ARTICLES

Measuring Compliance in Mobile Longitudinal Repeated-Measures Design Study

Michael Link¹

¹ Nielsen

Keywords: visual data, photos, repeated measures, mobile survey, smartphone

<https://doi.org/10.29115/SP-2013-0023>

Survey Practice

Vol. 6, Issue 4, 2013

Using a longitudinal, repeated-measures approach, we test whether a smartphone designed to deliver short surveys at periodic intervals throughout the day can be used to capture key information on respondent behaviors and attitudes over time. Leveraging a study of 384 South Africans conducted during the 2010 FIFA World Cup games, respondents were administered a brief survey five times per day for 33 days and asked to take a picture of what they were focused on at the time of the survey. Focusing on critical questions of task compliance, we examine completion rates and three additional indicators of survey compliance. The research demonstrates how Smartphones can expand traditional survey methods, capturing both cross-time and "in-the-moment" data.

Introduction

The digital age allows researchers to conduct new forms of "participant observation" studies, providing detailed insights into the thoughts, behaviors, and attitudes of those they study (Lai et al. 2010). Such an approach is tested here using a mobile-based, repeated-measures design, which combines several digital measurements, including repeated mobile phone surveys and visual data capture to provide real-time information about respondents.

Given that mobile-based survey designs are still in the early stages of development, the analysis focuses on various aspects of respondent compliance, including participation rates over time and across various segments of the day as well as potential longitudinal changes in data quality measures meant to capture the level of respondent engagement in the study.

Methods

In June and July 2010, a panel of nearly 400 South Africans was recruited across the four largest cities in South Africa: Johannesburg, Pretoria, Durban, and Cape Town. Recruitment was a two-phase effort, with telephone outreach to a quota sample of potential respondents that fit defined age and employment targets, followed by face-to-face orientation sessions to familiarize respondents with the smartphone device and study guidelines.

Each respondent was equipped with a BlackBerry Curve running a Techneos SODA® survey application (Confermit, Oslo, Norway). With the exception of the phoning, camera and survey capability, all other features were hidden and locked from respondent use to minimize potential changes in respondent behavior. Phones were programmed in one of four languages: Zulu, Sotho, Afrikaans, and English. An initial set of cell phone minutes was provided as an

upfront incentive for participation. Panelists were also told they could keep the Blackberry as a gift if they completed all 5 weeks of the study (\$85 value in U.S. dollars).

The self-administered mobile survey was delivered five times a day for the 33-day fielding period. Five times each day (6:00 a.m.–10:00 a.m.; 10:00 a.m.–1:00 p.m.; 1:00 p.m.–4:00 p.m.; 4:00 p.m.–7:00 p.m.; 7:00 p.m.–11:00 p.m.) a reminder alarm would sound prompting respondents to complete the survey. Although the time intervals were known, the actual timing of the alarm and reminders was conducted at random intervals within each time period to deter panelists from anticipating the survey and potentially changing their behaviors accordingly.

The survey collected “in the moment” information pertaining to respondent’s activities, where he or she was, what he or she was doing, and with whom, as well as questions about his or her current mood. The goal was to administer all questions within a 3–5 minute survey. The last question of each survey asked the respondent to take a picture of what he or she was currently focused on and to provide a caption to describe the photo.

The analysis focuses on issues related to respondent compliance, in particular assessments of increasing survey nonresponse, declining item completeness, changing mood, and decreasing compliance with specific tasks (providing a photo and caption).

Results

Survey completion rate

Data were collected from 384 respondents between June 10 and July 12, 2010. Over the 33-day period, respondents were asked to complete 165 surveys (five per day across different parts of the day), resulting in 63,360 potential surveys across all respondents. The average, respondent-level completion rate (number of fully completed interviews / number of surveys offered) was 78.7 percent.

The average completion rate varied significantly over time (Table 1). Over the 5-week period, we see a pattern of initial learning (week 1), followed by more intensive usage (weeks 2 and 3), followed by a slight drop-off in participation during the latter weeks.

There were no significant differences in participation levels across any of the weeks based on sex, work status, or language. The only significant difference noted for age was in week 1, where participation was significantly higher among those 16–24 years old (87 percent) versus those aged 25 and older (where participation was 79 percent or lower). This likely reflects greater familiarity with new technologies, such as smartphones, among younger versus older people. This notion is buttressed by the cross-time pattern of participation which demonstrates what appears to be a “learning curve” during the first week for the older group, but less so with the younger groups.

Table 1 Survey completion rate by demographic characteristics and week.

Demographic Characteristics	n	Completion Rates						Cross Week (Sig.) ¹
		Overall	Week 1	Week 2	Week 3	Week 4	Week 5	
Total	384	78.7	77.7	82.7	82.6	75.2	74.2	(<0.001)
Sex								
Male	190	77.4	76.3	82.6	81.4	72.8	72.3	(<0.001)
Female	194	80.1	79.1	82.8	83.8	77.4	76.1	(<0.01)
Between Group (Sig.)		(0.197)	(0.347)	(0.945)	(0.298)	(0.122)	(0.225)	
Age								
16–24	78	80.5	87.0	86.1	82.6	74.2	69.2	(<0.001)
25–34	93	77.5	79.0	82.2	82.2	70.4	72.2	(<0.001)
35–49	137	79.7	74.4	82.4	84.5	78.9	77.4	(<0.01)
50+	76	76.8	72.4	80.3	79.7	75.2	76.2	(0.151)
Between Group (Sig.)		(0.614)	(0.006)	(0.498)	(0.564)	(0.182)	(0.243)	
Work Status								
Full-time	188	77.1	75.6	82.4	80.9	73.1	72.1	(<0.001)
Part-time	52	80.1	78.7	82.9	81.3	75.4	82.7	(<0.05)
Not Working	144	80.4	80.1	83.0	85.3	77.8	74.0	(<0.01)
Between Group (Sig.)		(0.314)	(0.367)	(0.977)	(0.219)	(0.346)	(0.089)	
Language								
Zulu	72	79.7	80.5	83.7	82.5	77.7	71.8	(<0.01)
Sotho	117	78.2	79.2	80.8	81.7	74.2	73.8	(<0.01)
Afrikaans	121	80.0	74.6	84.2	83.7	78.9	78.0	(<0.01)
English	74	76.6	77.8	82.1	82.4	68.1	71.1	(<0.01)
Between Group (Sig.)		(0.689)	(0.508)	(0.709)	(0.933)	(0.070)	(0.400)	

Note: completion rate=number of surveys with complete data / number of surveys in time period.

¹Significance based on comparison of highest and lowest values across the 5 weeks. This indicates level of significance in variation over time, not necessarily significant trending.

Finally, looking at completion rates across the different times of the day, we find remarkable stability (Table 2). There were no significant differences across demographic groups or within subgroups over time. Nonresponse therefore appears primarily driven by whether the respondent chose to continue participating in the study, and not necessarily participating only at selected times of the day.

Recording activities at the time of the survey

To gain a better understanding of whether or not respondents were “learning” to navigate the shortest path through the survey – a sign of respondent fatigue and potential measurement error – we examined responses provided to the question: “What were you doing before starting the survey?” Respondents could choose all that apply from a list of 14 potential activities. Most of these

Table 2 Survey completion rate by demographic characteristics and time of day.

Demographic Characteristics	n	Completion Rates					Cross Time Sig. ¹
		6:00 a.m.–10:00 a.m.	10:00 a.m.–1:00 p.m.	1:00 p.m.–4:00 p.m.	4:00 p.m.–7:00 p.m.	7:00 p.m.–11:00 p.m.	
Total	384	77.6	78.3	78.8	79.5	79.5	(0.255)
Sex							
Male	190	76.0	77.5	77.6	77.5	78.3	(0.442)
Female	194	79.3	79.1	79.9	81.5	80.7	(0.290)
Between Group (Sig.)		(0.131)	(0.453)	(0.290)	(0.065)	(0.322)	
Age							
16–24	78	79.3	80.3	78.9	80.1	83.8	(0.218)
25–34	93	76.1	76.7	78.2	78.2	78.3	(0.461)
35–49	137	78.3	79.0	80.1	81.1	79.8	(0.563)
50+	76	76.6	76.8	76.9	77.7	76.0	(0.654)
Between Group (Sig.)		(0.750)	(0.628)	(0.741)	(0.630)	(0.229)	
Work Status							
Full-time	188	76.1	76.5	77.3	78.1	77.5	(0.454)
Part-time	52	77.6	80.0	80.3	79.0	83.4	(0.179)
Not Work	144	79.6	80.0	80.1	81.5	80.7	(0.161)
Between Group (Sig.)		(0.354)	(0.273)	(0.398)	(0.324)	(0.218)	
Language							
Zulu	72	80.7	79.9	79.6	79.2	79.1	(0.355)
Sotho	117	76.9	77.9	78.1	78.3	79.8	(0.329)
Afrikaans	121	78.4	79.1	79.6	81.3	81.5	(0.282)
English	74	74.6	76.0	77.6	78.8	76.0	(0.335)
Between Group (Sig.)		(0.368)	(0.665)	(0.875)	(0.717)	(0.496)	
[n]							

Note: completion rate=number of surveys with complete data / number of surveys in time period.

¹Significance based on comparison of highest and lowest values across the five weeks. This indicates level of significance in variation over time, not necessarily significant trending.

questions had specific follow-up questions to obtain more detail about the activity. A respondent wanting to minimize the number of questions he or she was asked could therefore limit the number of activities chosen and hence reduced the survey length.

Looking at the average number of activities reported by respondents over time, the weekly average across respondents was 1.63 activities over the 5-week period. This number was very stable across all weeks of the field period, with no significant variation noted over time overall (Table 3). Moreover, there was little significant variation across time within the various demographic groups

Table 3 Average number of activities reported at start of interview by demographic characteristics by week.

Demographic Characteristics	Average Number Reported Activities					Cross Time ¹ (Sig.)
	Week 1	Week 2	Week 3	Week 4	Week 5	
Total	1.60	1.63	1.65	1.64	1.61	(0.374)
Sex						
Male	1.58	1.60	1.64	1.63	1.64	(0.145)
Female	1.63	1.66	1.65	1.66	1.59	(0.219)
(Sig.)	(0.527)	(0.480)	(0.981)	(0.775)	(0.637)	
Age						
16–24	1.54	1.54	1.53	1.57	1.55	(0.247)
25–34	1.68	1.67	1.65	1.68	1.62	(0.179)
35–49	1.60	1.63	1.66	1.63	1.60	(0.215)
50+	1.58	1.68	1.73	1.71	1.69	(0.237)
(Sig.)	(0.657)	(0.643)	(0.499)	(0.773)	(0.826)	
Work Status						
Full-time	1.55	1.54	1.54	1.57	1.55	(0.226)
Part-time	1.68	1.73	1.75	1.74	1.76	(0.247)
Not Work	1.64	1.71	1.74	1.70	1.65	(0.230)
(Sig.)	(0.393)	(0.076)	(0.057)	(0.261)	(0.251)	
Language						
Zulu	1.65	1.62	1.59	1.63	1.63	(0.262)
Sotho	1.51	1.57	1.61	1.58	1.52	(0.283)
Afrikaans	1.69	1.70	1.75	1.78	1.74	(0.197)
English	1.58	1.62	1.58	1.58	1.52	(0.273)
(Sig.)	(0.249)	(0.574)	(0.377)	(0.438)	(0.196)	
[n]	[361]	[372]	[375]	[369]	[353]	

¹Significance based on comparison of highest and lowest values across the 5 weeks. This indicates level of significance in variation over time, not necessarily significant trending.

examined. This would seem to indicate that respondents did not alter their reporting behavior when it came to concurrent activities from the start to the conclusion of the field period.

Respondent mood

Another potential indicator of respondent engagement (or disengagement) in the survey process is his or her self-reported mood at the time of the survey. Halfway through each survey, respondents were asked to indicate on a 5-point scale the number that most accurately described their current mood. The scale had two descriptive anchors: 1=bored, 5=excited. If respondents were growing weary of completing surveys over the 5-week period, one would expect a higher percentage of “bored” responses or a shifting in the average rating to the lower end over time.

Table 4 Average mood rating by demographic characteristics by week.

Demographic Characteristics	Average Mood Rating					Cross Time ¹ (Sig.)
	Week 1	Week 2	Week 3	Week 4	Week 5	
Total	3.65	3.38	3.38	3.39	3.43	(<0.001)
Sex						
Male	3.69	3.43	3.40	3.43	3.47	(<0.01)
Female	3.61	3.33	3.35	3.36	3.38	(<0.01)
(Sig.)	(0.394)	(0.359)	(0.678)	(0.517)	(0.452)	
Age						
16–24	3.22	2.86	2.88	3.02	3.10	(<0.01)
25–34	3.61	3.25	3.30	3.20	3.25	(<0.01)
35–49	3.83	3.59	3.56	3.56	3.57	(<0.05)
50+	3.83	3.69	3.65	3.70	3.73	0.108
(Sig.)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	
Work Status						
Full-time	3.58	3.31	3.26	3.19	3.26	(<0.001)
Part-time	3.80	3.40	3.46	3.63	3.61	(<0.05)
Not Work	3.68	3.46	3.49	3.57	3.58	(<0.05)
(Sig.)	(0.263)	(0.416)	(0.128)	(<0.001)	(0.021)	
Language						
Zulu	3.45	3.34	3.37	3.39	3.41	(0.224)
Sotho	3.46	3.20	3.08	3.05	3.14	(<0.001)
Afrikaans	3.92	3.62	3.67	3.68	3.71	(<0.01)
English	3.72	3.28	3.36	3.44	3.43	(<0.01)
(Sig.)	(<0.001)	(0.011)	(<0.001)	(<0.001)	(0.002)	
[n]	[361]	[372]	[375]	[369]	[353]	

¹Significance based on comparison of highest and lowest values across the 5 weeks. This indicates level of significance in variation over time, not necessarily significant trending.

Over the 5-week period the weekly average rating for respondent mood was 3.45 (Table 4). There was a slight change over time across all respondents, with week 1 showing a marginally higher average rating when compared to other weeks. The trending overall was fairly flat, indicating that respondents did not appear to change in mood appreciably after the second week of the study.

Photos and captions

Finally, we examined the percentage of completed interviews accompanied by a photo and a brief caption recorded by the respondent. To measure compliance, we looked at the percentage of completed interviews that had both a photo and an accompanying caption.

On average 92.0 percent of all completed interviews were accompanied by a photograph, while 66.4 percent of all completes had both a photo and a caption (Table 5). Women were more likely than men across each of the five weeks to both take a picture and insert a caption. Like some of the other

Table 5 Average percentage of pictures with captions by demographic characteristics by week.

Demographic Characteristics	Percentage of Pictures with Captions					Cross Time ¹ (Sig.)
	Week 1	Week 2	Week 3	Week 4	Week 5	
Total	67.8	66.5	68.0	65.3	64.5	(0.115)
Sex						
Male	63.6	62.0	63.8	60.1	59.2	(0.150)
Female	71.9	70.9	72.5	70.3	69.5	(0.215)
(Sig.)	(<0.05)	(<0.05)	(<0.05)	(<0.05)	(<0.05)	
Age						
16–24	71.4	65.8	71.1	66.1	62.6	(0.077)
25–34	70.0	68.6	71.3	66.9	66.7	(0.216)
35–49	67.8	68.4	68.2	66.4	65.3	(0.281)
50+	60.9	61.2	59.9	60.1	62.5	(0.356)
(Sig.)	(0.351)	(0.572)	(0.235)	(0.684)	(0.885)	
Work Status						
Full-time	67.0	64.0	67.1	60.9	61.1	(0.073)
Part-time	65.4	68.2	66.8	65.9	65.0	(0.342)
Not Work	69.7	69.1	70.7	70.9	68.8	(0.341)
(Sig.)	(0.735)	(0.464)	(0.818)	(0.082)	(0.252)	
Language						
Zulu	68.8	68.9	69.5	64.4	62.1	(0.147)
Sotho	62.8	58.6	62.2	58.6	57.7	(0.201)
Afrikaans	71.3	67.9	72.2	68.2	67.9	(0.196)
English	69.5	72.9	69.2	71.7	72.1	(0.354)
(Sig.)	(0.376)	(0.109)	(0.255)	(0.128)	(0.084)	
[n]	[361]	[372]	[375]	[369]	[353]	

¹Significance based on comparison of highest and lowest values across the 5 weeks. This indicates level of significance in variation over time, not necessarily significant trending.

measures already discussed, the percentage of photos with captions remained remarkably stable over time, both overall and within the subgroups examined. Again, this is an indication that respondents were largely compliant with this specific task.

Implications

The 2010 World Cup study provides some critical insights into how researchers can utilize current technologies to expand the types of studies that can be conducted. Given the length of the panel and that surveys were required at multiple points during the day, the study achieved a marked rate of completion. Most of the other respondent engagement measures showed stability over time – both overall and within the demographic groups examined. Where there were differences in compliance, they tended to be across demographic groups (as opposed to cross-time changes within specific subgroups) and they tended to be minimal. Of interest for future research is the apparent “learning curve” among older respondents as compared to

younger ones. Participation levels among older respondents rose over the course of the first week, whereas for younger respondents started high from the outset. This “learning curve” is a phenomena that researchers need to understand better and potentially account for in their study designs when using mobile forms of data collection.

The study has several potential limitations which should be noted. First, it was conducted in one country (South Africa) and during a time of heightened national awareness and unity with the country hosting the World Cup matches. Second, the sample was drawn using a quota (non-probability) methodology. While this approach ensured the proper proportion of respondents in terms of sex, age, work status and language, the fact that these respondents were not chosen via a random selection process could affect the generalizability of the findings.

The promise of the digital age for researchers lies in how we can best harness these new technologies to unlock insights into various avenues of inquiry. An approach utilizing periodic surveys via smartphones combined with other survey and non-survey data offers one potential option. The approach allows researchers to capture rich, detailed data in a manner heretofore unattainable through traditional focus group or survey approaches. As those technologies continue to evolve, researchers will need to keep pace in will keep pace, leveraging new hardware and applications to meet changing data needs.

REFERENCES

- Lai, J., L. Vanno, M. Link, J. Pearson, H. Makowska, K. Benezra, and M. Green. 2010. "Life360: Usability of Mobile Devices for Time Use Surveys." *Survey Practice* 3 (1). <http://www.surveypractice.org>.