

This assessment covers three very different Local District initiatives, all with divergent mobile technology strategies and length of pilot testing. This report, therefore includes results that are specific to the individual initiatives plus general results of interest. Our review of the survey and workshop results across the three districts reveals much about the way participants used and evaluated the technology's effectiveness and performance. In addition, we were able to analyze data from the central database that provided evidence of improvements in work flow. These results vary according to the different time periods and pilot test situations. However they do provide some guidance for further technology deployment and testing.

Issues in Interpreting Assessment Results

The performance of any technology, especially the ones used in these pilot projects, is a result of a complex mix of human, organizational, and technical influences. The pilot tests were conducted in real field situations that reflect this complexity, not in laboratory experiments with elaborate and rigorous controls. Therefore, the assessment results, though valid and useful, do not answer many possible questions about the causes and implications of the impacts of the technologies. In addition, the number of test participants was small in some parts of the pilot, so as a result, some of the important work situations and associated challenges may not be represented in the results.

Timing was a particular issue in two of the Local District initiatives. In Monroe County, the participants had access to the various devices for such a short time that it is very difficult to distinguish between results due to improved efficiencies versus the disruptive effects of learning a new work method. Other factors include possible resistance to change by some workers or natural variations in workload. The timing problems were exacerbated by delays in deploying devices in Monroe County and the decision to rotate the various devices among the NYC/ACS participants on two week cycles. In addition, there was limited time for training and deployment support in all three Local District initiatives. The timing in the NYC/ACS test was further complicated because it occurred during a period the workers referred to as a "crisis." During the test period the workers were allowed to use paid overtime and were instructed to devote extra effort to reducing the backlog of open cases. As a consequence, it is not possible to separate the possible effects of the new technologies from the effects of these management actions.

Other issues are related to use of the data from the central CONNECTIONS repository. We extracted data about entries by all test participants for the month prior to and during their pilot test period to trace possible technology impacts on the timeliness and reporting work flow for progress notes. Our findings on timeliness and work flow impacts (presented below) include analysis of these data, however, the nature of the data supports only very rough conclusions about technology impacts for these tests. The repository records the timing and types of progress notes entered, but not their length or quality. During the pilot test period, the participants were working on a mix of cases, some open for long periods prior to the pilot test, some started and closed during the pilot, and others remaining open at the end of the test period. Therefore, the notes entered during the pilot test period applied to both new and older cases, ranging from as little as a day to over two months old. The number of notes per case varied widely, as did the types of notes entered. Moreover, the data does not include the ultimate disposition of the cases or any rating of the quality of outcomes obtained. Thus the analysis supports only very general conclusions about timeliness and workflow impacts.⁽¹⁾A more complete evaluation would require a considerably longer test period, some explicit control factors, and more detailed assessment of note quality and case outcomes.

Performance and Work Impacts

The effectiveness of the technologies examined in this assessment depend in large part on their overall capabilities and how well they fit with the users' normal work practices. As the summary in Table 1 shows (below), these technologies differ markedly in how they accomplish two main tasks: (1) converting the users' words into digital text, and (2) entering that text into the central database. Laptop activities are essentially the same as working at an office PC, except for connection and logon procedures. The other technologies involve additional steps to accomplish the analog-to-digital conversion and complete the database entry process. These additional steps add to the complexity of the work and introduce additional points of possible user error or technical malfunction. These use characteristics of the technologies are discussed in more detail below.

Table 1: Comparison of Technology Use Activities

Overall Assessment Results

Device	Data Entry Process		
	Analog-to-Digital Text Conversion	Intermediary Steps or Modifications	Placement into CONNECTIONS
Connected Laptop	type and edit at keyboard	none	simultaneous with typing
Digital Pen	write on special-purpose paper and download through software	compile and edit text	cut and paste
Voice Recognition Technology with Digital Recorder	speak with handheld recorder and download through software	edit text	cut and paste
Telephone Dictation Service	speak with cell phone to third party transcription service	retrieve and edit text	cut and paste

The technologies also differ in their capabilities, both as tested and as developed fully. Those differences are shown in Figure 1 (below).

Figure 1 – Comparison of Technology Capabilities

Figure 1 – Comparison of Technology Capabilities

The comparison is based on the use of the dictation and digital pen technologies in the field without a laptop computer. The data entry, retrieval, and interactivity capabilities are the ones described by the participants in the assessment. The figure shows the much higher overall capability of the connected laptop computer. Without a laptop or desk top PC, the other devices are effective only for the recording step of analog-to-digital conversion in the field. More specifically, in using Dragon Naturally Speaking (DNS) with a digital recorder, the only step that can be completed in the field is recording. Any conversion must take place in an office when connected to a PC. This is also true with a digital pen. The only variation that would extend the capability to perform functions in the field would be to use DNS or the digital pen in conjunction with a laptop with a wireless connection, but these capabilities were not tested in these initiatives.

Laptops

The overall results for laptop use, based on the NYC/ACS experience, are positive in terms of user ratings of device performance and support for increased work output. The case workers used the laptops primarily for documenting and reporting, plus for access to information resources both in the central database and on the Web. It was clear from the survey results and workshop comments that the positive ratings of the laptops were based primarily on three factors:

- using the laptop for reporting and documenting required only small adjustments in work practices, since laptop use is very similar to working with a desktop machine
- connectivity to the central database from outside the office provided new or increased opportunities for completing their reporting and documenting requirements
- access to the central database and other online information resources and email made workers more productive.

It was also clear, however, that case workers distinguished among the various locations for use and preferences among the laptop models. They generally did not use the laptops for recording or working in client contact situations, due to concerns that using the device would interfere with rapport or pose a security risk. The workers clearly preferred smaller, lighter laptop models as well, and recognized differences in connectivity among the models. However, the overall ratings for efficiency and recommendations for future use were quite high and very

similar across the four models tested. Only the tablet PC and Blackberry device received no or low ratings - the tablet PC due to the fact that it was not equipped to convert handwriting to digital text (recognizing or handling handwriting is also beyond the current capabilities of CONNECTIONS), and the Blackberry's restriction to email only.

The evidence presented below on timeliness and work product performance of the NYC/ACS technology suggests that there were productivity gains. However, there are too many other factors involved to attribute these gains solely to laptop use. The evidence on timeliness of progress notes entry (Figure 3), volume of notes entered per month (Figure 4), and clearing of case backlogs (Figure 5 and Figure 6) does point to a positive impact of laptop use. However, the case workers reported that because they used the laptops at home, the work time per day was increased and generated overtime pay. They had also been instructed to make an extra effort to clear backlogs during the "crisis" period. So we can say that the laptops supported more productive time during a work day (or week), but not necessarily that the laptops increased efficiency.

Dictation and Digital Pen Technologies

The dictation-based and digital pen technologies tested in Monroe and Westchester counties share sufficient characteristics and use issues to be considered together. Both rely on multi-step processes during which the devices transform the users' analog input (speaking or writing) into text in a digital format. The user then accesses the digital text files for review and cuts and pastes the text into the central database at a desktop PC. These could be called semi-mobile technologies, since they include small transportable devices for mobile analog input, but rely on desktop devices to complete the work process.(2)

The Monroe County results, involving both the voice recognition system and digital pens, reveal the close relationship among work practices, work situations, and technology use. The roughly one third of participants that used the digital pens found them to be useful for the following:

- note-taking in the field;
- other reporting tasks and occasional input for word processing;
- taking meeting notes that could be easily stored in electronic files, eliminating scattered paper notes;
- possibly completing locally designed forms, to be developed on the specialized paper.

Workers who preferred writing, used the pen in work situations where writing was an established and natural act. They did not, however, generally report pen use as particularly efficient or adaptable for a wide range of tasks

The main efficiency problems with pens appeared to result from the cumbersome process for transferring the text files created by the pen system into the central database, often requiring many intermediate steps. Some pen users also reported inaccuracies with writing recognition and some dissatisfaction with using them for long narrative notes. Pen users also expressed some concern for the cost and availability of the special paper needed for pen use. The users reported good results when the work situation, worker's preference for writing, and the pens unobtrusive nature aligned, however overall ratings of digital pens were not highly positive.

The telephone dictation system received a mix of high and moderately positive ratings and did contribute to increased work output in the following ways:

- users rated the system moderately efficient and usable;
- most users recommended keeping and expanding deployment of the service;
- several users found the cell phones valuable for other work-related tasks.

The workers also reported problems of poor cell phone connectivity and the complexity of data entry (cut & paste, etc.), which may have tempered the overall ratings. Some workers found dictation difficult and thus slower than typing, which could be mitigated by additional training. Problems with cell phone connectivity were also common, and some workers reported delays in the system processing of their dictation. These difficulties may account for the apparent drop in timeliness for the Westchester pilot seen in Figure 3. However, other measures of timeliness for the three initiatives presented in the following section show different patterns.

Dictation methods presented the same multi-step issues for moving the notes into the central database. Both the transcription service and Dragon voice recognition systems required users to cut and paste the text generated by the system into their reporting database, in one case from the transcription service Website and in the other from the PC running the Dragon application. In principle, the Dragon system could be used to dictate directly into the reporting database. But that would only be possible at a desktop PC with connections to the database, since no laptops with the Dragon system were deployed in the Monroe County test. In this context, the Dragon system by

itself is not a mobile technology. The digital recorders intended for mobile recording were not deployed in time to include in this assessment.

The Westchester County pilot test, with slightly over 13 weeks of dictation system use, provided an opportunity to examine how learning to use a new technology might have affected workflow. The other pilot tests did not include data for a sufficient length of time with each new device to show much in the way of learning effects. If there was a learning process affecting progress note production in Westchester, plotting the number of notes entered each week would show a gradual increase, with relatively lower production in the early period, then the volume of notes accelerating as skill and familiarity increase. In Figure 2 (below) we see little evidence of such learning, with an uneven pattern indicating little systematic variation in progress note entry from week to week, though there is some acceleration in the last week. This may be the result of a slow learning process, showing up only in the last week or two. However, it seems more likely from the user's comments that adaptation was very rapid for those that found the system easy, and much slower or nonexistent for those that essentially rejected its use altogether. The combination of these two effects would result in the same pattern.

Figure 2 – Number of Notes by Week of Pilot - Westchester County

Figure 2 – Number of Notes by Week of Pilot - Westchester County

There is some indication that the types of progress notes entered during the Westchester pilot were different from the pre-pilot period (see Figure 7 below). Our analysis treated all progress notes alike, so the trend in the above Figure may miss important productivity or learning effects.

In addition to use for dictation, some workers employed the cell phones for other work activities. They reported using the phones for setting up client meetings, following up voice mail while in the field, contacting the office for needed information, and even as a date book and appointment calendar. For these workers, an apparently good fit of a mobile technology's capabilities with a worker's needs and preferences can be a very effective combination. By contrast, other workers used the same phone for virtually no dictation or other tasks.

The test of the Dragon dictation system in Monroe County provided some data about the system's effectiveness and use, but none about its mobile potential, since the digital recorders were not deployed in time to include them in this evaluation. We concentrated our evaluation instead on the evidence for efficiency gains, though the time period was too short for firm conclusions. The participants' ratings of their work on progress notes during the pilot were uniformly lower than for the pre-pilot period. They also rated the Dragon system in the low ranges for improving efficiency and versatility. The timeliness and progress note volume data show virtually no difference between the pre- pilot and pilot period (Figure 5 and Figure 6 below). A longer period of use will be necessary before the effectiveness of the Dragon system can be adequately assessed.

A longer time period will also allow for adjustments to the use problems described by the workers. These included the complexity of the Dragon application itself, the need for additional training, lack of experience and skill with dictation, and the problems of using a dictation system in the office environment. Training and additional experience may mitigate the complexity and dictation skill problems. However, dictation with the Dragon system constrains multi-tasking (typing while on the phone, etc.), which was described as a common work practice. The open office environment also means dictation can be disrupted if the system picks up extraneous noise. The sound of many users dictating simultaneously may also be distracting and disruptive to other work. These problems appear to be reflected in the participants' low overall satisfaction with the system. They did see value, however, which was reflected in majority of testers who recommended continued use.

Evidence for Overall Work Flow Impacts

The data extracted from the central database provided some insight about three questions related to possible technology impacts: (1) Does timeliness of reporting change? (2) Are workers more or less productive with respect to progress notes and reporting? and (3) Does technology use affect the kind of reporting activities undertaken by the CPS workers? We were able to find some partial answers to these questions.

Timeliness of Progress Note Reporting

The indicator of timeliness we used was the number of days between an event and the entry of progress notes concerning that event. Progress note records in the central database include the date a note is entered, the type

Overall Assessment Results

of note, the related event date, and other case information. We focused on the number of days between each note and its related event. We then plotted the percent of all notes entered for each district by days from the related event. The higher the proportion of notes in the earlier days, the more timely the overall reporting process. The results for all districts, both prior to and during their pilot tests are shown in Figure 3 (below).

The differences in timeliness across the three districts fit well with our understanding of the pilot test situations. The pre-pilot versus pilot indicators for Monroe County are essentially the same— apparently no substantial technology impacts on timeliness. This would be expected, given the very short time and incomplete deployment of the technologies for this pilot test. There is an approximately 10 percent improvement in timeliness for the NYC/ACS, which is consistent with the reported effectiveness of the laptops used and the “crisis” period management instructions to the workers. There is a decrease in the timeliness indicator for Westchester County, which is consistent with the mixed reports on effectiveness and acceptance of the telephonic dictation technology, as well as with technical problems reported with the system and cell phones. After the first two days following an event, the timeliness indicators for the three districts are much closer together and are virtually indistinguishable over the longer term.

Figure 3 – Percent of Notes Entered by Day After Event

Figure 3 – Percent of Notes Entered by Day After Event

A different indicator of timeliness was also examined: the average elapsed time, in days, between event and progress note entry. This measure shows a different pattern altogether. The average number of days between an event and progress note increased for all three districts: from two to four days for Monroe, from two to over six days for NYC/ACS, and from two to over nine days for Westchester. For NYC/ACS, this rather large difference may be the result of what appears to be time devoted to completing documentation for older cases during the pilot period (described below). This does not, however, apply to the other districts. The Westchester County participants entered higher volume of notes during the pilot periods, compared to the pre-pilot period (described below). That larger volume may also include cases with much longer delays between event and note entry, which would raise the average event-to-note delay. There are also outliers—entries that occurred many days after the event—that have a disproportionate effect on the average.

Volume of Progress Notes

Possible work flow effects of the technologies appear in the different rates of progress note entry. There was a marked increase in the number of progress notes per month during the pilot test periods for NYC/ACS and Westchester, shown in Figure 4 (below). For both of these districts, the increase is substantial and not likely attributable to seasonal fluctuations in case load. The ACS increase may be caused in part by the “crisis” situation as well as by the technology use. It is also possible that the so- called Hawthorne Effect may account for some of these large increases in progress note entries.⁽³⁾ The difference for Monroe County is too small to be meaningful, as would be expected for the short pilot time period.

Figure 4 – Progress Notes Entered Per Month Prior to and During Pilot

Figure 4 – Progress Notes Entered Per Month Prior to and During Pilot

Age of Cases and Backlogs

The increases in efficiency reported for the technologies may provide the workers with the capacity to reduce case backlogs. If so, the age of cases closed during the pilot period should be higher, on average, than during the prior period. So we compared the proportion of cases closed before and after the 60 day closing requirement for the pre-pilot v. pilot test period. Those comparisons are shown in Figure 5 and Figure 6 below.

Figure 5 – Age of Cases Closed Prior to Pilot

Figure 5 – Age of Cases Closed Prior to Pilot

Figure 6 – Age of Cases Closed During Pilot

Figure 6 – Age of Cases Closed During Pilot

Only the ACS data show evidence of backlog clearing, with the proportion of older cases closed during the pilot at more than 65%, compared to just over 40% in the prior period. Just the opposite shift appears in the Monroe and Westchester data (Figure 6). The increase in the newer case closing in Monroe and Westchester can be seen as evidence of productivity increases as well, since the actual rate of case closing increased for these two counties during the pilot.

The percentages shown above are useful for comparison across districts, but they do not tell the whole story. The NYC/ACS case load during the pilot test period included a much larger number of cases open over 60 days (261), compared to 83 in Monroe and 66 in Westchester. Therefore, we would expect to see a substantial increase in older cases closed in the NYC/ACS district, given the combined impact of a substantial backlog to work on, new technology, and increased incentives and resources.

Types of Progress Notes

New technology is seldom neutral with respect to work practices. We would therefore expect the introduction of these new technologies to result in changes in what kind of work is done, as well as the speed or quantity. The descriptions of work impacts in the three districts provide a qualitative picture of some impacts. The data from CONNECTIONS shows some differences in types of impacts as well for Westchester and to a lesser degree for NYC/ACS. The patterns for the three districts are shown in Figure 7 (below). The figure shows the proportion of four kinds of notes: (1) attempted contacts, (2) contacts, (3) collateral contacts, and (4) summary notes both prior to and during the pilot test periods. These are only a few of the many types of notes, but the numbers of other note types were too small for meaningful comparisons.

Figure 7 - Percent of Notes Entered by Type-Prior to and During Pilot

Figure 7 - Percent of Notes Entered by Type-Prior to and During Pilot

The main technology impacts appear to have occurred in Westchester County. The increased number of notes seen for Westchester in Figure 4 above appears to come predominately from increased numbers of field contacts as well as increased documentation. There was a proportionate reduction in collateral notes. This could be the result of increased time for field work due to the mobile devices, reducing the need for travel to and from the office for note entry. In contrast to Westchester County, the ACS pilot period differed less compared to the prior month. There were proportional reductions in contact notes relative to collateral notes, which is likely related to the large increase in work on older cases in NYC/ACS. As expected, the Monroe County comparisons show very little pre-versus post pilot period differences. Larger differences may emerge over a longer test period.

Understanding Technology in the CPS Casework Setting

Though the technologies examined here differed widely, there are some common themes that help explain the ways technology can fit in and enhance CPS casework. We found these themes useful in interpreting the detailed data from each of the three districts. The themes are described here to assist in the district descriptions that follow and in future choices about technology deployment in these settings.

The Two Components of Mobility

For the purposes of working remotely in CPS, two work components are needed: a device with easy portability and connectivity with ubiquitous access. For instance, the voice recognition software is not a mobile device on its own, however, it becomes mobile if used in connection with a laptop or digital recorder. Also, a telephonic dictation service is essentially a mobile technology, but only when used in combination with a cell phone with excellent network coverage. If the two components are not present, the potential value of the mobile technology will not be obtained. This is particularly so for laptops that do not have reliable wireless access or are not set up for a data synchronization process. Deriving the most value from mobility appears to require mobile devices with ubiquitous access that provide real time entry into the central application.

Digitizing and Entering Information

Entering analog information (i.e., text from workers' paper notes or thoughts) into CONNECTIONS can require anywhere from one to as many as four or five separate steps. Different mobile devices require different

sequences, that may or may not fit well with the worker's skills, preferences, or work situation. Using a telephonic dictation service, for instance, involves at least three steps: speaking the progress notes, retrieving the transcribed notes from a Web site, then cutting and pasting them into CONNECTIONS. In some cases, the workers reported extra steps of writing out the notes in advance by hand to prepare for the unaccustomed task of dictating. The use of a digital pen is similar. After writing notes on special paper, the pen is put on a docking station (usually at a personal computer), the handwritten notes are digitized and presented in multiple MS Word files, which must then be merged and cut and pasted into CONNECTIONS. The only one-step process in the pilot was the connected laptop, where notes were typed directly into CONNECTIONS, thus digitizing and entering data occurred simultaneously. However, even in these cases, establishing the access connection required additional steps. These considerations illustrate how productivity is affected by the critical link between the nature of the technology and the users' skills and work practices.

Skills and Work Practice Compatibility

Not all caseworkers are comfortable working the same way. Some are more adept at dictating notes, some prefer typing, and others writing. Although all three methods can produce the same note, the work process differences are important. Speaking progress notes in complete coherent sentences comes naturally to some and is very difficult for others. Those who prefer typing describe it as a continuous editing process where cohesive thoughts are formed in the process of revising typed documents. Thus, asking people to change the way they work requires giving them time to learn and achieve proficiency in the new skill and recognition that this can engender substantial resistance.

Mobile technologies must also fit well with individual work practices to produce value. Some caseworkers believe the only way to pay the most attention in a home visit is to rely on handwritten notes; typing can be done after the visit. This belief shapes the choice of and the location where the technology is used. For other workers, a laptop is considered a status symbol that can interfere with establishing rapport with some clients. Typing may be the preferred method for data capture, but the caseworker may not want to bring a laptop into the house. Another caseworker may want to dictate notes but does not want to do so between home visits, preferring instead to wait until there is a large block of time to dictate continuously, documenting all the day's visits at once. These individual skills and preferences shape the use, and thus the potential value of a particular technology.

Work Activity Goals

The goals for work to be done in the field will help determine the appropriate technology to support field work. Defining **what** work activities are to be completed in the field will influence the choice of mobile technology. Some work activities are simple data entry like recording progress notes and safety assessments. For this kind of work laptops, dictation services and systems, and digital pens are all appropriate. But completing additional tasks such as sending and receiving email, reviewing new cases, and adding and relating persons in the database also require access to CONNECTIONS and other Web resources. Obtaining these broader capabilities requires a laptop with mobile connectivity.

Environmental and Contextual Factors

Mobile technology use is strongly influenced by the work environment and context. Workers who rely on public transportation say they have no private place to dictate or even use a laptop, while those who use cars can use both technologies while traveling. For possible use in public transportation, a laptop must be small, have a privacy screen, and have good wireless connectivity. Dictation, by contrast, cannot be adapted to public transportation but could be done in a private courthouse location, a park, or at home. Different technology configurations require different lengths of available time for efficient use. Several home visits in one day may not allow any blocks of time large enough to complete work nor an environment that provides enough privacy for dictating or typing. Time between visits may be too short to complete work activities, whereas large blocks of time waiting in court can be very productive. Physical comfort will also shape technology use. If a laptop is too heavy or big, it will not be used. A digital pen that is uncomfortable to use for long periods will not be used. Finally, if having to carry or use a laptop in an obvious way puts caseworkers at increased personal risk, then it will also not be used.

Summary

This pilot test was initiated to help achieve an important goal: improving protection for children at risk by increasing the productivity and effectiveness of protective service workers in the field. The pilot testing began in a

situation of much uncertainty about what might turn out to be the most effective mobile technologies for child protective service workers, due to a lack of research in this area. By participating in this testing and assessment, the OCFS and the Local District initiatives have added much to what is known about using mobile technologies and how to improve future technology testing, selection, and deployment. The new knowledge about mobile technology's effectiveness can be summarized in three key observations:

1. The effectiveness of any mobile technology strategy appears to depend on a combination of worker preferences, work practice demands, the capabilities of the various devices and systems deployed, and organizational support. No one technology strategy will be a good fit for all CPS workers. Some provision for individual variation should be a part of future strategies. And much attention is needed to training, technical support, and adaptation of work policies to support the mobile CPS worker.
2. In spite of numerous technical and organizational challenges in the deployment and use of these technologies, the testers most often rated them as effective and, with few exceptions, recommended continued deployment. The testers also reported sophisticated and nuanced assessments of the strengths and weaknesses of the various devices. One clear lesson is that the users' experiences and judgments must be an important part of any future tests and decisions about wider deployment.
3. The two mobile strategies that had the longest tests—the telephonic dictation system and connected laptops—showed evidence of improved timeliness in one case and overall greater work productivity for both. However, the limitations in available data about progress notes and work outcomes means that further data collection and analysis is needed before firm conclusions can be made about their effectiveness.

The promise of mobile technology to improve child protective services seems clear. But rapid progress toward achieving that promise will require significant attention to training and change management, continued investigation, and recognizing that the success of any new technology depends on human and organizational conditions at least as much as the devices and systems themselves.

(1) To compare the pre-pilot and pilot test periods it was necessary to assume that the two periods were the same with respect to the kinds of cases involved, the distribution of actions required for the cases, and the overall rates of cases opening and closing.

(2) The digital pens, Dragon Naturally Speaking voice recognition system, and the Santrax dictation service could all, in principle, deliver their digitized text to a laptop in the field, for entry into the central database from any location with connectivity. Dictation could also, in principle, be used to input directly to the central database, however none of these alternatives was tested during this pilot test period.

(3) This effect is caused by the tendency of workers who know they are being observed or involved in an experiment to perform at a higher level, regardless of other circumstances.