

**FJC Courtroom Use Study  
Final Report Technical Appendices**

**Technical Appendix Nine**

**Independent Observation of Twelve  
Courtroom Use Study Courts**

This page intentionally left blank.

**INDEPENDENT OBSERVATION OF TWELVE  
COURTROOM USE STUDY COURTS**

**November 2, 2007**

**Federal Judicial Center**

*David Rauma  
Project Director*

This report was undertaken in furtherance of the Federal Judicial Center's statutory mission to provide research and planning assistance to the Judicial Conference of the United States and its committees. The views expressed are those of the author and not necessarily those of the Judicial Conference, the Committee, or the Federal Judicial Center.

This page intentionally left blank.

## INTRODUCTION<sup>1</sup>

The purpose of this study is to provide an independent check on the data provided by the courts participating in the Federal Judicial Center’s courtroom use study. Independent data collection is a desirable goal for an evaluation, as it establishes a distance between the differing objectives of the evaluator and the organization under study. The scope of the courtroom use study precluded such an approach and required a dedicated data collection effort by the participating courts. To assuage any concerns about the court-provided data, we conducted a smaller, separate study that relied on independent data collectors to visit randomly selected federal courthouses, to observe and record information in randomly selected courtrooms, distinct and apart from the overall court data collection effort.

This observation study was not designed to replicate the level of detail in the court-supplied data. Rather, it was intended to check the courtroom use data by recording basic information in the selected courtrooms. The independent observers visited courtrooms randomly selected from among those with scheduled proceedings and recorded whether the courtroom was in use at any time during their visits. The independent observers’ data can be compared against the study courts’ reported use data for the same courtrooms at the same times on the same dates.

### Observation in the District Courts

These 12 districts were selected at random for observation from among the 26 districts participating in the two waves of the courtroom usage study:

Wave 1	Wave 2
District of Arizona	Central District of California
District of Colorado	Southern District of Florida
Northern District of Georgia	District of Oregon
Northern District of Illinois	Western District of Pennsylvania
Western District of New York	Eastern District of Tennessee
Western District of Oklahoma	Western District of Texas

---

<sup>1</sup> Vashty Gobinpersad provided invaluable assistance in the preparation of the data for this project. James B. Eaglin gave detailed and insightful comments on earlier drafts of this report. And 54 student observers worked diligently to observe and record information in several hundred courtrooms in 12 federal district courts.

## *Independent Observation of Twelve Courtroom Use Study Courts*

The Wave 1 districts collected courtroom use data for the period January 15, 2007 to April 15, 2007. The observation study in the Wave 1 districts began on March 12, 2007 and concluded on April 13, 2007. The Wave 2 districts collected courtroom use data for the period April 16, 2007 to July 15, 2007; the observation study began on April 30, 2007 and ended on July 13, 2007.

In each of these 12 districts, we chose for observation the district courthouse, or courthouses, in the city that is the primary place of holding district court. To do the actual observations, we contracted with law school and graduate students in the selected cities to serve as independent observers. Using information that the district courts had provided to the Center about scheduled courtroom events, we sent the observers to randomly selected courtrooms at predetermined times, for one-half-hour intervals, to record whether the courtrooms were in use. No one in the 12 selected districts was given any advance notice that the observers would be in their courtrooms. We matched the information collected by the observers to the courts' reported use data for the same courtrooms, on the same dates and at the same times. These jointly-reported half-hour periods are the unit of analysis for this study.

### **Summary of Findings**

The major findings of this study are the following.

- When courtroom use is defined as use by a judge, the overall degree of concordance between the independent observers' reports and the court-reported data in the 12 selected districts is 89%.
- When courtroom use is defined as use by anyone, including court staff, attorneys, and judges, the overall degree of concordance is 85%.
- In the Wave 1 study courts, the degree of concordance between the two data sources is 88% for courtroom use by a judge and 83% for courtroom use by anyone. In the Wave 2 study courts, these figures are 89% and 86%, respectively.
- Several districts have somewhat higher or somewhat lower degrees of concordance, compared to the overall averages, but the majority of districts differ little from the overall averages.
- Over time, the results are relatively stable. Week-by-week, from the first week of observation to the last week of observation in each study wave, the degree of concordance on each measure of courtroom use typically falls within 3 to 4 percentage points of the overall average.

*Independent Observation of Twelve Courtroom Use Study Courts*

- The independent observers were varied in their individual degrees of concordance with the court-supplied data, but none were exceptionally low. Put differently, the observers' performance does not appear to have had an impact on the data analysis.

Based upon our findings, we conclude that the courtroom use data provided by the study courts to the Center reliably represents what actually occurred in the courtrooms under study. The next sections contain detailed discussions of the methodology and the findings of the data analysis that lead us to this conclusion.

## **METHODOLOGY**

### **Selection of Districts and Courthouses**

We randomly selected 6 districts from each of the 2 study waves for the independent observation study. The number of study districts is too small for a sub-sample to be representative of the study districts, but random selection ensured that the choice of districts for observation was neutral. Consequently, we will not generalize any specific findings from the randomly selected districts to all study districts. However, we feel confident that our conclusions about the concordance between the independent observers' data and the court-reported data apply to those 14 districts that were not selected for independent observation.

In each of the 12 randomly selected districts, we designated for observation the courthouse, or courthouses, in the city that is the principle place of holding district court. Tables 1 and 2 list the districts, cities, and the courthouses selected for Waves 1 and 2, respectively. In 10 of these 12 cities, there is one United States District Courthouse. Los Angeles, in the Central District of California, has two district courthouses and Miami, in the Southern District of Florida, has three district courthouses clustered together at Federal Courthouse Square. We included all district courthouses in these 12 cities in the procedures used to schedule courtroom observations.

**Table 1**  
**Selected Courthouses, Wave 1**

<b>District</b>	<b>City</b>	<b>Courthouse</b>
District of Arizona	Phoenix	Sandra Day O'Connor U.S. Courthouse
District of Colorado	Denver	Alfred A. Arraj U.S. Courthouse
Northern District of Georgia	Atlanta	Richard B. Russell Federal Building and U.S. Courthouse
Northern District of Illinois	Chicago	Everett McKinley Dirksen U.S. Courthouse
Western District of New York	Buffalo	U.S. Courthouse
Western District of Oklahoma	Oklahoma City	U.S. Courthouse

**Table 2**  
**Selected Courthouses, Wave 2**

<b>District</b>	<b>City</b>	<b>Courthouse</b>
Central District of California	Los Angeles	Roybal Building
Central District of California	Los Angeles	Spring Street Courthouse
Southern District of Florida	Miami	David W. Dyer Federal Building and U.S. Courthouse
Southern District of Florida	Miami	James Lawrence King Federal Justice Building and U.S. Courthouse
Southern District of Florida	Miami	U.S. Courthouse
District of Oregon	Portland	Mark O. Hatfield U.S. Courthouse
Western District of Pennsylvania	Pittsburgh	U.S. Courthouse
Eastern District of Tennessee	Knoxville	Howard H. Baker, Jr. U.S. Courthouse
Western District of Texas	San Antonio	John H. Wood, Jr. U.S. Courthouse

### **Independent Observers**

To observe courtrooms in the selected districts, we used local, independent observers who are not current or former employees of the Federal Judicial Center or the federal courts.<sup>2</sup> Apart from their independence, a primary advantage of using local observers is that they could be readily available to observe in the courts over the course of several months, and be available at a much lower cost than, for example, Center staff who would

---

<sup>2</sup> We were able to recruit only two local observers in Chicago (Northern District of Illinois). To gather additional data in this large court, we sent two observers to Chicago for a week of observation. One of these two observers was a temporary employee of the Federal Judicial Center; the other was hired as an independent contractor for this project. Also, after her observations in Chicago were complete, we hired one of the local observers as a temporary summer intern.

## *Independent Observation of Twelve Courtroom Use Study Courts*

have to be sent to these locations. Another advantage is that the observers would not be known to court staff, nor to judges, and could make their observations anonymously. Judges and staff in the study courts were advised at the beginning of Wave 1 and Wave 2 that the Center would use outside independent observers. The disadvantage is that these independent observers were not expected to have the knowledge and experience to make detailed observations in a federal district courtroom. This limited the amount and type of data that could be collected in the courtrooms by the observers. However, because the primary goal of the observation study was to record whether or not courtrooms were in use at particular times, and by whom, we did not expect that lack of experience would prevent observers from recording this basic information.

We contracted with 2<sup>nd</sup> and 3<sup>rd</sup>-year law students and advanced graduate students to serve as independent observers. Advanced students, whether in law school or graduate school, would have the level of maturity needed for such a task and could have flexible schedules that would allow for court observation during the normal hours of court operation. Where possible, we chose law students because of their additional knowledge of legal matters. We recruited these students by contacting faculty and administrators at universities in the selected cities. The goal was to hire five observers in each city. Table 3 shows that we met this goal in 8 of the 12 selected cities, that there were at least three observers in each city, and that there were a total of 54 observers.

**Table 3**  
**Court Observers, Waves 1 and 2**

	<b>Number of Observers</b>
Wave 1	
District of Arizona	5
District of Colorado	3
Northern District of Georgia	5
Northern District of Illinois	4
Western District of New York	5
Western District of Oklahoma	4
Wave 2	
Central District of California	5
Southern District of Florida	5
District of Oregon	5
Western District of Pennsylvania	3
Eastern District of Tennessee	5
Western District of Texas	5

**Selection of Courtrooms for Observation**

We randomly scheduled observations in all courtrooms in each courthouse, with a few exceptions. We did not schedule observations in:

- Courtrooms identified by the court as unavailable (*e.g.*, under renovation) or unused (*e.g.*, vacant or unassigned);
- Ceremonial courtrooms that were not also assigned to a judge; and

- Courtrooms that did not have conventional room numbers or other identifying information.

Courtrooms without identifying information were excluded for one reason. Without such information, we could not readily direct observers to a particular room or floor of the courthouse. The observers were to make their observations without drawing attention to themselves, and we did not want them asking questions that might, in turn, cause questions to be asked of the observers. As we will describe below, one of the instructions to the observers was, if asked, to be forthcoming and describe why they were in the courthouse or in a particular courtroom. On average, one courtroom was excluded from any given courthouse by applying these three guidelines.

### **Scheduling of Observations**

Court observers received a schedule for each week of observation that listed specific, randomly selected courtrooms and the times at which they were to make their observations in these courtrooms. The schedules were made during the week before the observations were to take place and were based on the court scheduling information supplied to the Center by the courts.<sup>3</sup>

The courtrooms were scheduled for observation in contiguous sets of half-hour blocks, either in the morning or afternoon, during regular courthouse hours. For example, in Wave 1, an observer would be scheduled each week for a three-hour block of time and would visit six different courtrooms for one half-hour each. We shortened this block of time to two and one-half hours for Wave 2 because Wave 2 observations extended over ten weeks rather than the five weeks of Wave 1. The shorter blocks still produced a large number of observations over ten weeks.

The scheduling process was designed to use random selection as a neutral selection process whenever choices had to be made. Thus, the selection of time of day (*i.e.*, morning or afternoon), courthouse (in Los Angeles and Miami), courtrooms, and the order in which the courtrooms were to be visited was randomly determined. Observers were each scheduled for a different day of the week during each week of observation. Thus, if there were 5 observers available, Monday through Friday would each have a block of time scheduled for observation. Ordinarily, we scheduled observations for courtrooms that, according to court-supplied scheduling information, were to be in use during the half-hour observation period. However, on one randomly selected day each week, the courtrooms were randomly chosen without regard for the court schedule. We wanted to have a portion of the observations independent of the court schedule, to possibly capture courtroom use that was scheduled after the observers' weekly

---

<sup>3</sup> The scheduling data were available in DISCUS, a software package that the study courts used exclusively for this project, to record courtroom schedules and courtroom usage. DISCUS and the scheduling and use data are described in detail in Chapter 2 of the courtroom use study report.

## *Independent Observation of Twelve Courtroom Use Study Courts*

observation schedules were completed, or perhaps scheduled at the last minute. Once the observation schedules were complete, we assigned the independent observers to a daily schedule based on their availability.<sup>4</sup>

Under certain circumstances, the random selection of courtrooms and/or the time of day would have to be changed for a given courthouse on a given day. Here are some example circumstances.

- If there were too few courtrooms scheduled for use to fill a daily block of time, observers would either return to a courtroom they had visited earlier or stay in a courtroom for an hour. Observers could also be scheduled for visits to randomly selected courtrooms that did not have events scheduled at the time of the visit.
- If the scheduled courtrooms were scheduled for relatively brief events (*e.g.*, less than an hour), and there were too few of them to fill the selected block of time, the courtroom observations would be scheduled to match the available courtroom schedules rather than on a random basis.
- If an observer's schedule of available times did not fit with the randomly selected morning and afternoon blocks of time, a block of time would be shifted to match the observer's schedule. If possible, the courtroom visits would be rescheduled using the random selection process, but might have to be scheduled to fit the court's schedule for that block of time.

For example, if there were five or six short events scheduled in a courthouse for a selected morning, and those events were spread out over the morning, completely random selection might not fill the observer's schedule. Instead, we would schedule the observations in the order in which the events were scheduled to take place. Tables 4 and 5 show the final distribution of observations in each courthouse in Waves 1 and 2, respectively.

Changes to the randomly selected schedules were more common in the smaller courthouses, and these changes occurred in at least one district every week. It is important to keep in mind that this study is not designed to analyze courtroom use, and that random selection of courtrooms and times is not the same as random sampling. This study is about the degree to which two data sources are reporting the same information. Random selection in constructing the observers' schedules is merely a neutral selection

---

<sup>4</sup> The observations in Chicago (Northern District of Illinois) were collected by a somewhat different procedure. We recruited two local observers and, to supplement their observations, we sent two additional observers to Chicago. Each observer, local or not, did their observations over the course of a single week, observing all day, every day. The courtrooms were selected in the same manner as in other districts, and each observation period was the standard one half-hour in length. One consequence of this procedure was that we were not able to observe in all 44 courtrooms in the Chicago courthouse that met our requirements for observation (see Table 4).

device. One benefit of random selection is that scheduling will not favor, for example, lower-numbered courtrooms, lengthy proceedings over shorter proceedings, or any other courthouse, courtroom, or event-specific features that might exist. In other words, random selection will distribute evenly the scheduling of observers' visits over the available courtrooms. When purely random selection would not work, for reasons listed above, we took care to schedule courtrooms without regard to anything except whether an event was scheduled.

**Table 4**  
**Independent Observations in Wave 1 Districts**

<b>District</b>	<b>City</b>	<b>Courthouse</b>	<b>Number of Courtrooms Observed</b>	<b>Number of Observations</b>
District of Arizona	Phoenix	Sandra Day O'Connor U.S. Courthouse	17	145
District of Colorado	Denver	Alfred A. Arraj U.S. Courthouse	15	66
Northern District of Georgia	Atlanta	Richard B. Russell Federal Building and U.S. Courthouse	21	145
Northern District of Illinois	Chicago	Everett McKinley Dirksen U.S. Courthouse	36	196
Western District of New York	Buffalo	U.S. Courthouse	8	138
Western District of Oklahoma	Oklahoma City	U.S. Courthouse	13	120

**Table 5**  
**Independent Observations in Wave 2 Districts**

<b>District</b>	<b>City</b>	<b>Courthouse</b>	<b>Number of Courtrooms Observed</b>	<b>Number of Observations</b>
Central District of California	Los Angeles	Roybal Building	17	100
		Spring Street Courthouse	30	145
Southern District of Florida	Miami	David W. Dyer Federal Building and U.S. Courthouse	9	71
		James Lawrence King Federal Justice Building and U.S. Courthouse	6	84
		U.S. Courthouse	9	80
District of Oregon	Portland	Mark O. Hatfield U.S. Courthouse	13	219
Western District of Pennsylvania	Pittsburgh	U.S. Courthouse	14	150
Eastern District of Tennessee	Knoxville	Howard H. Baker, Jr. U.S. Courthouse	6	239
Western District of Texas	San Antonio	John H. Wood, Jr. U.S. Courthouse	8	250

**Court Observation Protocol**

Each independent observer received a package of instructions for conducting the courtroom observations. Included in the package was a one-page form that observers were to use to record their observations for each half-hour observation period. The form captured this information about the courtroom visit:

- Information identifying the observer, city, courtroom, and date of visit;
- The observer’s time of arrival at the courtroom;

## *Independent Observation of Twelve Courtroom Use Study Courts*

- Whether or not the observer could enter the courtroom and, if not, why (e.g., courtroom locked, proceeding closed to the public);
- If the observer could enter the courtroom, whether the courtroom was in use;
- If the observer could not enter the courtroom, was it possible to determine that the courtroom was in use;
- If use of the courtroom began during the observer's visit and the time at which it began;
- If the courtroom was in use, whether a judge was present;
- If the courtroom was in use, whether attorneys or members of the public were present and how many;
- If the courtroom was in use, whether that use ended during the observation period and, if it ended, the time it ended;
- If available, the posted description of the event scheduled for the courtroom;
- The observer's time of departure from the courtroom; and
- The observer's notes describing situations that could not be accurately described within the available categories of information or anything unusual that occurred.

Observers were instructed to use their best judgment to determine whether a courtroom was in use. As for what constituted use, their only guideline was that, if a judge was present and presiding over a proceeding, the courtroom was in use; other situations were to be assessed on a case-by-case basis and were to be described on the observation form.

We kept the data collection basic for several reasons. First, a basic data collection reduced the training and background knowledge necessary to complete the task. Second, a basic data collection effort could minimize inter-observer variation in the data collection, and, in turn, reduce or eliminate discrepancies between the observers' recounting of courtroom use and the courts' description of courtroom use. As noted above, it was sufficient for purposes of this study to know whether a courtroom was in use during a given observation period. Anything more could have introduced a source of error into the comparison of observers' reports and court-reported data.

The observers were instructed to draw no attention to themselves, either in dress or demeanor. But if they were asked their purpose for being in the courthouse or in a particular courtroom, they were instructed to be forthcoming and respond that they were

there to observe courtroom usage as part of the Federal Judicial Center's courtroom use study. The observers were also to report any such incidents as soon as possible after their observations were concluded for the day. Fewer than half of the observers reported inquiries by court staff or by judges, but at least one observer in each district was asked why they were in the courthouse.

### **Data Entry and Data Matching**

Observers were required, within 8 hours of completing a set of observations, to enter the information into an Internet database created specifically for this portion of the Center's study. This process of remote data entry served two purposes. It served as a check on the observers' work, but more importantly it created a permanent record of the independent court observations that could not be altered.<sup>5</sup> The observers sent to the Center copies of their observation forms, which were used to check the observers' data entry and resolve any errors or discrepancies in the data.

After the observers' data were checked, these data were matched with the courtroom use data collected by the study courts. The matching was done on the basis of date, time, and courtroom. The courtroom use data supplied by the courts are described in detail in Chapter 2 of the courtroom use study report.

### **Measuring Concordance Between the Data Sources**

We will focus on two measures of concordance between the observers' descriptions of courtroom use and the courts' reports of courtroom usage.

1. To what extent do the two data sources agree whether a given courtroom was being used by a judge during a randomly selected half-hour period?
2. To what extent do the two data sources agree whether a give courtroom was in use for any purpose, either by a judge, staff, attorneys, or others during a randomly selected half-hour period?

There are several reasons for this choice of measures. First, the goal of the observation study was to determine whether courtrooms were in use and by whom, and not to collect detailed information about the nature of that use. The independent observers did not collect the detailed information contained in the court-reported data. Thus, measures of comparison between the data sources are limited by design to whether a courtroom was in use.

---

<sup>5</sup> We used WebSurveyor for the remote data entry. WebSurveyor is designed for Internet-based opinion surveys, but works well for data entry. According to the technical staff at WebSurveyor, the original responses to an online survey cannot be altered by the customer, only copied. WebSurveyor is now marketed by Vovici as EFM Feedback.

Second, determining who was using a given courtroom can be difficult under some circumstances. While use of a courtroom by a judge for a proceeding should be evident to an observer, there can be a variety of other courtroom uses by staff, attorneys, or even outside groups (*e.g.*, a moot court conducted by a local law school) that may not be easy to categorize or perhaps recognize by courtroom observers without an inquiry. Consequently, we will use these two general categories of users—judges and anyone, including judges.

For purposes of this analysis, courtroom use during a randomly selected half-hour period is defined broadly as use, whether by a judge or someone else, for any amount of time during that half-hour period. Nonuse is exactly that—depending on the definition of use, the courtroom was not in use by a judge or not at all during a given half-hour period. In some instances, observers could not enter the courtrooms. If a courtroom was locked or inaccessible for some other reason, the observers were instructed to note that lack of access. Unless the observer could determine by some other means that the courtroom was in use, the courtroom was assumed not to be in use in these situations. As a result, there may have been cases in which some nonpublic use occurred that could not be detected by the observers.

The calculation of the degree of concordance between the data sources is straightforward—the percentage of jointly-reported half-hour periods in which the observers and the courts agree that a courtroom was or was not in use, for each measure or definition of use. Consider the use of a courtroom by a judge. If both data sources report that a courtroom was in use by a judge during a selected half-hour period, they are counted as in agreement for that half-hour period. If both data sources show that a courtroom was not in use by a judge during a selected half-hour period, they are also counted as in agreement. The data sources disagree when one source reports use by a judge and the other does not. Note that if one source reports courtroom use by a judge and the other reports use but not by a judge, the data sources are counted as disagreeing. The other measure of concordance expands the definition of use to any use, including by judges, but otherwise utilizes the same conditions for agreement and disagreement. If both data sources report that a courtroom was in use, whether by a judge or others, or that a courtroom was not in use, they are in agreement about courtroom use by anyone.

The degree of concordance between these data sources has to be interpreted as exactly that and nothing more. It is the degree to which the two data sources are reporting the same thing. It is not a measure of performance by the courts, nor by the independent observers. It is also not a measure of courtroom use. The task of the observers was to provide an independent check on the courtroom use data that was supplied by the courts to the Center, and this study was designed for that purpose alone.

## DESCRIPTION OF FINDINGS

Table 6 contains the degree of concordance between the observers' reports and the court's data overall as well as separately for the two waves of the study. The overall concordance between the data sources is 89% for courtroom use by a judge and 85% for use by anyone. For the first wave of study courts, the figures are 88% for judge use and 83% for any use. These figures rise in the second wave to 89% and 86%, respectively.<sup>6</sup>

**Table 6**  
**Degree of Concordance Between**  
**Observers' Reports and Court-Reported Data**

	Degree of Concordance		
	All Districts (n=2,148)	Wave 1 Districts (n=810)	Wave 2 Districts (n=1,338)
Courtroom In Use by a Judge	89%	88%	89%
Any Courtroom Use	85%	83%	86%

The overall percentage for courtroom use by a judge means that the observers' reports and the matching court-reported data agreed 89% of the time whether a judge was using a particular courtroom during a randomly selected half-hour period. In other words, 89% of the time the two data sources reported the same thing. The same interpretation applies to the 85% figure for court use by anyone, as well for the separate figures for the first and second waves.

Tables 7 and 8 show the distribution of agreement and disagreement between the observers' reports and the court-supplied data, overall, for each wave, and for each measure of concordance. Consistently, for each wave and for each measure of use, where these data sources agree about courtroom use, two-thirds or more of these cases are agreement that the courtroom was not in use. Consider the overall result for judge use. In 63% of the jointly-reported half-hour periods, the observers and the courts agree that the courtroom was not in use by a judge. This compares to 26% of the half-hour periods in which the two sources agree that the courtroom was in use by a judge. This result is consistent for both measures of concordance across Waves 1 and 2. When the two data

---

<sup>6</sup> Our sample of observations qualifies as a stratified sample. In studies such as this, where the strata (courts) are of unequal size (*i.e.*, number of courtrooms), it is common to weight the data before performing any analysis. The weighting compensates for the fact that sampled observations from strata of different sizes represent different numbers of observations in the population. We experimented with weights that took into account the numbers of courthouses, courtrooms, and observers, but the weighted results were virtually identical to the unweighted results. For the sake of parsimony, we report the unweighted results throughout.

*Independent Observation of Twelve Courtroom Use Study Courts*

sources disagree, the most frequent basis for that disagreement is that courts report courtroom use and the observers do not. The figures in Tables 7 and 8 cannot be taken as measures of courtroom use per se. This study was designed for the purpose of comparing two data sources, not measuring anything specific about courtroom use. The larger courtroom use study was designed for that purpose.

**Table 7**  
**Comparison of Independent Observers' Reports and Court-Reported Data:**  
**Courtroom Use by Judges**

	<b>All Districts (n=2,148)</b>	<b>Wave 1 Districts (n=810)</b>	<b>Wave 2 Districts (n=1,338)</b>
Sources Agree: Courtroom In Use by a Judge	26%	27%	25%
Sources Agree: Courtroom Not in Use by a Judge	63%	61%	64%
Sources Disagree: Court Reports Courtroom In Use by a Judge	9%	9%	8%
Sources Disagree: Observer Reports Courtroom in Use by a Judge	2%	3%	2%

**Table 8**  
**Comparison of Independent Observers' Reports and Court-Reported Data:**  
**Courtroom Use by Anyone**

	<b>All Districts (n=2,148)</b>	<b>Wave 1 Districts (n=810)</b>	<b>Wave 2 Districts (n=1,338)</b>
Sources Agree: Courtroom In Use	28%	28%	28%
Sources Agree: Courtroom Not in Use	57%	55%	58%
Sources Disagree: Court Reports Courtroom In Use	14%	15%	13%
Sources Disagree: Observer Reports Courtroom in Use	2%	2%	1%

In the next sections, we will examine these measures of concordance 1) for individual districts, 2) over time, and 3) for the individual observers, to assess whether these results are stable and not the result of outliers or other data problems.

### **Concordance Within Individual Districts**

The measures of concordance for each of the 12 study courts show remarkably similar results. Tables 9 and 10 contain the results for the study courts in the first and second waves, respectively. Among these 12 courts, two courts stand out with lower degrees of concordance on both measures: the Northern District of Illinois in Wave 1 and the Western District of Pennsylvania in Wave 2. The remainder of the study courts cluster together around the overall figures on each measure of concordance, typically differing from that overall figure by 1 to 3 percentage points. The Western District of Texas stands out a bit in Wave 2, with 94% concordance on courtroom use by a judge. Similarly, the Northern District of Georgia has 89% concordance on any courtroom use. Otherwise, these data, accumulated over 5 to 10 weeks by different observers in different courts, produce results that differ by only a few percentage points in the majority of the courts.

**Table 9**  
**Degree of Concordance Between**  
**Independent Observers' Reports and Court-Reported Data,**  
**Wave 1 Districts**

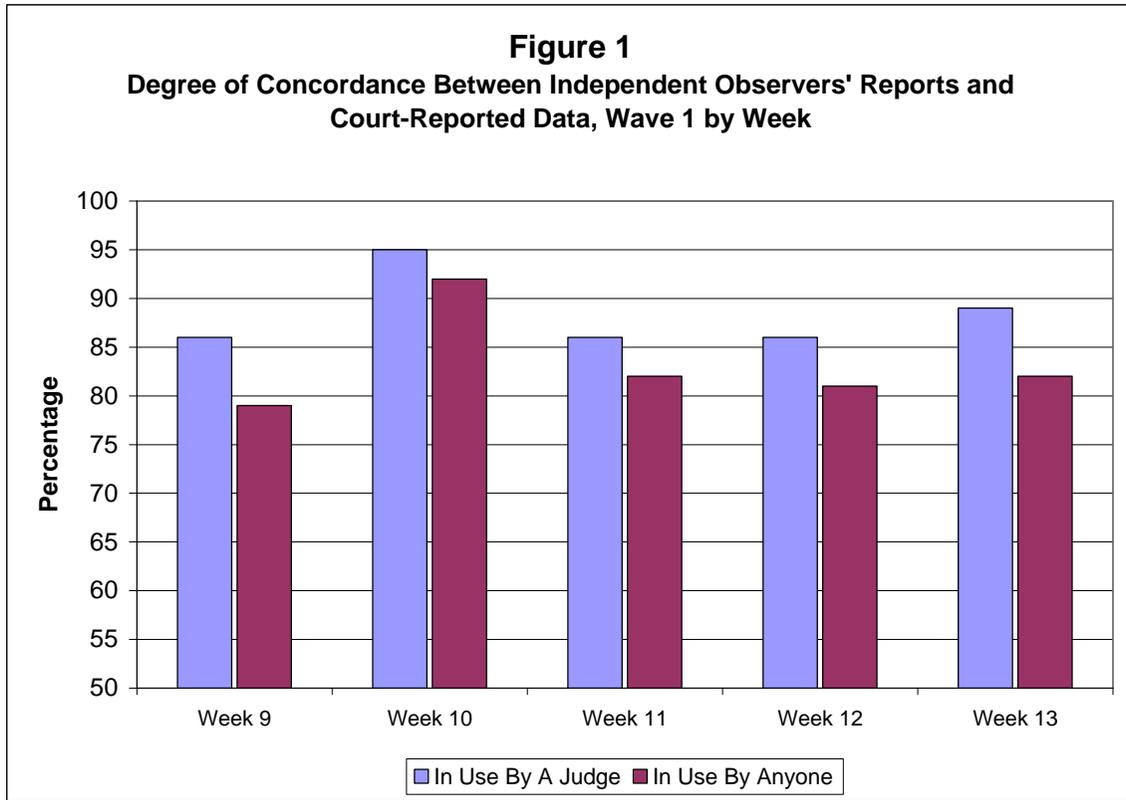
	<b>Degree of Concordance</b>	<b>Difference from Wave 1 Average</b>
District of Arizona (n=145)		
Courtroom In Use by a Judge	91%	+3
Any Courtroom Use	83%	0
District of Colorado (n=66)		
Courtroom In Use by a Judge	89%	+1
Any Courtroom Use	83%	0
Northern District of Georgia (n=145)		
Courtroom In Use by a Judge	90%	+2
Any Courtroom Use	89%	+6
Northern District of Illinois (n=196)		
Courtroom In Use by a Judge	83%	-5
Any Courtroom Use	79%	-4
Western District of New York (n=138)		
Courtroom In Use by a Judge	91%	+3
Any Courtroom Use	86%	+3
Western District of Oklahoma (n=120)		
Courtroom In Use by a Judge	86%	-2
Any Courtroom Use	80%	-3

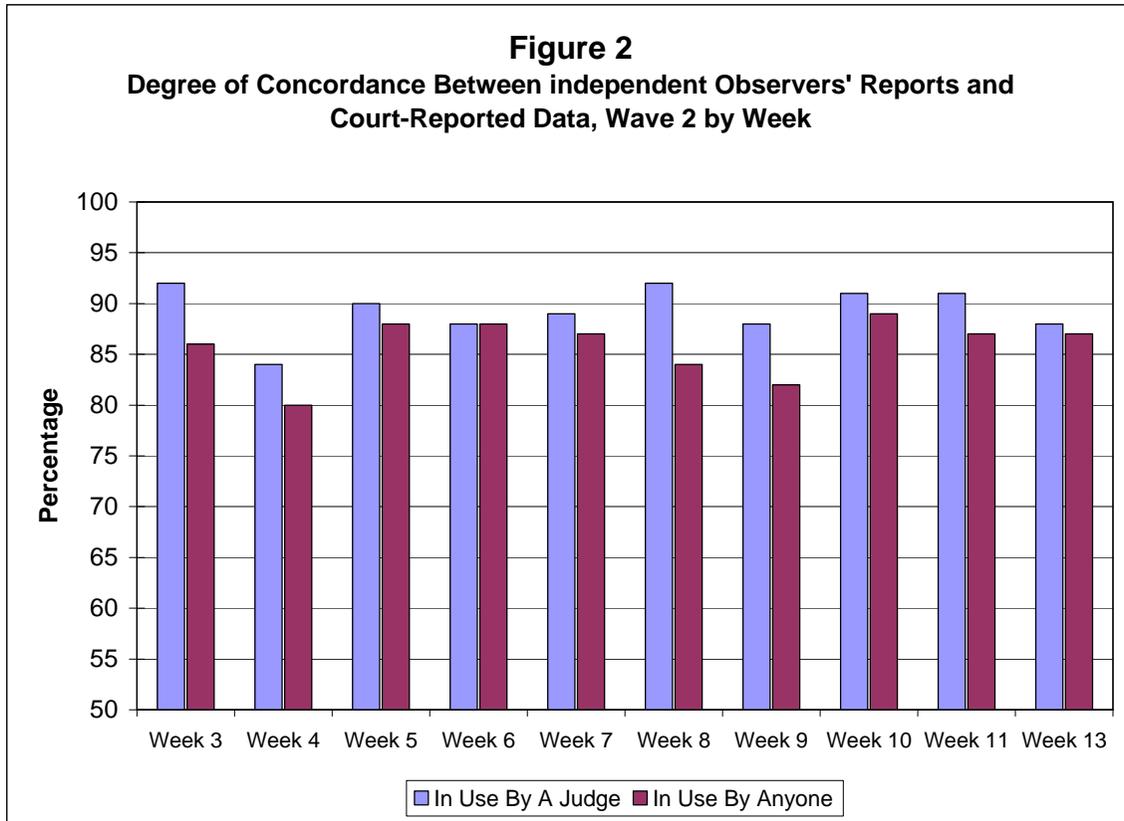
**Table 10**  
**Degree of Concordance Between**  
**Independent Observers' Reports and Court-Reported Data,**  
**Wave 2 Districts**

	<b>Degree of Concordance</b>	<b>Difference from Wave 2 Average</b>
Central District of California (n=245)		
Courtroom In Use by a Judge	89%	0
Any Courtroom Use	88%	+2
Southern District of Florida (n=235)		
Courtroom In Use by a Judge	90%	+1
Any Courtroom Use	87%	+1
District of Oregon (n=219)		
Courtroom In Use by a Judge	90%	+1
Any Courtroom Use	88%	+2
Western District of Pennsylvania (n=150)		
Courtroom In Use by a Judge	80%	-9
Any Courtroom Use	77%	-9
Eastern District of Tennessee (n=239)		
Courtroom In Use by a Judge	89%	0
Any Courtroom Use	86%	0
Western District of Texas (n=250)		
Courtroom In Use by a Judge	94%	+5
Any Courtroom Use	85%	-1

**Concordance Over Time**

Figures 1 and 2 show the degree of concordance by week for Waves 1 and 2, respectively. Note that observations began during Week 8 of Wave 1 and Week 3 of Wave 2, and the horizontal axes of these figures are labeled accordingly. Similar to the comparisons across districts, there is consistency on the measures of concordance, albeit over time rather than across districts.





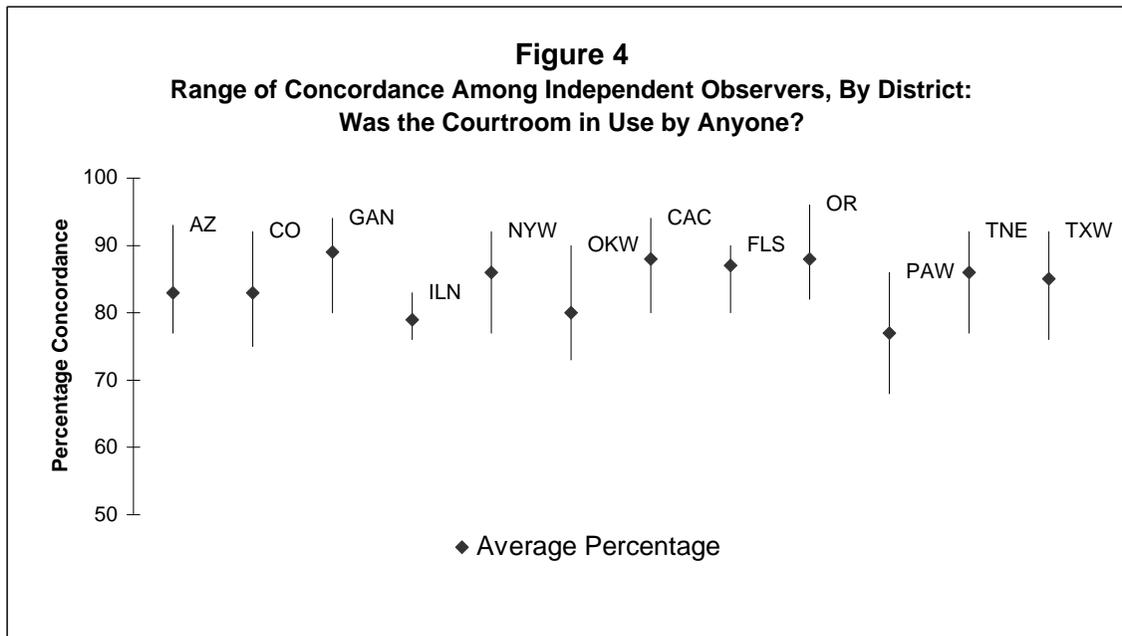
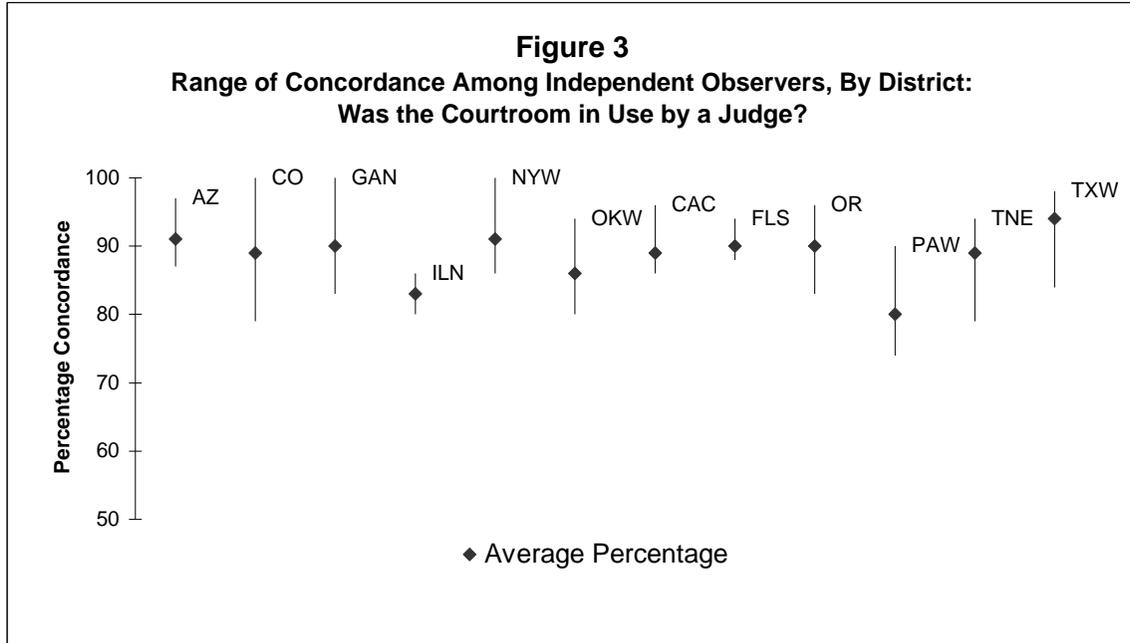
Several features of Figure 1 and 2 stand out. First, a pattern of higher concordance about judge use compared to any use of courtrooms is repeated in all but one week. In Week 6 of Wave 2, the percentages are equal. Second, only the second week of observation in each wave stands out with appreciably higher or lower degrees of concordance. In Wave 1, the percentages are highest in Week 6. In Wave 2, the percentages are lowest in Week 4. Third, apart from these two weeks, the percentages are very stable over time for both measures. If these two weeks are excluded from the calculations of the overall percentages neither measure of courtroom use differs more than 2 percentage points in any given week from the recalculated overall average for each wave.

### Concordance Among the Independent Observers

Measuring inter-observer reliability—the degree to which observers agree in their reports of courtroom use—is impossible, for one reason. The observers were not viewing the same events, which would allow their reports to be compared to one another for accuracy. In other words, there is no standard against which to judge the observers' performance. In lieu of such a measure of reliability, we have plotted for each district the observers' range of concordance within each district for each measure. The results are in Figures 3 and 4.

*Independent Observation of Twelve Courtroom Use Study Courts*

These plots can give insight into how different, or not, the observers were compared to each other within each district.



In Figure 3 the vertical lines represent the difference between the highest and lowest percentages of concordance among individual observers in each district for the measure whether a courtroom was in use by a judge. The district averages, taken from Tables 8 and 9, are also displayed as a reference point on each vertical line. The districts are clustered by wave. Figure 4 shows the same information for the measure whether a courtroom was in use by anyone.

As an example, consider the first district on the left-hand side of Figure 3. There were 5 observers in the District of Arizona, with an average degree of concordance of 91% for courtroom use by a judge. The highest degree of concordance among these 5 observers was 97%; the lowest was 87%. These two percentages are the endpoints of the range represented by the vertical line in Figure 3 for the District of Arizona.

Perhaps the most important features of these graphs are that the ranges, from highest to lowest percentages, are typically in the mid-teens for both measures of concordance, and that there are no extremely low percentages in any of the districts. The District of Colorado stands out in Figure 3 with a range of 21 points between the highest and lowest percentages, but this range may be a function of two things. First, this district had only 3 observers, one of whom had 100% concordance on this measure. Such a small sample is very sensitive to an outlier, represented by this observer. In fact, only 3 of the 54 observers had 100% on the measure whether a courtroom was in use by a judge; none of the observers had 100% on the other measure.

## **SUMMARY**

Using independent observers in a randomly selected subset of the courtroom use study courts, we collected information on courtroom use for randomly selected half-hour periods. These data provided a basis for comparison with data reported by the study courts for these same half-hour periods. Overall, the observers' reports agreed with data from the study courts 89% of the time when courtroom use was measured as use by a judge. These two data sources agreed 85% of the time when the courtroom use was measured as use by anyone, including judges. These results are consistent when the data from these sources are compared according to study wave, district, and over time. While we have no exact measures of their performance, the observers did not seem to be a factor, or at least an important factor, that could skew the results. We therefore can conclude that the courtroom use data provided by the study courts to the Center reliably represents what actually occurred in the courtrooms in the study.