. Rosenbaum's (1987) findings in this regard provide an important reality check. In a randomized study, he found a shift in the attitudes and behavioral intentions expressed by newly trained police officers concerning how to better handle victims. However, an ensuing telephone survey of actual victims whose responses could be linked to the treatment or control status of their responding officer found that this training had no notable impact on measures of victims' emotional state, fear, feeling of vulnerability, or crime prevention behavior. The evidentiary contributions of training evaluations will be greatly enhanced when linkages like this can be pursued.

Because we found little guidance in past research, the measurement process for these studies necessarily had to involve exploratory scaling of officers' responses. As a result, some scales – which were already constrained by the required brevity of the training academy questionnaire – did not include many items. The reliability measure Cronbach's Alpha counts the number of items heavily in its calculations, so it was often relatively low. However, all of the items in the scales were substantially correlated and (when there were more than two) single factored, and -- as the discussion of example questions for each scale illustrated – responses to them displayed considerable variation between officers.

A limitation of the long-term survey study is that it involved only Police Officers and Sergeants with district assignments. Officers serving in other units, including detectives, organized crime, the narcotics bureau, the traffic division, and other branches of the department were trained at the police academy, but they were not included in the follow-up survey. However, the Bureau of Patrol includes about 82 percent of all department personnel, and those currently serving in other units all once served in patrol as well. As a result, findings based on patrol officers could be reflective of Chicago policing in general.

Another limitation of the long-term survey project is that – in contrast to the quasi-experimental training study – it used propensity scores and other control variables to adjust for measured differences between trained and untrained district officers. However, the study had the significant advantage of using as a control group not-yet-trained officers serving in the same areas and jobs as those who had already passed through training. As a result, there were no strong measured selection-related differences among them. This, along with the general consistency of the long-term findings with the results of the short-term experiment, increases our confidence in the regression analyses. Nonetheless, there could remain unmeasured differences between trained and not-yet-trained officers that were not accounted for by the selection model.

Discussion

The first study reported here found that, in the short-term, training increased officer support for all of the procedural justice dimensions included in the experiment. These were voice, respect, neutrality and trust. By our measures, training "moved the needle" among officers, and many held different views when they left the building. The second study explored whether these effects persisted once officers returned to their duties. In a follow-up survey we found that, post-training, officers who reported participating were more supportive of three of these dimensions, but were not statistically more likely to endorse the principle of trust. However, training is certainly not the only strategy for redirecting police behavior. Even staff members of the training academy were not sanguine about relying on one day of training to change the organization. In their view, change would at least require follow-up, reinforcement training. At the time of this writing this project was still in development . Evaluations of the effectiveness of such new training regimes are required if policing is to claim to be evidence-based with regard to one of its most fundamental organizational processes.

However, in the long term, monitoring, supervision and discipline also have to be part of the behavior change mix. Call centers and help desks listen in on their call takers' encounters with customers; foremen pace the assembly line; university students complete course-and-teacher evaluation forms. The difficulty is that everything about policing makes officers' actions on the street hard to penetrate, and police organizations have always struggled to maintain control of their field force. To fill this lacuna, supervisors need to do their jobs: allocating the workload correctly, staying alert for problems, and managing the needs of both employees and the organization. It is probable that supervisors need procedural justice training as much as anyone in the organization, but with an eye toward exercising its principles internally rather than externally. Supervisors and their managers set the tone of the work environment, and studies of the effectiveness of training suggest that it is in combination with effective management that training works best. Likewise, randomized experiments in responding to domestic violence demonstrated that careful training and monitoring can control how officers handle cases. In this multi-site project, officers successfully alternated between arrest, counseling and other strategies in handling domestic violence, following assignment rules given to them by researchers (Garner at al. 1995). Research by Mastrofski (1996) touching on the effectiveness of training officers to make drunk driving arrests concluded that it depended upon the organizational environment. The training worked better in agencies that monitored the arrest practices of their officers and had policies and practices in place that supported and rewarded focusing on DUI cases. In

organizations that were not aligned in support of what they said they wanted their officers to do, the amount of training officers received had no influence on their behavior.

In Chicago, one mechanism for aligning leadership in the districts with the priorities of top management is CompStat. At CompStat sessions, various administrative statistics flash on giant LCD screens as unit commanders are questioned about the shortcomings they reveal. There has been discussion about incorporating evidence of the units' performance in delivering procedurally just service, but to date it has proven difficult to figure out how to do this in any practical way. A central problem with CompStat is the "measuring what matters" issue. There are valued things that we want police to do that are not reflected in the data flow that drives CompStat, and building legitimacy in the community is high on that list. Holding officers to account for the procedural justice they do or do not deliver, and rewarding good service, could be the ultimate key to moving an agency in that direction, but – in Chicago, at least – it is unclear how this could happen.

Acknowledgments

References

Bradford, Ben. 2011. "Voice, Neutrality and Respect: Use of Victim Support Services, Procedural Fairness and Confidence in the Criminal Justice System," Criminology & Criminal Justice, 11, 345-366.

Cohen, Jacob. 1988. Statistical Power Analysis for the Behavioral Sciences, Second Edition. Hillsdale, New Jersey: Lawrence Erlbaum.

Garner, Joel, Jeffrey Fagan and Christopher D. Maxwell. 1995. "Published Findings From the NIJ Spouse Assault Replication Program: A Critical Review." Journal of Quantitative Criminology, 8, 1-29.

Haarr, Robin N. 2001. "The Making of a Community Policing Officer: The Impact of Basic Training and Occupational Socialization on Police Recruits," Police Quarterly, 4, 402-433.

Hinds, Lyn and Kristina Murphy. 2007. "Public Satisfaction With Police: Using Procedural Justice to Improve Police Legitimacy," The Australian and New Zealand Journal of Criminology, 40, 27-42.

Jackson, Jonathan, Ben Bradford, Betsy Stanko and Katrin Hohl. 2013. Just Authority: Trust in the Police in England and Wales. London: Routledge.

Lipsey, Mark W. 1990. Design Sensitivity: Statistical Power for Experimental Research. Newbury Park, CA: Sage.

Mastrofski, Stephen D. and R. Richard Ritti. 1996. "Police Training and the Effects of Organization on Drunk Driving Enforcement," Justice Quarterly, 13, 291-320.

Mazerolle, Lorraine, Sarah Bennett, Jacqueline Davis, Elise Sargeant and Matthew Manning. 2013a. "Procedural Justice and Police Legitimacy: A Systematic Review of the Research Evidence," Journal of Experimental Criminology, 9, 245-274.

Mazerolle, Lorraine, Emma Antrobus, Sarah Bennett and Tom R. Tyler. 2013b. "Shaping Citizen Perceptions of Police Legitimacy: A Randomized Field Trial of Procedural Justice," Criminology, 51, 33–63

Murphy, Kristina, Lyn Hinds and Jenny Fleming. 2008. "Encouraging Public Cooperation and Support for Police," Policing & Society, 18, 136-155.

Rosenbaum, Dennis P. 1987. "Coping With Victimization: The Effects of Police Intervention on Victims' Psychological Readjustment," Crime & Delinquency, 33, 502-519.

Skogan, Wesley G. and Kathleen Frydl. 2004. Fairness and Effectiveness in Policing: The Evidence. Washington, DC: National Academies Press.

Tyler, Tom R., Patrick E. Callahan and Jeffrey Frost. 2007. "Armed, and Dangerous (?): Motivating Rule Adherence among Agents of Social Control," Law & Society Review, 41, 457-492.

Tyler, Tom R. and Jonathan Jackson. 2013. "Future Challenges in the Study of Legitimacy and Criminal Justice." In Justice Tankebe and Alison Liebling (Eds.), Legitimacy and Criminal Justice: An International Exploration. Oxford: Oxford University Press, 83-102.

Wheller, Levin and Julia Morris. 2010. Evidence Reviews: What Works in Training, Behaviour Change and Implementing Guidance. London: National Police Improvement Agency.

Wheller, Levin, Paul Quinton, Alistair Fildes and Andy Mills. 2013. The Greater Manchester Police Procedural Justice Training Experiment: Technical Report. London: The College of Policing.

Table 1: Participation in the Classroom Experiment

	number of officers	number of classrooms	mean age	median age	pct in patrol
treatment	1392	66	42.3	42.5	89.8
control	1289	67	42.1	41.8	89.7

Table 2: Summary of Experimental Findings

	mean	standard. deviation	signifi-cance	Cohen's d	eta squared	classroom level R-sq
neutrality						
treatment	3.1846	.12277	.00	1.60	.40	9%
control	2.9908	.1177				
respect						
treatment	3.1492	.14700	.00	1.49	.36	8%
control	2.9331	.14338				
trust						
treatment	2.7995	.17898	.00	1.65	.41	10%
control	2.5217	.15634				
voice						
treatment	3.3325	.11764	.00	1.42	.34	6%
control	3.1635	.12032				

Note: 66 treatment and 67 control classrooms in every comparison

Table 3: Summary of Long-Term Effects Findings

	trust			voice			respect			neutrality		
	b	sigf		**b**	sigf		**b**	sigf		**b**	sigf	
training	**0.200**	(.06)		**0.216**	(.00)	**	**0.418**	(.00)	**	**0.212**	(.04)	*
selection	-0.255	(.33)		-0.193	(.29)		-0.576	(.06)		-0.253	(.29)	
white	0.166	(.26)		-0.019	(.86)		-0.326	(.06)		-0.120	(.39)	
black	0.532	(.00)	**	0.171	(.14)		0.526	(.01)	**	0.582	(.00)	**
latino	0.118	(.50)		0.144	(.24)		0.289	(.16)		0.276	(.08)	
age35	0.223	(.01)	**	0.125	(.15)		0.283	(.04)	*	0.136	(.21)	
college	-0.042	(.67)		-0.043	(.53)		-0.168	(.15)		-0.173	(.05)	*
female	-0.104	(.34)		0.166	(.03)	*	0.163	(.21)		-0.073	(.46)	
intercept	2.396	(.00)	**	3.946	(.00)	**	3.180	(.00)	**	3.631	(.00)	**
R-squared	.04			.03			.08			.07		

* $p<.05$

** $p<.01$

Fig 1: Four dimensions of procedural justice, before and after training

67 control and 66 experimental classes