FI Tracker:
A Tool for Displaying and Monitoring Survey Interview Paradata

Joe Murphy, Susan Myers, and Barbara Bibb

RTI International (a trade name of Research Triangle Institute)
Contact: jmurphy@rti.org

Acknowledgment

Too Much Paradata, Not Enough Time…

Most field, phone, and web surveys have available rich data on call circumstances – the characteristics of each attempt to contact or complete an interview with a respondent. These "paradata" are typically stored in a record of calls (ROC) datafile. Paradata can be utilized in many ways, including the development of data collection strategies to maximize response, quality, and efficiency. However, data collection managers are often ill-equipped to make operational decisions based on this wealth of data.

Exhibit 1 shows an example of a typical ROC file. There is one record for each call attempt made to a case. The first column includes the ID number for the case. The second column shows the ID of the field interviewer (FI) who made the call. The "Event Code" column shows a disposition code for the case (e.g. 320 = appointment made, 321 = noncontact). The "EvtNote" column shows the FI's notes for the call that were entered into the ROC. Finally "EvtDttm" provides an automatic date and time stamp for each individual call.

Exhibit 1. Example ROC

	A	B	C	D	E
1	**CaseId**	**FIID**	**Event Code**	**EvtNote**	**EvtDttm**
2	00005003	UNCMJAM1	999	Created by CMS after case import	7/23/08 5:51 PM
3	00005003	UNCMJAM1	330	appt 7/28 @ 630	7/28/08 11:50 AM
4	00005003	UNCMJAM1	321	Respondent would not be at home until 930 per his wife	8/2/08 7:38 PM
5	00005003	UNCMJAM1	321	Called the respondent, he said he would be home in 30 minutes	8/2/08 7:39 PM
6	00005003	UNCMJAM1	321	family member said r is still not at home, called cell, no answer	8/2/08 7:40 PM
7	00005003	UNCMJAM1	321	waited for the respondent on his block, did not return home	8/2/08 7:41 PM
8	00005003	UNCMJAM1	330	appt on 8/9 @ 6pm	8/4/08 10:45 AM
9	00005003	UNCMJAM1	320	Respondent was not a home, I called him and he said he was at the beach. He said he forgot although, I had called him the night before. He said to call next week and he was very sorry	8/9/08 7:47 PM
10	00005003	UNCMJAM1	330	appt 8/17 @ 6pm	8/16/08 6:01 PM
11	00005003	UNCMJAM1	320	Called to remind of the appointment	8/18/08 7:43 AM
12	00005003	UNCMJAM1	320	Went to respondent's home for appointment, no one home. I called and the respondent did not answer phone. I waited, then left the sorry I miss you card	8/18/08 7:44 AM
13	00005003	UNCMJAM1	888	Created by CMS after case got exported	9/11/08 6:02 AM
14	00005003	UNCMJAM1	980		9/16/08 4:57 PM

The potential uses for ROC data, coupled with the underutilization of paradata for operational survey decisions led us to develop a method for more efficient display and interpretation of call record information. To this end, we developed a tool called *FI Tracker* that converts raw ROC data into an interactive tool for data managers and supervisors.

Our tool begins with the ROC file of events for all cases with associated time stamped notes. In most data collection systems, this file is saved in a raw text or delimited format. We use SAS software to process the ROC and "flatten" all records associated with a single case into a single record. Numeric codes in the ROC are translated into lettered codes for easy understanding and data are further processed to display by day and time of day. Formatting produces color codes events and special cases which might require scrutiny by data quality reviewers. The end result is a single interactive Microsoft Excel chart for each interviewer showing all of his or her work in a single view. We chose Excel as the software to display *FI Tracker* because it is commonly available for those needing to review field work and allows for copying and pasting of relevant data. This tool greatly increases the efficiency and objectivity with which a review can evaluate the work of interviewers during the course of data collection. Further details on the programming of the *FI Tracker* can be found in Murphy et al. (2006).

Data

For the examples in this paper, we use data from the Community Advantage Panel Study (CAPS) which is a longitudinal survey studying accumulation of wealth and social capital. The survey follows new home owners of low to moderate income along with a matched sample of renters. For the cases included here, FIs made calls to households and could also arrange appointments by phone. The data in our examples come from the sixth round of the survey from 2008; some values were suppressed or masked to protect respondent confidentiality. The field component of the 2008 survey included over 2,000 respondents, 50 FIs, and three field supervisors. For more details on the CAPS survey, see Riley, Ru, and Quercia (2009).

***FI Tracker* Examples**

In this section, we present examples of *FI Tracker* using CAPS data and specific scenarios where it can be utilized. Exhibit 2 shows a partial view of what the reviewer sees upon entering the system. The name of the study is at the top of the chart, followed by the instruction to click on the interviewer ID to open his or her individual report. For each interviewer, the total cases work and rates of any noncontact or refusal and completion rates are displayed. The data for each of these summary statistics come directly from the ROC and only represent examples of what can be displayed in *FI Tracker*. It is customizable to projects' needs as long as the source of the information resides in the ROC. The All Interviewers View allows one to make comparisons across interviewers and to identify interviewers that may require extra review. For instance, interviewer UNCKBUR1 has a pending refusal rate of 14%, highest among these interviewers. The *FI Tracker* allows for drilling down into the case data to determine whether there might be something about this interviewer's work patterns or approach that contributed to this relatively high refusal rate.

Exhibit 2. *FI Tracker* **All Interviewers View**

FI Tracker: UNC Community Advantage Panel Study

Click on Interviewer to Open Report

Interviewer	Total Cases	Non-contact		Refusal		Complete	
		n	Pct	n	Pct	n	Pct
UNCGMCD1	42	24	57%	4	10%	37	88%
UNCJPAD1	32	24	75%	1	3%	25	78%
UNCJSTA1	20	7	35%	2	10%	13	65%
UNCKBUR1	43	28	65%	6	14%	32	74%
UNCLATK1	36	23	64%	1	3%	32	89%
UNCLHAL1	27	14	52%	2	7%	23	85%

Next, we present the single interviewer view which can be accessed by clicking on the interviewer ID. For this example, we will look at the data for interviewer UNCLHAL1. The single interviewer view is shown in Exhibit 3. At the top of the screen is the study name, a link to return to the previous screen, the interviewer's ID and total cases worked, and several color-coded letters referring to case dispositions. For example, an orange N denotes a non-contact disposition, a purple R denotes a refusal, and so on. Below the legend, in the first column, appears a list of CASEIDs, one for each case worked by this interviewer. Next to that are summary and case-level counts of pending non-contacts and refusals, and completion. Again, these statistics can be customized to show anything captured within the ROC. For instance, we could add a breakoff field to show how many breakoffs an interviewer had overall and for each case being worked. To the right of these codes is a timeline of the data collection field period with a column for each half-day in the period. It starts with the first day of data collection, in this case Thursday July 24, and each day is split into daytime (before 4 pm) and nighttime (after 4 pm). These splits are also customizable for projects.

Exhibit 2. *FI Tracker* Single Interviewer View – UNCLHAL1

UNC CAPS FI Tracking Sheet
Click here to return to List of Reports
FI = UNCLHAL1 27 Cases
N = non-contact
R = refusal
A = appointment
C = complete
F = final incomplete
O = other
B = breakoff

CASEID	N	R	C	Thu 07-24 Day	Thu 07-24 Night	Fri 07-25 Day	Fri 07-25 Night	Sat 07-26 Day	Sat 07-26 Night	Sun 07-27 Day	Sun 07-27 Night	Mon 07-28 Day	Mon 07-28 Night	Tue 07-29 Day	Tue 07-29 Night	Wed 07-30 Day	Wed 07-30 Night
	52% 14	7% 2	85% 23														
00003676	2	0	1									N					
00003762	0	0	1							C							
00003763	0	0	0														
00003767	0	0	1	A													
00003811	0	0	1														
00003812	0	0	1													A	
00003881	1	0	0														
00003883	1	0	1														
00004029	2	0	1	N								N,N					
00004030	1	0	1	N								A					C

Within the cells of the tracking sheet are the call outcomes for a particular case for a particular day and time. Looking at case 00004030, we see an N in the Thursday July 24 Day column, indicating one call made to this case before 4pm on this day, resulting in a non-contact. Moving to the right, or forward in time, we see the second call occurring during the evening of Monday July 28, resulting in an appointment. Going even further to the right, we see a final call made to this case on the evening of Wednesday July 30 with an outcome of complete. As shown in the case directly above this one (00004029), if there were two calls during a half-day, both calls would appear, separated by a comma.

The Single Interviewer View provides the history for this case 00004030 conveniently coded on a timeline in the context of all other cases worked by this interviewer. This would be useful to a reviewer of the interviewer's work to determine the approach taken towards call timing and frequency. To obtain more about the specific circumstances of this case, the user can click on the CASEID and access the raw ROC and read through the call notes for this case. The call notes for case 00004030 are presented in Exhibit 4. This shows the date the case was initialized in the system, followed by the first call attempt ending in a non-contact and its associated call note, and so on for the appointment, completion, and assignment of final disposition within the case management system.

Exhibit 4. ROC Call Notes for Case 00004030

UNC Renter Record of Call Notes
Click here to return to Main Report
caseid=00004030

Event Date	Code	Note
22JUL2008:09:46:44	999	Created by CMS after case import
24JUL2008:12:59:06	320	Called number in file and checked with address operator to make sureit was the right number. It is the correct number. No voicemail. Will callback this evening.
28JUL2008:17:22:29	330	Called and spoke with respondent. Scheduled appt. on 8/5 at 4p.
29JUL2008:16:05:46	491	Created by CMS after event/status code was modified by interview software
08AUG2008:08:30:16	980	

To learn more about this interviewer's characteristics, experience, contact information, and other information, clicking on his or her ID from the Single Interviewer View will pull up the Interviewer Demographics View. These are the only data in FI Tracker not originating from the ROC, but a separate project database with interviewer demographic data. However, the data are simple to merge with the ROC by interviewer ID. The Interviewer Demographics View is show in Exhibit 5.

Exhibit 5. FI Demographics View

FI Demographics
Click here to return to Main Report

FIID	UNCLHAL1
FIFirstName	Deborah
FILastName	Gale
FSLastName	Johnson
Assignment	GA
Bilingual	N
Interviewing__Experience	Y
Phone	555-123-4567
RTI_Experienced	Y
Region_	200
State	GA
Traveler	Y
UNC_Renter_Experienced	N
Address	45 Taylor Road
City	Springfield
ZIP	30001

Below we present one more example from the Single Interviewer View. Interviewer UNCNNOD1 shows some immediate differences with UNCLHAL1. This interviewer appears to have made many more calls early in the data collection period, but also had many more pending noncontact dispositions. Apparent are some cells with yellow highlighting; we call these yellowlight cells. These are specific scenarios that can be flagged where a reviewer may want to pay special attention to an interviewer's work pattern. The example here has three noncontact dispositions within the same half day. This might suggest a suboptimal strategy on the part of the interviewer because if someone is not at home during a short period, it does not necessarily make sense to repeatedly try that case during that time. Another example of a yellowlight would be a pending refusal immediately followed by a complete within

the same half day. Such a pattern is unexpected and may raise flags about the interviewer's work or coding. The yellowlights are fully customizable to the needs of a project and can be set up for any call pattern of interest.

Exhibit 6. *FI Tracker* Single Interviewer View – UNCNNOD1

UNC CAPS FI Tracking Sheet
Click here to return to List of Reports
FI = UNCNNOD1 32 Cases
N = non-contact
R = refusal
A = appointment
C = complete
F = final incomplete
O = other
B = breakoff

CASEID	N	R	C	Thu 07-24		Fri 07-25		Sat 07-26		Sun 07-27		Mon 07-28		Tue 07-29		Wed 07-30	
	66%	0%	63%	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night
	21	0	20														
00004270	5	0	0				N			N,N,N	N						
00005208	2	0	0														
00005247	0	0	1			A											
00005267	5	0	0				N,N			N,N,N	N						
00005279	2	0	0														
00005280	1	0	1	N,N		A								A		C	
00005303	0	0	1	O		A					C						
00005320	2	0	1	N,N		N				A							C
00005321	0	0	1	A	C												
00005399	1	0	1	N,N,N													

Finally, we can take a look at the notes associated with one of the yellowlight scenarios. This is a scenario that would not immediately be obvious simply looking at raw ROC data. Case 00005399 has a yellowlight showing three noncontacts during the day of Thursday July 24. Scrolling to the right (not shown here), we see no further action on this case until the night of Wednesday August 6 when an appointment was set and then a complete during the day on Friday August 8. Clicking on CASEID 00005399 brings up the ROC call notes for this case, as shown in Exhibit 7.

Exhibit 7. ROC Call Notes for Case 00005399

UNC Renter Record of Call Notes
Click here to return to Main Report
caseid=00005399

Event Date	Code	Note
22JUL2008:21:18:13	999	Created by CMS after case import
24JUL2008:18:35:01	320	no answer, will try later.
24JUL2008:19:56:46	320	no answer, will try later.
24JUL2008:20:30:42	320	no ans, will try tomorrow evening,
06AUG2008:20:17:03	332	he works midnight shift..come by about 9am friday.
08AUG2008:12:10:19	491	Created by CMS after event/status code was modified by interview software
11AUG2008:17:37:53	980	

In Exhibit 7, we see the three non-contacts, appointment, and completed interview for case 00005399. At this point, the supervisor may want to contact the interviewer to talk more about these cases, especially the yellowlight cases followed by periods of inactivity to determine if these were the most efficient call strategies. This communication would benefit the supervisor by letting him or her know precisely what is happening in the field for particular cases that stood out and would benefit the interviewer if some retraining was necessary.

Future Enhancements
We would like to further develop FI Tracker to incorporate more project-specific rules and yellowlight conditions to best meet the needs of projects faced with analyzing a vast quantity of survey paradata. We will look to customize the tool to specific types of cases and interviewers and consider applying it in a phone survey environment and in other modes.

References

Murphy J., A. Safir, S. Myers, and C. Hottinger (October 2006). "Using SAS to Create a Flexible Survey Interviewer Tracking Chart." Proceedings of the Annual Midwest SAS Users Group Conference, Detroit, MI.

Riley, S.F., H. Ru, and R.G. Quercia (2009). "The Community Advantage Program Database: Overview and Comparison with the Current Population Survey" *Cityscape*. Volume 11, Number 3.
http://www.huduser.org/portal/periodicals/cityscpe/vol11num3/ch13.html