

DARYL FOX: Good afternoon, everyone, and welcome to today's webinar, "Update on the NCVS Instrument Redesign: Additional Findings from the National Field Test and Plans for Implementation," hosted by the Bureau of Justice Statistics. At this time, it's my pleasure to introduce Heather Brotsos, Chief of the Victimization Statistics Unit within BJS, to begin the presentation. Heather?

HEATHER BROTSOS: Good afternoon. Thank you, Daryl. Welcome, everybody, to today's webinar on the NCVS Instrument Redesign. As Daryl said, my name is Heather Brotsos. I'm the Chief of Victimization Statistics at BJS. I have three esteemed speakers with me today to walk through key findings from a set of reports that were recently released on the National Crime Victimization Survey Instrument Redesign effort. David Cantor is Co-Principal Investigator for the NCVS Instrument Redesign at Westat and has more than 40 years of experience in designing surveys and evaluating the impact on social policy. Pamela Giambo is a Survey Methodologist at Westat with more than 30 years of experience in survey design, testing, and analysis. Jennifer Truman will round out the discussion with an update on the BJS implementation. She is a Statistician at BJS, and her current research focuses on stalking victimization, firearm violence, the measurement of demographic characteristics, and the redesign of the National Crime Victimization Survey Instrument. As Daryl mentioned, we'll save time at the end for questions. So, please go ahead and put those in as they occur to you over the course of the presentation. We'll address these at the end of the presentation. And with that, I will turn it over to David to kick us off.

DAVID CANTOR: Okay. Thank you, Heather. It's great to see everybody here. I'll first provide a brief overview of what we'll be talking about today, and a little bit more about the—overall about the NCVS Redesign.

So, this has been a long process. It started in 2014 with discussions with the Census Bureau and doing some mock interviews. And you can see we did quite a bit of testing over the next two to three years, culminating in a field test, which was done in 2020. After the field test, we ran a number—we produced a *Topline Report*, which was out at least a year ago, I think. And then we subsequently put out these other reports to provide more information on the redesign. The analysis of the pilot was primarily in 2021. And then, from that, we came up with final recommendations for the instrument. And we're in this last box now. We're in the homestretch on the Census implementation of the redesign. And Jennifer will be talking about that at the end of today's session.

These are six reports that have been put out so far. The *Topline Report*, which we talked about in a previous webinar. Today, we'll be talking about the Police Performance and Neighborhood Safety measures that were added to the survey, *Measuring Crime in*

*the NCVS*, which talks about some of the major changes that were done to the survey, the Instrument Redesign Test Methodology, which Pam will be talking about in a minute, which basically describes the methods. And then there are two other reports that are out there which we will not talk about, but we recommend it for those interested to look at those. One talks about how we redesigned the Crime Incident Report and then a couple of experiments that we ran with advance letters and incentives.

So, just a little bit about each of these. I'll talk about the Police Performance and Neighborhood Safety a little later on in more detail. But this basically involves adding two different sets of questions to the survey, one on Police Performance and one on Neighborhood Safety. This was introduced—I'll talk about this a little bit later, but these questions will come before the victimization screener and will be rotated between each of the two. So, a respondent will get one of these sections at one interview, and then when the interviewer comes back for the next visit, they'll get the other on Neighborhood Safety.

This report, *Measuring Crime in the NCVS*, we'll be talking about after Pam presents the methodology. This talks about the changes that were made to the screener, and some of the changes on the detailed incident form, and sort of the nature of each of those changes, the results from the field test, and recommendations that came out from that. In a minute, Pam will be talking about the Field Test Methodology and basically how the pilot test was designed, the sample design, the data collection methods, the processing of the data, and the weighting and variance estimation.

And then, one of the reports that we will not be talking about is the report on the Crime Incident Report or the CIR. And we did make a number of changes to the CIR. A lot of the changes have to do with how the flow of the CIR works, which we'll talk a bit about today. But we did make changes to a number of sections related to the victim-offender relationship, offender characteristics, self-protection, and you can see the list there. So, there were a number of changes. We greatly enhanced the section on victim services, which is one of the sections that probably received most of the attention—a lot of the attention—on the redesign.

And then finally, there's a report on an experiment that we ran with the web version of the survey that we had tested in the field test. The experiment involved looking at two different aspects of the survey. One was the advance material that's sent—the advance letter that's sent to respondents before the interviewer arrives. One version was the traditional paragraph format. The second was an icon format which tried to highlight different aspects of the letter, trying to make it a little bit easier for the reader to find information. And then the second experiment had to do with an incentive experiment

that we ran with the web survey, where we tested a \$20 promised incentive for each person completing the interview versus no incentive. I won't give away the results for those. But a hint, the letter didn't work so well, but the incentive really did work great. So, I encourage you to go ahead and read those reports.

And then just as a reminder for those of you who aren't familiar with the NCVS, this is the basic flow of the redesigned instrument. The flow itself at this level only differs in this one part where the Police Performance or the Community Safety measures are going to be asked. Those come right before the screener and right after the person characteristics, but otherwise, the flow is very similar to what it is now. So, once they get to the victimization screener, if they report any incidents, they get a Crime Incident Report for each incident that they report on the screener. Once they do not have any other incidents report, or if they don't report any incidents at all, then they go on to the next section, which depending on the person, it could be additional person or household characteristics. And that's it for me. And I'll hand it over to Pam. We'll talk about the methodology for the field test.

PAMELA GIAMBO: All right. So, this is a very high-level summary, the Field Test Methods. So in the field test, we administered both the current NCVS as well as the redesigned versions so that we would have a control. As in the current NCVS, the field test included a household roster instrument, which the Census calls the control card. And then, as David mentioned, the redesigned instrument we tested included new modules placed upfront in the survey that were relevant to everyone, victims and non-victims alike, which is why we nicknamed them the "Ask-all" modules. As David mentioned, these modules ask questions about respondent opinions about Police Performance or about Community Safety, and each household that was sampled for the redesigned instrument was randomly assigned to receive one of these modules.

Each respondent was asked a victimization screener that they call the NCVS-1 instrument. And as David said, if they screen in with a potential incident, they are also asked the Crime Incident Report, the CIR, the NCVS-2 instrument. Sampled addresses were randomly assigned to receive either the current or the redesigned version. And then within the redesigned instrument, we had two subversions with different approaches to the screener. There was an interleaved version of the screener, which included follow-up items every time a respondent said yes to a screener question. And then we had a non-interleaved version, which didn't include follow-up questions within the screener itself. And then finally, the sample excluded Alaska and Hawaii because it costs more.

Let's see. Okay. So, like the NCVS, the population of inference was all residents age 12 or older. But unlike the NCVS, the field test was only one wave of data collection. So, it wasn't a panel approach. Also, unlike the NCVS, the reference period was a full 12 months versus the NCVS, which is six months. On the back end, the process was similar to the NCVS processing. Type of crime codes, or TOCs, are assigned using [INDISTINCT] algorithm. And then these TOCs are subject to revision based on coder's review of the narrative information, or in our case, audio recordings, because we had audio recordings of most of the interviews.

Okay, Instrument Development. As David showed in his upfront slide, there's a lot of upfront work that went into the redesigned instruments. It was analysis of the raw NCVS data. We interviewed NCVS field staff. We held many discussions with a very patient Census team responsible for editing and coding. Early work included literature reviews, formative research, support from a great technical review panel, who helped guide survey content and looked at data collection approaches. After the draft instruments were ready, we used cognitive and usability testing as well as expert review to fine-tune the final questions. The roster and control instruments—the current NCVS—were programmed into a survey software system called Blaise. And the redesigned instruments were programmed using a language called PHP, which is a programming language used to create interactive websites not specific to surveys.

Field Test Design. For the field test, both the control instrument, the NCVS, which we called Condition 1, and the experimental version, which we called Condition 2, were administered by a field interviewer. The goal was to collect 3,000 interviews using the control version and 5,000 with the experimental version—2,500 each with the two different screener approaches. Most of the interviews were conducted in person. But just like how Census handles the NCVS, interviewers could administer the survey by telephone. Sample Design. Since it was in person, it was clustered—first stage based on counties or groups of counties, then tracts or tract groups, and then households.

So, first, we did a small pilot that served as a dress rehearsal of the data collection and training procedures. After the dress rehearsal, we assigned field interviewers to either the Condition 1 control group or the Condition 2 experimental group. And then the trainings were implemented separately for each group. After advance letters were mailed, interviewers visited the household. And before they went in, they completed a mini-survey about the neighborhood and the dwelling unit. They were asked questions about presence of trash, graffiti, whether the interviewer thought the neighborhood looked safe, etc. And then finally, during data collection, Westat used many different methods to validate the interviews and then also to assess data quality and how the instrument was working.

Field test response rate for the control and experimental version were similar, at around 25%. The control group was slightly higher, but not by much.

Okay, Data Processing. On the back end, we used the Census hierarchical algorithm for assigning type of crime to the Condition 1 control version. And then we just adapted that for the experimental Condition 2 version. For the field test, we included any survey in the analysis that had sufficient data. So, for purposes of the field test, we just define that as respondents who had at least one usable TOC as well as respondents who didn't have a TOC, a type of crime. But they answered all the key items in the screener, and at least, all the items in one CIR, if applicable to them. And then as I mentioned, we had coders review all the variable narratives collected by the interviewer. And the coders could revise the type of crime as needed. And then in our case, this review also included listening to the audio recordings.

And we created weights for both the control group and the experimental group. And the weights we generated were at the household level, person level, and incident level.

And then finally, replicate weights were created to support variance estimation. And there's a lot more details about the Field Test Methodology available on the BJS website. And now, we'll turn it back to David Cantor. He's going to talk more about some of the findings.

DAVID CANTOR: Thanks, Pam. So, this is in a report called *Measuring Crime in the NCVS*. And it's talking about some of the major changes that we found. So, this is a repeat of the first slide that we talked about for this report, just summarizing how the NCVS works. And what we concentrate on in this particular report, and it has to do with collecting data on incidents and classifying those incidents into a type of crime code that BJS is interested in—uses when they publish their annual estimates—as well as what users have available when they use the public use tape. So, I'll talk about the nature of the changes, the results from the field tests, and some of the recommendations that came out of this.

So, what were some of the major changes to the victimization screener? Well, one was based on the idea that the current NCVS on the screener does not really use the information collected on the screener when they do the classification. And this was based on—for reasons that were reasonable at the time. But as the feedback we got from the NCVS interviewers, as well as from the Census staff and others, this does cause problems, because if you're not using the information from the screener, respondents can get frustrated because you can ask the same question twice, once in the screener and then once in the Crime Incident Report. And that can lead to both

frustration and measurement error. So, we know from survey research that if you ask people the same question twice, many times they reinterpret the second question in a different way, and that can lead to measurement error. And this actually occurs when—two examples on the current CIR, where this occurs was—when asking about presence at a particular incident or the location of where it occurred, and those were used as filter questions for measuring violent crime and for burglary, and if someone didn't answer those questions correctly, they would skip over all the questions about violent crime, which would make it very difficult to—or burglary—so it would make it very difficult for classifying that particular incident.

So, what we decided to do for the screener was to reorganize it slightly. So, each section of the screener is related to a particular type of crime. You can see those type of crimes listed here. And note vandalism is a new type of crime that we added as part of the redesign. And then when they answer those questions, the responses to those questions are used to route them through the Crime Incident Report. So, if they told us in the screener, "Yeah, someone broke into my house," or "Yeah, someone robbed me," we don't ask them questions about whether they were present, or we don't ask them about certain kinds of attack questions, because they've already told us about that in the screener. We ask additional questions which provide more information about what happened and allow us to classify the incident.

Another change was, within each of the screener series, we ask separate questions within each series. So, in a couple of instances on the current screener, there are long lists of cues or examples that are read to the respondent. And in all cases, the interviewer is asked to record a single yes/no from the respondent. So, once listening to all these cues, a question is finally asked to the respondent whether anything happened that falls under those particular types of crime. Well, that seems very awkward and it also leads to pre-empting the interviewer, as reading the cues many times, respondents will interrupt and say yes or no because they think they're being asked about that particular cue. So, we redesigned and asked respondents to provide a yes/no to a particular series of cues of like-type crimes, so that they're answering more questions as part of the screen.

We updated the cues. The cues hadn't been updated since the early 1990s, so we put in things like cell phones and other sorts of things that brought things up to date in terms of things that are going on out there in the world. And then a very big expansion was to expand the cues that are used to asked about unwanted sexual contact. And those cues were developed as a result of some of the research that BJS has been doing over the last 10 years or so on measuring rape and sexual assault.

Some of the major changes to the CIR: We reorganized it so that most of the questions are at the beginning of the CIR that are used to classify the type of crime, or the TOC code. Each of the sections have to do with those major types of crimes that I just listed on the screener—we call them the “What Happened” sections, and they basically ask additional questions that provide information on how to classify the incident.

As I said before, we're using the screener and the follow-up probes to drive the skip patterns on the CIR. So, there are parts of the CIR that may never get asked for a particular incident because they weren't relevant. The screener—perhaps the person said in the screener that someone stole their laptop from their house and there was no indication that they were ever confronted or anything like that. So, the questions about robbery or anything like that are not asked, they're just skipped. And that's a big change from the current CIR where questions are asked even though there was no indication in the screener that anything like a robbery occurred, for example.

As I had mentioned, we changed the definition of “presence” and removed it as a requirement both for presence and for location when asking about violent crime and burglary respectively. This eliminates the possibility that if they answer those questions incorrectly, they'll get skipped out of questions that they need to answer. And we added a lot of details about the use of weapons. I'll mention, there's also another report that talks in more detail about some of the substantive changes that were made to the screener that I encourage you to take a look at.

Another big change to the CIR was on the Rape and Sexual Assault measurement. There's a specific section now on rape and sexual assault, whereas before there was no such section. In the current NCVS, there's just an open-ended question asking what happened, and the respondent has to volunteer information that would provide the interviewer enough information to classify it as a rape or a sexual assault. So, it's asking more directly about things like the type and direction of penetration, if penetration was involved; other types of unwanted sexual contact; and it also gets into specifics about the tactics and the behaviors that constitute rape and sexual assault. So, things like use or threat of force. Was the person blacked out, unconscious, or unable to consent? Other forms of coercion, and whether the offender had stopped after being told to stop. So, there's a lot more direct measurement of the elements of a rape and sexual assault that are not currently on the Crime Incident Report.

So, we did the pilot test. Pam described the design of the pilot test. And what were some of the big takeaways? Well, first of all, the redesign estimates came up with higher estimates for the current NCVS. Now, let me remind you that the design of the field test was such that our control condition was the current NCVS, that's what we

called Condition 1. And then there were two other conditions related to the redesign, which was the interleaved version and the non-interleaved version. And the results of the pilot test were such that the rates for the redesigned survey were higher than the estimates for a number of crimes, including rape and sexual assault, which wasn't that surprising given the changes we made to the screener and the CIR; robbery; burglary; simple assault; and household theft of lower valued items, so if you want to say less serious kinds of crimes. We weren't necessarily expecting higher estimates in any of the other types of crime besides RSA, so we were surprised that we saw higher estimates. We attribute some of this to more individual questions on the individual screener, so that procedure that we changed whereas the respondent has to answer yes/no to a number of questions about specific cues rather than just giving one global yes or no could've led to a higher rate. Some of the follow-up probes that we put in there might've helped identify things like burglary and robbery, which are a bit more complex to classify because they involve several different elements, such as breaking in and stealing, for burglary, or for robbery, use of force and taking property, which involves several different elements that you have to be asked about.

And we also removed the CIR's dependence on the presence and location questions, and that helps for both violent crime and for burglary. Of course, some of these results have to be verified when the Census Bureau actually implements the NCVS as they're about to do in the next few months.

Other field test takeaways: The redesigned questionnaire did identify more victimization incidents and that's what led to a higher victimization count. I'll show you a little data on that in a minute. We found that the classification of the incidents was more efficient for the redesigned survey—we asked fewer questions and it took less time. So, we ended up recommending to proceed with the redesigned questionnaire, and we recommended of the two versions that we tested, we tested the non-interleaved version, and we decided to use that because it was better able to identify these more complex crimes like robbery and burglary.

So, just to remind you, the interleaving, the interleaved version had to do with if someone said yes to a screener response, we would follow up that yes response and ask about whether certain elements of—if there were other elements to the crime besides what they told us about. So, for example, if they said to the theft question on the screener, something was stolen, we then ask, well, were you attacked in some way or did someone break into your house? And the purpose of that was to make sure that we got all the elements of the crime in the screener, which would then help us drive the skip patterns on the CIR. So, the interleaved version asked those probes right after each yes question on the screener, whereas the non-interleaved version move those



probes to the beginning the Crime Incident Report, so it was only asked for those questions that—so those questions were administered before getting into the CIR and helped drive the skip patterns. And it turned out that the NIL, the non-interleaved version, seemed to work better. So, that's the version that we're recommending.

We had other changes based on the field test experiment that I'm not going over here and encourage you to read the report.

Here's just a little bit of data. This shows the higher rates of reporting incidents for the two test versions—Condition 2 IL, Condition 2 NIL, are the two test versions—and Condition 1 is what currently is done on the NCVS. You can see that the mean positive responses per respondent to the screener was significantly higher for the two Condition 2 versions, 0.20 versus per Condition 1 and 0.30 and 0.32 for the two Condition 2 versions. So, that's really what's driving the higher victimization rates, and that's one reason why we think the additional probing or asking questions after each shorter set of cues is driving some of this difference. And you can see that once someone said yes, the number of incidents that are reported is about the same between the three versions, so 1.81 versus 1.78. So, what's driving a lot of the difference, at least in our field test, was just saying yes to a screener question.

Major changes to the TOC Classification. There was a complete overhaul of the RSA TOC in terms of what's used to define each of the different types of rape or sexual assault. The actual TOCs don't change very much—they still include things like completed rape, attempted rape, completed assault, sexual assault, attempted sexual assault. But we also added as part of this some non-NCVS crimes, including things like "Coerced Penetration" and "Coerced Sexual Contact." These are unwanted sexual contacts but don't necessarily rise to the level of being a crime.

The Personal Theft was expanded to any items on the person that was not coded as a robbery. So before, the personal theft was related to pocket picking and purse snatching and we expanded that a bit. As I said before, the Violent Crime codes don't require a report of presence because it's implicit in what's answered in the questionnaire, and we added Vandalism. And that's it for the *Measuring Crime* report.

Take a second to take a breather here. And now I'll talk about the Police Performance and Neighborhood Safety questions that were added to the NCVS. And as I mentioned at the beginning, these were added right before the screening questions. These are going to be asked of everybody 12 and over, and they have to do with Police Performance and Neighborhood Safety. And these items were added because they address two significant priorities of the NCVS: One was to provide data for local areas

and for analysts when doing policy research. BJS convened a group of police chiefs and other major local policy makers before the redesign, and they recommended that these kinds of questions be added to the survey. And Groves and Cork, in their review of the NCVS and what should be done with it in the future, also recommended adding questions that provided more analytic utility as well as data for local policy makers. So, these particular items will be used for both purposes.

The other priority is to engage respondents at the beginning of the interview. So, most people to the NCVS don't report any crimes. And after going through the survey once, for example, the first time [INDISTINCT] and the interviewer comes back and they have nothing to report again, they start losing interest. So, the idea here was to engage respondents at the beginning of the interview, even if they didn't have anything to report, and hopefully this will help motivate them to continue with the NCVS.

So, let's first talk about the Police Performance. And we have contact with the police because we ended up adding that. I'm sort of giving you a preview about what I'm about to say. But the Police Performance questions are based on procedural justice concepts. These were developed with a special subcommittee of the technical review panel. Wes Skogan worked closely with us to develop, test, and finalize the wording of the questions. They're based on the idea that popular assessment of the police is affected by what they perceive as the process and how the individuals are treated by the police. So, things like perceived fairness of the process and how the individual is actually treated. The original draft of this module included measures of seven dimensions—things like voice, trust, respect, neutrality, legitimacy, value alignment, and distributive fairness. These are the basic tenets of the procedural justice perspective.

But when we tested these items, they tended to be a little problematic. So, when we pre-tested, a number of respondents found the items difficult because they didn't have the direct experience with the police in order to make judgments about some of the questions asked. And a number of items seemed redundant to respondents—they didn't quite get the distinction between the different items. So, the changes we made were, first, we reduced the number of questions so that they did not seem as redundant as the cognitive testing showed. And we added a direct question on police effectiveness. And I'll illustrate what that is in a second when I go over the questions. We also added questions on prior contact with the police that are administered before we ask about police performance. And the idea here is that it'll allow analysts to assess opinions about whether there has been contact with the police and the type of contact and whether that affects attitudes about the police. The items that we used for these questions are an abbreviated set of items that were adopted from the Police Public

Contact Survey, or the PPCS, which is a supplement that is regularly administered by the NCVS.

So, first, here are the “contact with the police” questions that we asked. So, during the past 12 months, have you contacted the police in your area to report a crime, disturbance, or suspicious activity? Have you contacted the police in your area to report something else, such as a traffic accident or medical emergency? Have you been stopped by the police when you were driving or when you were a passenger in a motor vehicle? Have you been stopped or approached by the police for some other reason? And have you been at a community meeting, neighborhood watch, or other activities where the police took part? So, you can see there's two types of questions, one is citizen-initiated contacts and then police-initiated contacts, just like the current PPCS divides the questions.

And then in terms of Police Performance, we asked these questions, and the dimensions that they represent are shown there. So, for Respect is, how respectfully do you think the police in your area treat people? Sort of obvious on its face there. Voices: how much time and attention do the police in your area give to what people have to say? Distributive fairness is, how fairly do the police in your area treat people regardless of who they are? Police effectiveness: how effective are the police at preventing crime in your area? Trust: how much do you trust the police in your area? And then this is the direct question which we added: just overall, taking everything into account, how would you rate the job that the police in your area are doing.

So, in terms of results: contacting the police, around 43% of the respondents said that they had contacted the police in the last 12 months. About half said police initiated, about half citizen initiated. This is quite a bit higher than what is measured on the PPCS—the last version was around 20%, 21%. And we were a little surprised to see how much higher it was, but given results from other surveys that we have done and even with the pilot test and comparing victimization rates between what our administration gets versus what the Census Bureau or BJS gets in the normal rotating panel design of the NCVS, it wasn't entirely surprising, because even with the victimization rates, they generally tend to be much higher in the surveys administered outside the Census Bureau. And there are a number of reasons why that certainly could be. The placement in the interview, in the field test, the questions on the contact are at the beginning, whereas on the PPCS, those are supplements after the NCVS has already been administered. There's big differences in response rates between the two—the field test was around 27%, the PPCS was around 51%. Then there's just the normal differences in design. So, time and sample, the field test was one interview. People

getting the PPCS could have had up to six other NCVS interviews done before they are administered the PPCS. So, that could very well affect how they view the questions.

There's also a difference in the reference periods, as Pam mentioned, six months for the NCVS, 12 months for us. A big one could be that there are differences in the field staff experience with the NCVS. So, many of the NCVS interviewers are highly experienced administering surveys. Our interviewers were first time—they were inexperienced in that sense because this was the first time they administered the NCVS or anything like the NCVS. And we know from NCVS research that even within the Census Bureau, when new interviewers come in, they tend to get significantly higher victimization rates than the experienced interviewers. So, this could also be an issue. And then there's just general house effects, which is sort of a term that we don't really have good explanations for, but we know different organizations use different ways to train interviewers, monitor interviewers—all of that could very well have effects on these differences.

But the Police Performance items worked in ways that we had expected and hoped. So, for example, people who had contact with the police had less missing data. And these are especially evident for those items that respondents have less certainty about how they answer the question. So, you can see for Distributive Fairness, Voice, and Respect, there are big differences on how willing people are to answer the questions. And this in a sense validated the reasons for why we included the police contact questions.

And then, in general, the Police Performance items worked as we expected. They were correlated with the number of characteristics that we had hoped they would be correlated with, which research expects them to be correlated with. Here is just an example. You can see that there are much higher negative opinions of the police by those people living in areas where whether the interviewer thought the block was well kept or not. So, those people living in areas where the interviewer thought it was not well kept had more negative opinions than those areas where they thought it was well kept. So, this seemed to work as expected.

So, recommendations for the police items. Overall, we thought they worked as hoped. Something I didn't talk about here was the Police Performance items were related, although not consistently, with the police contact items. But I think that may not be totally surprising. There was some concern that the contact items had such high rates relative to the PPCS, so that's something that BJS will be monitoring as the survey is instituted in the normal production environment and see if that persists. If it does persist, then there needs to be some studying for the reasons behind this. And then we also

recommend that there be increased monitoring of those people, of the youth, 12- to 17-year-olds, to further assess the quality of their ability to report on police performance, especially. Our sample of 12- to 17-year-olds was not the best. It had a very low response rate because we had to get active parental consent in order to talk to youth. The Census Bureau is much more successful in working around those consent procedures. So, that's something that should be looked at in a little bit more detail when it's actually implemented.

So, what about the Neighborhood Safety measures? These are measures that were also helped or developed through a subgroup of the technical review panel. Janet Lauritsen and Wes Skogan worked closely with us to develop these items and finalize the wording. We implemented perceived risk of victimization through fear of crime and neighborhood disorder. And here, just like with the Police Performance measures, BJS anticipates producing regular reports at both the national and subnational level on trends and levels. A third measure that we instituted was something on collective efficacy based on the work by Sampson and colleagues. And this is composed of two different dimensions here, social control and neighborhood cohesion. And these weren't—BJS isn't necessarily going to be publishing annual reports on, but they were kept because they provide analysts with more information to look at victimization risk.

So, the Fear of Crime questions. Here are the fear of crime questions. Are you worried about being mugged, robbed in your local area? Threatened or attacked in your local area? Something stolen from inside your home or having something stolen from your porch, lawn, garage, or other part of your property? So, this has to do with what you might think of as the worry dimension. And you can see some of the tabs here. Most people are not that worried, from our field test. And then we asked about things that might have to do more with behavior, such as is there any place within a mile of your home where you would be afraid to walk alone at night and how often are you concerned about—how often does concern about crime prevent you from doing things that you like to do? These are probably familiar questions to many of you. These are taken from a number of different surveys.

So, in terms of the results, about 60% to 65% of the respondents said not at all worried about being mugged, robbed, threatened, or attacked—so for violent crime. There was somewhat more concern about property crime. All of the concern was—especially for violent crime—was correlated with things that we would have expected: gender or sex, Black and Hispanic persons, unmarried persons, those having personal experience with crime, and those on blocks that the interviewers reported as unsafe and not well kept were all correlated in the expected direction for the fear of crime.

The Neighborhood Disorder questions: Here are the questions that we included. In your local area, how common a problem is vandalism, graffiti, or other deliberate damage to property? Being drunk, rowdy in public places? Burned, abandoned, or boarded up buildings? People using or dealing drugs illegally? You can see the majority of the people said either not common or not common at all for most of these. But these were also correlated with characteristics that we would have expected neighborhood disorder to be correlated with, including a couple of the interviewer observations about being well kept and not feeling safe.

And then, finally, for Collective Efficacy, we had questions on social control and neighborhood cohesion. This is not by any means a full scale or measure of collective efficacy, as Sampson and colleagues have laid out or others, but we were limited in space and we did some analysis of some of the surveys that have used these scales, and we picked these out because they seem to be the best predictor. So, if children or teenagers in your local area were skipping school or hanging out on the street corner, how likely is it that any of your neighbors would do something about it? How likely is it about damaging others' property? And if there was a crime in your local area, how likely is it that any of your neighbors will call the police? And then for neighborhood cohesion, there were two questions: people around here are willing to help their neighbors and people in this local area can be trusted.

So, the results for the Social Cohesion measures, item nonresponse was low, and they were correlated with what we would expect.

So, for the Neighborhood Safety measures, these are the final recommendations. We basically recommend retaining all of the fear of crime and neighborhood disorder items. These would have direct policy and utility for users of the NCVS as well as tracking important trends over time. We did recommend retaining all the collective efficacy items because they do add in terms of covariance and we think they're questions that respondents enjoy answering.

And just as a final note, after we produced the report, BJS decided to rename this set of modules as Community Safety as opposed to Neighborhood Safety. So, in the future, don't get confused when you read things about neighborhood safety, even though we've been talking about—and the report is called—*Neighborhood Safety*. And that's it for me. I think Heather is—or Jenna is going to talk about implementation at this point.

JENNIFER TRUMAN: Great. Thanks, David. Hi, everyone. I'm going to be talking about the implementation of this instrument redesign and all the work that Pamela and David have presented today. So, I will cover kind of our decisions from BJS on the new

survey, our implementation plans, also some respondent communication updates, recent web testing efforts we've been doing, and then really next steps and the timeline. And we do have an NCVS Instrument Redesign web page available on the BJS web page. We have updates there. All the reports that we've been talking about here today are published there. And as we move forward, any updates will be there as well.

So, just as some reminders for anyone who's attended the previous webinar, key updates to what we're doing with this instrument redesign. So, we are maintaining the two-stage measurement approach in screening and classifying crime. So, we will maintain that screener and the crime that's in our report, as David discussed. We have significantly improved and updated the crime screening questions, including the quality of information that's collected, and really just allowing for more efficient flow through the Crime Incident Report and using the information that we're collecting to guide that CIR. We're also collecting expanded information on victimization incidents and help-seeking. And as David just talked about, we are adding those new questions on non-crime—the non-crime questions, we're calling them—on Police Performance and Community Safety.

You've seen this once before today, but as a reminder, the flow will be from the household roster to some initial characteristics, rotating the two new modules on police and community, and then into the screener. And from there, if someone did report an incident into the Crime Incident Report for each of the incidents they reported. And, finally, ending in the additional person and household characteristics.

So, the new NCVS instrument: The version that we tested in this field test is available in the *Field Test Topline Report* that was released last year. The Post-Field Test version. So, all of the recommendations coming out of that field test have been incorporated into a new version of the instrument, and that'll be released on our next phase of testing, which I'll talk about in a little bit. And then because we know it's really important for data users, we are working with our colleagues at Census and Westat to develop a crosswalk of the new NCVS instrument in comparison to the old, because eventually when we do release these data, we recognize that that will be super important for people who want to look at the data from the kind of current NCVS to the new NCVS. So, we'll have that crosswalk available and we're working on that as well.

So, next steps for us. Next month, which is just a couple days from now, July through September of this year, we will be conducting an operational pilot test. The purpose of this pilot test—will be conducted with the Census Bureau, who is our data collector—and it's really to assess the new survey instrument and the protocols within the Census Bureau data collection environment. So, as Pamela and David just mentioned, the field

test that we did was with the Westat team and their staff and their field interviewers and within their environment. And so, this is the test that was in the Census data collection environment. So, we want to make sure to test all the systems and the procedures within Census Bureau to make sure everything is set before we go through a full implementation. And this small pilot test will then inform the field protocol as well as interviewer training for the full implementation.

So, our sample: We're sampling—again, it's a smaller sample of about 3,500 households, with 45% coming from our existing samples (so those households that are already in the NCVS sample) and then about 55% from our reserve sample. And our reserve sample is selected and set aside for any research as we go along. So, something similar to this. So, primary sampling units, or PSUs, were selected based on available field staff, hiring, their current workloads, and also crime rates. Because it is a small sample, we do want to try to get as much data as possible. The sampled households for this pilot will be excluded from any official estimates. So, this test and these sampled households will not be included in any official estimates. Because it is a small sample, it's not representative of the U.S. and it's not designed to test for differences between the new and the current. Again, it's really just operational to make sure things are working before we go into full implementation. Because of that, it's not feasible for us to go through every combination of questions in the instrument, and so as a part of that, we're also doing testing with the Census Bureau to make sure the instrument is working as well.

So, from here, the pilot will then inform the next step, which is the split sample. So, January through December, an entire year of 2024, we will be doing a split sample. And that means that we will be both collecting the current NCVS instrument and the new NCVS instrument, and it's going to be applied at the second stage sample, and that's the stage where we select housing units. And so the instrument will be designated at the housing unit level and will be randomly split within the primary sampling units. So, half of NCVS sample cases within each PSU will receive the current instrument and half will receive the new instrument. And the interviewers will be administering both new and the current instrument. And our aggregate 2024 or aggregate annual sample will remain steady. So, right now, we're sampling about 240,000 persons in about 150,000 households, and that will remain steady through 2024 as well.

There are a couple things that impacted this final design. First, the estimate precision. So, we do know by splitting the sample, the results decrease in estimate precision because we are splitting the sample in half. However, adding sample does introduce complexities and risk. As David mentioned kind of earlier, and as we've seen with other sample redesigns, if we add sample, it means adding field staff, and that can also



impact some of the victimization estimates that we get. So, really, the stability of maintaining our current sample size does outweigh the loss of precision for us. Also, just the Census field operations, their structure and logistics. Designing the split sample this way allows us to maintain procedures for making interviewer-level piece assignments within the PSUs and it also keeps workflow consistent and preserves rapport for those on the [INDISTINCT] cases. So, any of those cases that are in their time and sample two through seven, the interviewer has established that rapport, so this helps maintain that as well. And then in terms of training procedures, all of the interviewers will be trained on the new instrument prior to the split sample, and BJS and Census will be monitoring various quality indicators throughout this transition.

So, really, part of the main purpose of this split sample is so that we can measure change in the estimates. So, this year, in 2023, we'll be remaining and reporting estimates on the current instrument. Next year, 2024, we'll have the split sample. And so by doing that, it allows us to compare data collected through the current and the new instrument from 2023 and 2024. And then it also allows us to compare the change over the years. So, the current to the current in 2023 and 2024. And then as we move forward in 2025, we will be fully on the new instrument. And so, again, the split sample then allows us to compare the new and the current instruments and then the change in estimates from 2024 to 2025 using the new and the new instrument.

The full implementation is planned for 2025. So all interviews will be conducted with the new instrument. 2024 estimates, for anyone who pays attention and downloads our data, will likely follow a delayed schedule. We will maintain as much as we can, but most likely we'll be on a delayed schedule. Estimates, as I mentioned, for both current and new instruments will be compared, as well as estimates examined over time to look at change estimates. And because we're going through this full instrument redesign, BJS is not administering any NCVS supplements in 2023 or 2024.

In addition to some of the instrument redesign, we're also refreshing some of our respondent communication. So, this year, we put in a new fact sheet in the field as of January, and that's an example of what the new fact sheet looks like. We have this both available in English as well as Spanish, Arabic, Chinese, Korean, and Vietnamese. All of these are available on our new NCVS participant web page. And we are also working with the Census Bureau to put a new advance letter in the field later this year. And that updates—traditionally, the letter is kind of a traditional paragraph letter in an update, so it's a little easier to read, with we call them kind of icons—it has more bulleted information for the respondents.

This is the new NCVS respondent web page. So, it just lays out information for anyone who is selected into the NCVS—why it's important, what they're supposed to do if their household is selected, what they should expect, to let them know that it is a panel interview, so once their household is selected, the household will remain and be interviewed every six months. And then just has resources for victims as well as participant materials, which includes the fact sheet that I just showed you, the letters currently, and will eventually include the updated letters as well.

So, where are we going now? In addition to the field test that we talked about here today, we also worked with our team at Westat to test a self-administered, web-based mode as a third mode for the future. So, right now, we collect face-to-face interviews and telephone interviews, and that will remain the same for the redesign at this moment. But we wanted to understand more about whether or not respondents were actually willing to complete the NCVS on the web. Do respondents exhibit signs of inattention or satisficing when they're completing it on their own? And then, importantly, how do victimization rates estimated from a self-administered version compare to those from an interviewer-administered version?

So, we did this with a two-sample design. It was an address-based sample and a probability-based online panel in the U.S. The household roster for the address-based was completed on the web or paper, and then all person-level interviews—or, excuse me—surveys were completed on the web. The data were collected January through May of last year. And, again, this is really research for the future. So, our Census Bureau team anticipates having technology in place for a web-based mode likely no earlier than 2027. And so BJS is working to do more research and testing to determine if and how web mode could be used in the future for the NCVS.

In terms of our longer timeline, as you can see, we are fully through research and testing. We have been working hard with the Census Bureau to test the instrument. Next month, July, we start this pilot test, which is where we're at right now. Next year will be the split sample. And, as I mentioned, the full implementation of the new instrument will be in 2025. And really, the goal here is the sample redesign happens in 2026. So, every year with the decennial census—or excuse me—every 10 years of the decennial census, all of the household surveys that the Census collects goes through a sample redesign to address any changes to the population based on the Census. And so that gets phased in for us in 2026. So, we really want to make sure that the new instrument is fully phased in before we get to that part of this as well.

Lastly, I will just say that this is the 50th anniversary of the NCVS and we are holding a celebration on September 27th. It's a two-part event. There's a morning session that's

going to be at the main Department of Justice and an afternoon session at the Bureau of Justice Statistics within the Office of Justice Programs. We encourage everyone to attend. There is an option for in-person or virtual attendees. So, please, we would love to see you all there, whether it's in person or virtual. And with that, I will thank you all again for attending, and we will move on to a Q&A session.

HEATHER BROTSOS: Thanks to all of our speakers. We're getting some great questions in. So, we'll see how many of them we can get to before the end of the webinar. I'm going to start with the list that we're getting in through the Q&A. And so the first question is, "During instrument development, were there conversations with trauma-informed experts—for example, psychologists, social workers, etc.—to inform question phrasing or order training procedures and to better understand potential measurement error?" And then I think maybe a related question is regarding—second question here is, "During cognitive testing, were you able to test the new question series on sexual assault victims?" So, I don't know who wants to start with those. Maybe start by talking a little bit about the instrument development process and what testing was involved. I'll let maybe—looks like David wants to go first.

DAVID CANTOR: Sure. Thanks. That's all good questions. In terms of the development of the questions, we had several different people on the technical review panel who are experts in asking these questions as well as working with victims. And so they were part of a small subgroup that reviewed the questions and helped us come up with wording for both the screener questions and the questions on the Crime Incident Report. I would also add that—both sets of questions were used on several different surveys that BJS had administered on rape and sexual assault in the last five or six years. And we did do cognitive interviews with victims of rape and sexual assault as part of the testing procedure.

HEATHER BROTSOS: Pamela or Jenna, anything you want to add to that?

JENNIFER TRUMAN: I would just say that—in the methodology report that Pamela had talked about and Daryl had linked to in the chat, it lays out the process for the instrument review and the technical review panel that David just mentioned. So if you want some more details, that's available there as well.

HEATHER BROTSOS: Thanks, Jenna. Okay. Next question. "Are the slides available after this webinar?" We will be posting the recording of this webinar, as well as the PowerPoint slides, on the BJS website. So, keep a lookout for that. We will make all of this information available to attendees as well as others that access the website.

There is another question about collective efficacy. This question asks, "While the results on collective efficacy will not be published, are the results available to researchers or academics?" Jenna, do you maybe want to talk a little bit about the data that we collected, and then if David or Pamela has a follow-up to expand more on collective efficacy, we can go from there.

JENNIFER TRUMAN: So, everything that we've presented here today is published and available on the BJS website. So, the results that David talked through on the police and community measures, and specifically on the collective efficacy, are available in that report. We are also working to release the data from the field test. So, we are working to make sure everything's good and we plan to release that as a public-use file eventually, and that would be—we archive all of the data through the National Archive of Criminal Justice Data, so that will eventually be up later. Hopefully this year, but just keep an eye on it for any of our kind of updates. If you do sign up for Justice Stats, we would do any updates there, so—David, I think you're on mute, if you're talking.

DAVID CANTOR: Sorry. The only thing I would add, and BJS, Jenna or Heather can weigh in, but when we were making decisions about what questions to add, I mentioned the idea that, for a lot of the police engagement and community measures, there are reports that we envision will be produced that show trends and looking at things at small areas that people will have a lot of interest in. This was not the main reason to include the collective efficacy items. They might be at some point, but at least right now they're primarily seen as questions that can be used by researchers. So, they'll be available on the public use tape and they can be used for whatever reason. That's the main reason to include them.

HEATHER BROTSOS: Thank you. Okay. We've got another question here, specifically about the Police Public Contact Survey supplement. Jenna, can you speak to whether this will continue to be fielded once the new modules are incorporated into the NCVS?

JENNIFER TRUMAN: Yes. So, the plan is to continue to field the PPCS, or the Police Public Contact Survey, while also including the Police Performance measures in the new instrument. As we've kind of seen here, there is overlap in terms of the contact, but the rest of it isn't overlapping, so we do anticipate collecting both. Part of the reason for the rotation is to try to avoid having them in the field at the same time. So, the goal would be that when we're asking the Police Performance measures in the field, that the PPCS would not be in the field at the same time. That it would kind of rotate on a different pattern and would be in the field when we're asking the community measures. But, yes, we do anticipate collecting that in the future.

HEATHER BROTSOS: Thanks, Jenna. Okay. Looking through some of our other questions. Questions about timing of data and things. I think the best advice we can provide at this point is to just keep an eye out on our BJS website and please subscribe to JUSTSTATS for emails that announce the release of reports and datasets as they become available. We don't have specific timing on any of those at this point in time. And then another question related to the split sample in 2024. Jenna, is there more that you can provide about BJS' plans during that split sample as far as what will be collected and what will be done with the data collected from both of those instruments?

JENNIFER TRUMAN: So, the goal, as I kind of mentioned, is to collect both instruments so that we can compare across. So, we want to understand if there are differences between the current and the new instrument. So, having the split sample will allow us to do analysis to compare the two instruments, similar to how it was done back in the 90—the 92-93 redesign. And then it will also allow us to compare the 2023 current instrument to the 2024 current instrument so that we can see if crime is changing over time. And then as we move forward, we'll also have a half sample of the 2024 new instrument to then compare to the 2025 new instrument estimates. So, that's kind of our goal is to be able to not only compare the two instruments to understand the differences and what we're getting, because as David mentioned, we did collect the field test with Westat, with Westat interviewers, and so Census will be our data collector and their interviewers—so we want to be able to test within the Census environment the two. But it's also important, right, for us to be able to say whether crime is changing. So, this allows us to have data from both the new and the current instrument as we move along. So, that's kind of the goal there. And as I did warn, there's a chance—we typically release data in the fall. It's possible that as we continue to analyze this, as we get through this split sample, we might be on a delayed schedule. But we will continue to maintain updates on the instrument redesign website and then as things are released, we put out notices, as Heather just mentioned, in JUSTSTATS.

HEATHER BROTSOS: Great. Well, I think that we have gotten to the end of the questions that have come through the chat and the Q&A. So, I want to thank all of our speakers for incredibly great presentations and a lot of really interesting information. I'll give one last plug for the 50th anniversary celebration of the NCVS. We hope to see you all there, either in person or virtually. As Jenna mentioned, we're going to have some big surprises, so definitely mark your calendars and come on out to celebrate this huge milestone with all of us.

With that, I will thank our speakers and thank Daryl for all of his support throughout this webinar. And if you do have further questions, you can always reach out to us via [askbjs@usdoj.gov](mailto:askbjs@usdoj.gov). And those will be routed directly to our team and we'll be happy to

answer any further questions that you have on this topic or any others related to the NCVS. Thank you so much for taking time this afternoon to hear more about what we're doing. Take care.

DARYL FOX: So, on behalf of the Bureau of Justice Statistics, Westat, and our panelists, we want to thank you for joining today's webinar. This will end today's presentation.