Using Paradata to Evaluate Youth Participation in a Digital Diary Study

Marta Tienda and Dawn Koffman
Office of Population Research
Princeton University
September 2019

Abstract

The expansion of digital technology and Internet access has amplified opportunities and lowered the cost of longitudinal data collection. Mobile technologies are appealing for administering surveys to youth because they align well with their media and communication habits. This paper uses rich paradata derived from a year-long intensive longitudinal study (*mDiary*) that used a mobile-optimized web app to administer 25 bi-weekly diaries to adolescents recruited from the *Fragile Families and Child Wellbeing* study. Specifically, we investigate which aspects of teen recruitment experiences are associated with enrollment and longitudinal response patterns that consider both interim missingness and attrition; whether compliance behavior of teens who require multiple nudges to enroll differs from that of their peers who enroll on the first invitation; and what social circumstances facilitate the highest levels of longitudinal compliance. The conclusion highlights implications for future intensive longitudinal study designs that use digital platforms to deliver surveys to adolescents.

Using Paradata to Evaluate Youth Participation in a Digital Diary Study

Marta Tienda and Dawn Koffman

Introduction

Increasing difficulty recruiting respondents for representative surveys coupled with declining survey response rates have led to numerous studies that try to understand the reasons for these trends and the implications for data quality and statistical inference (Czajka and Beyler 2016; Groves et al. 2004). Although the expansion of digital technology enabled both by widespread access to the Internet and the proliferation of mobile devices has amplified opportunities and lowered the cost of data collection (Link et al. 2014), widespread use of call filtering and call answering systems permit prospective respondents to screen telephonic solicitations and evade recruitment (Czajka and Beyler 2016). Even as cross-sectional survey response rates continue a downward spiral, wave-to-wave longitudinal response rates appear to have remained steady, partly buoyed by use of paid incentives, identification with the study, and topic salience (Schoeni, et al. 2013). For adults there is evidence that pre-paid financial incentives boost survey response rates across modes of administration (Singer et al. 1999; 2000), and there is some evidence of carry-over persistence across repeated surveys, particularly when topic salience is high (Schoeni et al. 2013; Laurie and Lynn 2009; Singer and Ye 2013), but comparable evidence for teens is limited.

Research involving youth faces even higher participation hurdles because adult (usually parent) consent is required before requesting teen assent is allowed. Evidence about topic salience and incentives in maintaining youth participation in longitudinal surveys is both limited and inconsistent, particularly for designs that involve frequent measurements over several months (Halpern et al. 1994; Boys et al. 2003; Post et al. 2012; Powers and Loxton 2010;

Goldberg et al. 2019). Although youth are intense users of digital devices for communication, news, and entertainment (Lenhart 2015; Anderson 2016; Anderson and Jiang 2018; Rideout 2015), they are seldom the subjects of intensive longitudinal studies with year-long time horizons (Jaccard et al. 2004; Bergdall et al. 2012). A notable exception is the Relationship Dynamics and Social Life (RDSL) study, which administered a multi-year weekly diary study using electronic and telephonic methods to study the processes leading to unintended pregnancy (Barber et al. 2016; Wagner et al. 2019). This study was restricted to women ages 18 and 19, hence did not require adult consent. To our knowledge, no study has administered a year-long diary study (intensive longitudinal survey) to minors, nor have analyses of compliance behavior considered high frequency data.

This paper uses rich paradata from a year-long digital diary study about romantic relationships to evaluate adolescents' longitudinal response behavior. Paradata, which is process data compiled through subject recruitment and respondents' interaction with the survey instruments and portals, is proving valuable for understanding nonresponse and attrition in longitudinal surveys (Callegaro 2013; Lugtig 2014; Kocar 2019; Bristle et al. 2014). In addition to recruitment process indicators that are customarily generated through computer-assisted telephone interviews (CATI), such as modes and number of contacts, time elapsed until consent/assent, and survey response times, digitally-administered surveys produce further paradata that is relevant for understanding compliance in longitudinal studies. Examples include types of devices used to access web portals, email providers, log-in attempts, familiarity with authentication procedures, and opportunities for pre-notification of new surveys, among others (Callegaro 2013; Kocar 2019).

We evaluate youth compliance in an intensive longitudinal survey by addressing several specific questions that distinguish between actions that are under the control of the researcher and the respondent. First, conditional on assenting to participate, what aspects of the recruitment experience are associated with enrollment in the study? Second, do recruitment experiences carry over to response behavior patterns, such as longitudinal compliance, interim missing diaries and attrition? Third, do teens who enroll only after receiving multiple invitations exhibit different diary compliance behavior compared with teens who register after a single invitation? Finally, what respondent circumstances are associated with compliance in a year-long diary study?

Following a summary of recent literature about the value of paradata for understanding longitudinal compliance behavior, we describe the design and instrumentation of the digital diary study, emphasizing aspects of the recruitment and enrollment protocols hypothesized to influence longitudinal compliance. The conclusion summarizes key finding and highlights general lessons about the promise of paradata for understanding adolescents' compliance in mobile-enabled surveys.

Paradata and Survey Response Behavior

Digital technologies are appealing for administering intensive longitudinal studies to adolescents both because they facilitate respondent convenience in taking surveys and are well aligned with youth media habits (Anderson 2015) and also permit timely data retrieval (Link et al. 2014; Raento et al. 2009). Importantly, web-administered surveys facilitate the capture of information about various aspects of the data collection process, including recruitment experiences and modes of Internet access, that have furthered understanding of longitudinal response behavior in adults (Callegaro 2013; Kocar 2019; Lugtig and Blom 2018). Because most intensive longitudinal studies involving youth use signal-triggered methods that capture multiple

daily measurements over short durations (Hensel et al. 2012; Runyan et al. 2013), there is limited information about adolescents' compliance behavior in web-administered surveys that span several weeks or months (Goldberg et al. 2019; Barber et al. 2016; Wagner et al. 2019).

Survey researchers generally agree that attrition risk accumulates over waves and that high frequency data produces complex nonresponse patterns (Lugtig 2014; Wagner et al. 2019). For example, analyses of attrition in multi-wave studies that collapse wave-on-wave attrition as a binary outcome assume uniform response propensities across waves and ignore intermittent nonresponse that does not result in permanent withdrawal (Lugtig 2014). The mechanisms creating intermittent non-response and attrition in longitudinal surveys change over the span of the study. These include extraordinary personal events (e.g., death of a family member, job losses, medical emergencies, and relocation disruptions) along with response fatigue, and topic salience (Lugtig 2014; Kocar 2019; Barber et al. 2016; Wagner et al. 2019).

Correlates of nonresponse in web-administered surveys differ somewhat from those for face-to-face and CATI surveys not only because of individual variations in internet proficiency by age and socioeconomic status, but also because automated notifications may go to spam folders; because of connectivity failures; or because respondents change service providers and/or electronic mail addresses (Lugtig and Blom 2018; Callegaro 2013). Furthermore, there is some evidence that web-administered surveys that require access to a desktop or laptop appear to dampen longitudinal compliance (Turner et al. 1998; Barber et al. 2011; Link et al. 2014).

For youth who came of age as digital technology became socially ubiquitous, dubbed "digital natives" by Prensky (2001), access to smartphones is a *sine-qua-non* for entertainment, communication, as well as forming and maintaining social and romantic relationships (Anderson 2015; Anderson and Jiang 2018; Goldberg and Tienda 2017). We hypothesize that access to text-

enabled mobile devices not only increases teens' willingness to participate in a diary study about romantic relationships, but also their longitudinal compliance. To address whether, how much, and in what ways adolescents participate in a web-enabled digital diary study about romantic relationships, we draw upon leverage theory about topic salience (Barber et al. 2016), new insights about the power of paradata to understand response behavior (Lugtig and Blom 2018; Kocar 2019; Callegaro 2013), and established findings about the social and economic correlates of survey participation (Groves et al. 2004; Lugtig 2014). We also examine whether incentives, which

are positively associated with adult longitudinal compliance (Singer and Ye 2013; Kocar 2019), also boost compliance among youth.

Data

The analyses draw on three sources of data: (1) response data collected by the *Fragile Families and Child Wellbeing Study* (FFCWS), a prospective birth-cohort study that followed almost 5,000 children from birth through approximately age 15 across six survey waves; (2) response data from the *mDiary Study of Adolescent Relationships* (mDiary) which used an intensive longitudinal design to query a sample of the FFCWS youth for a year-long period starting when the youth were 16 to 17 years old; and (3) paradata generated during the recruitment, enrollment and data collection processes of the mDiary study. The FFCWS birth cohort study followed a cohort of children born at the turn of the millennium in 20 medium-to-large U.S. cities (Reichman et al., 2001). By design, births to unmarried mothers were oversampled at baseline. Index children and their primary caregivers were surveyed over six waves, most recently when the youth were approximately 15 years of age. The FFCWS surveys provide rich information about target youths' socioeconomic background, living arrangements,

school behaviors, and age-specific measures of socio-emotional development, among many others. Our analyses use information from the baseline survey (FFCWS baseline) and Year 15 surveys conducted with primary caregivers (FFCWS Y15-parent) and target youth (FFCWS Y15-teen), as well as the *mDiary* enrollment survey and the 25 bi-weekly diaries.

mDiary was designed to investigate the nature and dynamism of teenagers' romantic relationships. Recruitment occurred over a 17-month period (November 2015 – April 2017) on a rolling cohort basis, lagging the field operations of the FFCWS parent study by approximately a year.[1] Access to a private email address was the only requirement to participate in the study. All surveys were administered via a mobile optimized custom web app (mdiary.org) linked to the Qualtrics web survey platform via API calls. Teens could take the surveys on desktops or mobile devices, such as tablets or smartphones, provided the latter were not shared. The enrollment survey included several non-repeating baseline questions that replicated items in the parent study to gauge change during the intervening year.[2] To incentivize compliance, respondents received Amazon e-gift cards, disbursed via email or text. Following the enrollment survey ($5), respondents earned $2 for each completed diary; the Amazon gift cards were delivered upon completing three ($6) or four ($8) consecutive diaries. Respondents who completed the last survey received a bonus gift card of $10.[3]

The process of recruiting and enrolling respondents for the *mDiary* study generated rich paradata that is relevant for understanding longitudinal response behavior (Durrant and Kreuter

[1] mDiary respondents were recruited from 13 of 20 cities in the parent study. These include Baltimore, Boston, Corpus Christi, Indianapolis, Jacksonville, Milwaukee, Nashville, Newark, New York, Norfolk, Philadelphia, Pittsburgh, Richmond, San Antonio, and San Jose); FFCWS Year-15 participants with contact information known to be invalid were excluded from the sampling frame. In nine of the thirteen target cities, *mDiary* sampled 100% of eligible adolescents; adolescents from Newark, Philadelphia, Baltimore, and Richmond were randomly sampled at a rate of 44%.

[2] Median completion time for the enrollment survey was 7.3 minutes compared with 2.5 for the diaries.

[3] Respondents were randomly assigned to the 3 and 4 consecutive survey incentive groups.

2013; Callegaro 2013). These include various aspects of the recruitment experience, enrollment latency, digital proficiency and mode of Internet access. *Recruitment experience* measures captured by the *mDiary* study include mode of assent and the number of telephonic contact attempts required to obtain assent. Because less than five percent of families granted consent and assent via the materials mailed in the welcome package (or via email), for the vast majority of recruited teens consent and assent were obtained telephonically.[4] The survey team recorded the date and time of each contact attempt as well as the as the date and mode (paper, email or phone) of assent. Using the date assent was obtained, we calculated the time lag (in days) between the date of assent and the date of first enrollment invitation. We define this invitation lag, which is a function of study operations rather than respondent behavior, as either short (0-7 days), medium (7-14 days) or long (15 or more days).[5] We also recorded whether the first invitation to register for the study occurred during summer, when many teems enjoy more free time than when school is in session.

Texting is adolescents' preferred mode of communication, and while young people seldom check their email (Anderson 2015; Coyne et al. 2017), many do so if required. Most teens use commercial email addresses like Gmail or Outlook, but some rely on accounts provided by their schools. Our paradata captures these aspects of data collection. Paradata used to gauge respondents' *digital access and proficiency* include teens' preferences (text vs. email) for notifications about an open diary survey; whether teens use a school email address (versus private or public providers); and whether teens' required assistance in the dual authentication

[4] For some families consent and assent were obtained concurrently because both the PCG and youth were present when called by the study team; however, the majority required numerous follow-up calls to obtain teen assent even after parents granted consent.

[5] The lag arose because rather than sending teens an electronic invitation to enroll as soon as assent was obtained, on a bi-weekly basis the research team formed a cohort of recently assented teens who would each be sent an enrollment invitation at the same time (alternating Sundays at 4 PM).

process.[6] Paradata also reveal the type of device used to enroll in the study, which we use to construct a binary indicator that distinguishes between smartphones and other devices.

mDiary Sample and Measures

Figure 1 provides a summary of the recruitment, enrollment and response behavior of the Fragile Family teens recruited for the *mDiary* study. Of 869 respondents with valid contact information, 689 (80%) assented to participate in the study. Assented teens were usually invited to enroll in the study within two weeks of their assent to participate.[7] The enrollment instructions were provided by email, but respondents who provided a cell phone number during the recruitment process also received a text message alerting them to check their email to view the instructions.

Figure 1 About Here

To enroll in the study, assented teens were required to select a user name and password and to complete an enrollment survey. Conditional on assent, over three-fourths of teens (531) enrolled in the study. The enrollment survey as well as the diaries opened on Sundays at 4 PM local time and remained open for one week. Diaries not completed by the end of the week-long response window were considered missing.

Response Behavior

We analyze two types of response behavior: *enrollment* conditional on assent, and *longitudinal compliance*. To evaluate who did and did not enroll in the mDiary study, we create a binary measure where 1 indicates that a teen registered on the user website and completed the

[6] mDiary used a dual authentication process, which required youth to recall an image they selected in order to access the website. Youth who did not recall their security image were locked out of the system after three guesses and required a manual reset of their security image at the time of enrollment.
[7] During the first recruitment quarter, many respondents experienced a lag in excess of two weeks because the recruitment team was not fully staffed during the 2015-2016 holiday season.

enrollment survey and 0 indicates nonregistration. We further examine *enrollment* by creating a binary measure that distinguishes between respondents who registered at the first invitation to enroll (on-time) and those who required multiple invitations before they enrolled (delayed). Although the vast majority of study participants enrolled within a week of receiving their first invitation (designated on-time enrollees), about 9% required multiple invitations before registering for the study. Teens who failed to register within the allotted time window were added to the next cohort and re-invited to enroll up to five times before they were designated nonenrollees.[8]

We use several indicators to portray longitudinal compliance, including the total number of diary surveys completed (range = 25); the last survey completed after the enrollment survey (enrollment survey is coded zero); and response behavior patterns that consider interim missingness and attrition (Barber et al. 2016; Wagner et al. 2019). In order to capture the intermittent participation and nonmonotone attrition (Lugtig 2014; Kocar 2019), we examined the distribution of completed surveys and patterns of interim missingness. Extensive data diagnostics produced a 5-category compliance scheme that considers the variation in response consistency and persistence and defines attritors as teens who did not complete the last survey.

- Full compliance: completed 25 diaries
- Highly engaged: completed 22-24 diaries
- Interim missingness, <22 completed diaries
- Attrition, <22 completed diaries[9]
- Attrition + interim missingness, <22 completed diaries

[8] A handful of assented teens was invited up to 7 times, but the number of invitations was capped at 5 as it became clear that recalcitrant teens were not likely to enroll after repeated invitations.
[9] We separately identify a subset of this group, namely respondents who completed the enrollment survey, but none of the diaries.

The lower panel of Figure 1 reveals a high level of compliance overall: 44 percent of respondents completed all 25 diaries, and an additional 12 percent completed between 22 and 24 (88 to 96 percent). The overall compliance rate (person surveys) was 71 percent.

Social and Behavioral Correlates of Survey Participation

Because *mDiary* respondents were sampled from a 15-year birth cohort study, information about personal and family circumstances associated with compliance was available from the FFCWS baseline survey and the Year-15 primary caregiver (mostly parents) and Year-15 youth surveys. Psychosocial characteristics and topic salience also influence response behavior (Lugtig 2014; Kocar 2019). The key measures drawn from the FFCWS baseline survey are informed by a vast literature about survey response behavior and attrition (Groves et al. 2002; Singer and Ye 2013). These include mother's self-reported racial identification, educational attainment, and marital status at the birth of the teen. Parents reported about the availability of home Internet service at the Year-15 interview, but teens reported on their living arrangements (with both biological parents, with one biological parent or with neither); the amount of time spent alone (often, sometimes, rarely or never); the level of chaos in their home; and their perseverance on several tasks ("grit").

We construct a measure of grit based on three self-reported items—keeps at schoolwork until done; sticks with plans to get things done; and finishes whatever begins—to assess respondents' propensity to remain in the *mDiary* study. These items were measured with declarative statements that used a 4-point scale to measure level of agreement: strongly agree, somewhat agree, somewhat disagree, and strongly disagree. Topic salience is captured as a binary indicator designating whether respondents had ever dated. By design, this information was obtained both in the FFCWS Year 15-teen survey and in the *mDiary* enrollment survey. The

former measure is used to predict enrollment in the diary study, and the *mDiary* measure, captured in the enrollment survey, is used to predict longitudinal compliance.

Appendix A summarizes sample characteristics for recruited teens according to enrollment status.[10] Several differences between enrollees and non-enrollees are noteworthy. First, enrollment rates were higher for girls than for boys. Second, white teens enrolled at higher rates than their minority counterparts. Third, teens with college-educated mothers enrolled at higher rates than teens with lesser educated mothers (49 percent vs. 33 percent) and teens whose mothers were married at the time of their birth also enrolled at higher rates (34 percent vs. 25 percent). These differentials are consistent with those observed for adult samples (Groves et al. 2002; Watson and Wooden 2009; Singer and Ye 2013).

Convenient access to the internet is necessary for compliance in a digital diary study. Nationally over 85 percent of U.S. households subscribe to Internet services, and approximately 60 percent both subscribed to broadband services and owned multiple devices to access the Web (Ryan and Lewis 2017). Over 90 percent of assented teens had Internet service at home, with enrolled teens displaying a modest advantage over their non-enrolled counterparts (94 percent vs. 88 percent, respectively). Nearly three-quarters of assented teens reported having dated at the Year-15 interview, with a 10-point difference between enrolled and non-enrolled teens favoring the latter. The dating rate of *mDiary* enrollees inched up about up 4-points in the year following their Year-15 interview.

[10] Because the sample was drawn from a birth cohort study, there was limited age variation among respondents and is not reported or modeled in any analyses. The median age at the first mDiary survey was 16.7 years, approximately one year had elapsed since the FF Year-15 interviews.

Analytical Strategy

The empirical analysis proceeds in two stages. After describing the sample of assented teens, we use logistic regression to assess whether, which and to what extent recruitment experiences and variations in digital proficiency are associated with enrollment in the study. Our analyses sequentially model teen and family background characteristics associated with response behavior in adults to evaluate the robustness of the results (Groves et al. 2002; Watson and Wooden 2009; Barber et al. 2016). Subsequently we evaluate whether recruitment experiences and enrollment timing (delayed vs. on-time) are associated with overall compliance, intermittent participation and attrition. We use linear regression for interval measures (total and last survey completed).

Results

1. Enrollment Behavior

The paradata reported in Table 1 offers several insights about recruitment experiences of assented teens who did not register for the *mDiary* study or registered only after receiving multiple invitations. First, nonenrollees were harder to reach than their enrolled counterparts, as evident by the higher median contact attempts and the higher maximum range to obtain assent. Second, among enrollees, the mode of assent differs between teens who registered at the first invitation to enroll (on-time) and the 9% who required multiple invitations (delayed enrollees). Enrollment delays were largely incurred for logistical reasons, such as dysfunctional email addresses, incorrect or shared cell phone numbers, and requested postponement for extracurricular activities. For some teens the dual-authentication procedure incurred delays because teens neglected to select a security icon to verify their identity upon registration for the

study. These instances required additional contact between respondents and the study team, often lasting several days owing to teenagers' busy schedules.

Table 1 About Here

Third, the paradata reveal one striking difference between assented teens who did and did not enroll, namely capacity to receive text messages. Virtually all enrolled teens provided a *both* an email address and a private cell phone number, compared with only 79 percent of non-enrolled teens. Furthermore, compared with teens who enrolled at the first invitation, a lower share of delayed enrollees was recruited during the summer, when teens presumably have greater time flexibility that could facilitate enrollment. Lastly, compared with on-time enrollees, delayed enrollees appear to be less digitally proficient and have more limited access to mobile technology. Approximately 20 percent required enrollment assistance and two-thirds enrolled using a smartphone compared with three-quarters of on-time enrollees. Given teens preference to communicate using smartphones and text messages (rather than email), both circumstances may also influence their longitudinal persistence.

Table 2 reports the results of a logistic regression that reveals which aspects of subject recruitment are associated with enrollment. Three aspects of recruitment and digital proficiency are associated with enrollment propensities. First, the amount of effort expended to assent teens is inversely associated with their enrollment odds, which calls into question the value of pursuing subjects whose recalcitrance may signal reluctance to participate. Second, long lags between assent and the invitation to register dampened teens propensity to enroll. Teens who received their enrollment invitation 15 or more days after assent were approximately half as likely to register for the mDiary study compared with their counterparts with invitation lags of a week or less. The strongest predictor of enrollment is ability to receive text messages, which presumes

uninterrupted access to cellular services. Teens whose communication with the study team included both email and text messages were over 6 times as likely to enroll as their counterparts lacking this capability. Consistent with the burgeoning literature about teens use of technology, access to text-enabled devices is vital for enlisting teenagers in a web survey because they seldom check their email (Anderson 2015; Anderson and Jiang 2018; Coyne et al. 2017). These associations are robust to specifications that model respondent and family characteristics.

Table 2 About Here

Girls who assented to participate were over twice as likely as their male counterparts to register for the *mDiary* study (see also Appendix A). Topic salience, measured by having ever dated by age 15, was unrelated to enrollment in the study, however. Enrollment odds differed in significant and substantively meaningful ways according to mother's education. The lower odds of enrollment corresponding to teens who live with both parents reflect differences in mother's education rather than parental supervision. Enrollment odds for teens whose mothers attended or completed college were two to three times the odds of adolescents whose mothers did not complete high school. Home access to the Web did not boost enrollment odds, most likely due to the widespread penetration of home Internet services (Ryan and Lewis 2017).

As shown in Table 1 (and Appendix A), delayed enrollees differ from on-time enrollees in several ways that are related to compliance. Compared with on-time enrollees, higher proportions are minority, male, and have mothers with low education—attributes that prior studies associate with lower participation propensities (Lugtig 2014; Groves et al. 2002). Although similar shares of on-time and delayed enrollees had access to text-enabled devices, the latter appear to be less digitally proficient: 21 percent required help registering for the study

compared with 6 percent of on-time enrollees and two-thirds enrolled via smartphone as compared with nearly three-quarters of their on-time peers.

To assess how enrollment timing might signal differential longitudinal response propensities, we model the enrollment decision using a tripartite outcome that distinguishes among on-time (reference group), delayed and nonenrollment groups. Results reported in Table 2a clarify that the temporal lag between assent and first enrollment invitation contributed to nonenrollment. A lag of 15 or more days boosted the odds of nonenrollment 1.7 times relative to enrollment, but no such association obtains for delayed enrollment. Rather, limited digital proficiency and having assented by paper, appear to drive enrollment delay. These two aspects of recruitment are related because the paper assent form required teens to select a security image that was used for dual authentication during the registration process.[11] Teens who contacted the research team to reset the security icon were over four times as likely to enroll with delay than their peers who did not need assistance. These results robust to inclusion of teens' personal and family circumstances.

Table 2a

Results reported in the right panel of Table 2a provide additional insights about respondent circumstances associated with enrollment in the study. Compared to their White peers, Black youth were, respectively, 2.3 and 1.8 times as likely to enroll with delay or not enroll. Mother's educational status is associated with whether or not youth registered for the study, but not delayed registration. More specifically, youth whose mothers attended or completed college had higher enrollment propensities, but mothers' education was inconsequential for the timing of enrollment. That the effort to assent is inversely associated with

[11] It is likely that some teens forgot which image they selected and were locked out after three failed guesses.

the likelihood of enrollment raises questions about the trade-off between boosting sample size and participation. Each additional contact to obtain assent raised the odds of nonenrollment by approximately 4 percent. Although the corresponding point estimate for delayed enrollment is imprecisely estimated, the positive sign suggests that delayed enrollment might be associated with lower longitudinal compliance.

Table 3 provides supporting evidence about lower longitudinal persistence for delayed enrollees, who averaged 7 fewer completed surveys, than on-time enrollees. Delayed enrollees also experienced higher and earlier attrition than their on-time peers. Over half of delayed enrollees failed to complete the final diary compared with approximately one-third of on-time enrollees. By contrast, over 60 percent of on-time enrollees completed at least 22 diary surveys compared with about one-in-four delayed enrollees. What is more, fully 10 percent of delayed enrollees did not complete any of the 25 diaries after enrolling, over twice the rate of on-time enrollees.

Table 3 About Here

In addition to averaging fewer completed surveys, teens who enrolled with delay attrited at a faster rate than teens who enrolled on-time, as Figure 2 illustrates. The steepness of the two curves between one and six completed surveys is striking. Only 55 percent of reluctant teens completed at least 6 surveys compared with 85 percent of on-time enrollees. Less than half of delayed enrollees completed at least 13 surveys compared with over three-quarters of on-time enrollees. Although the number of delayed enrollees is relatively small, it is striking nonetheless that only 15 percent completed all 25 diaries as compared with almost half (47 percent) of on-time enrollees. After describing variations in response behavior, we address whether their early

disengagement reflects a lack of interest in the study or other circumstances associated with participation, such as convenient Web access or unsupportive home environments.

Response Behavior

Table 4 sorts respondents according to the scheme that classifies respondents by the number of diaries completed, whether they skipped diaries in the course of the study, and whether they persisted through the final diary (#25). The average number of completed diaries varies between 23 to 25 for the highly and fully engaged to a low of 3 for the attrition category. Interim missingness adds considerable heterogeneity to the response behavior patterns, as evident in the third and fourth columns. Interim missingness often eventuates in attrition, but not always, depending on whether participation is incentivized and whether intermittent nonresponse signals loss of interest. Both groups that skipped diaries intermittently completed less than the sample mean of 17 (see Table 3). Possibly because the last diary was compensated with a $10 incentive, nearly 10 percent of teens who participated intermittently actually completed the last diary. The group that attrited in the absence of intermittent nonresponse completed three surveys, on average, and included 26 respondents who enrolled but did not complete a single diary.

Table 4 About Here

Paradata provides some insight into variations in longitudinal response behavior. First, over one-quarter of fully and highly engaged teens assented via materials provided in the welcome package, signaling their interest in the study, but a comparable share of the intermittent respondents did as well. By comparison, between 11 and 19 percent of the attrition group provided paper or electronic assent. Second, there is little variation among the compliance groups in the median number of days between assent and the first invitation to enroll; however, the range in elapsed days varies appreciably owing to factors beyond the control of the

recruitment team.[12] This is meaningful because prolonged delay in communication with the survey team may dampen teens' interest in participating. The descriptive tabulations indicate that recruitment during summer either boosts or lowers longitudinal compliance.

Finally, although access to text-enabled devices boosted the odds of enrollment, the association with longitudinal compliance is less clear. Over 95 percent of enrolled teens had access to text-enabled devices, but enrollment using a smartphone did not guarantee the highest response rates. Over three-fourths of the fully engaged group enrolled using a smartphone, as did a slightly higher share of the most disengaged group. Both response fatigue associated with participation in the birth cohort study and *mDiary* topic salience may partly explain this puzzle. Eligibility for mDiary was restricted to teens who completed the Year-15 FFCWS survey. Moreover, the mDiary enrollment survey, which collected some background information to supplement and link with the FFCWS Year-15 youth survey, was longer than the bi-weekly diaries and also did not ask about current romantic relationships.

Appendix B reveals noteworthy differences in the sociodemographic profiles of teens according to compliance groups. Three-fifths of fully compliant teens were girls; however, among respondents who only completed the enrollment survey, over 60 percent were boys. Paralleling findings for adults, teens with white and college-educated mothers completed more surveys than their minority peers whose mother did not complete high school. Finally, virtually all of the fully and highly compliant teens had home Internet service, as compared with 87 percent of respondents who attrited and completed fewer than 22 surveys. In the remainder of the

[12] Most notable was the need for follow-up calls to obtain missing information required for enrollment (e.g., valid email addresses), but other reasons include verification of contact information (cell phones, name spelling, etc.) and requested delays for extracurricular activities.

paper we address how these individual and familial circumstances contribute to understandings of adolescents' longitudinal compliance in the diary study.

Longitudinal Response Behavior

To evaluate whether recruitment experiences carry over to longitudinal compliance, we regress the total number of surveys (Table 5) and the last survey completed (Table 6) on paradata measures that capture variations in recruitment experiences and digital proficiency/access, along with respondent and family background characteristics associated with survey participation (Barber et al. 2016; Groves et al. 2002; Lugtig and Blom 2018). The two outcome measures are not normally distributed, hence violating the assumptions of ordinary least squares (OLS) estimation procedures. Therefore, we evaluate the robustness of our findings by modeling the odds of completing the sample means for the total number and number of the last survey completed (Barber et al. 2016; Wagner et al. 2019). Because the results are substantively similar, we discuss the OLS estimates, which have a straightforward interpretation, and report the logit results in supplementary appendix tables.

Tables 5 and 6 About Here

Conditional on enrollment in the *mDiary* study, the difficulty of recruitment, the time lag between assent and first enrollment invitation, and access to text-enabled devices are unrelated to longitudinal compliance. However, the influence of assent mode carries over to longitudinal compliance. Signaling their interest in the study and disposition to participate, youth who assented to participate via the welcome package materials completed two more diary surveys, on average, compared with their peers who provided assent via telephone. Estimates for last survey completed differ in magnitude by fractional amounts.

Enrollment with delay undermines longitudinal compliance appreciably, and can have consequences for data quality. Youth who registered for the *mDiary* study only after repeated invitations completed seven fewer surveys than youth who enrolled on-time. That the last survey completed by delayed enrollees was five to six fewer than their on-time peers indicates a faster attrition from the study. These rather substantial differences suggest that enrollment delay results from factors other than digital proficiency and also reflects hesitancy about joining the study. Participation, of course, is voluntary and respondents were advised that they could withdraw at any time.[13] Associations for both paradata indicators persist even after modeling teen and family characteristics associated with survey compliance, and are also replicated by the logistic regressions that estimate the odds of completing at least the mean number for both outcome variables (see Appendix Tables 5a and 6a).[14]

Surprisingly, we find that topic salience, captured by having dated between the year-15 Fragile Families interview and their enrollment in the mDiary study, was unrelated to longitudinal compliance. Girls were more likely than boys to enroll (Table 2a), and completed 1.4 more surveys than their similarly situated male counterparts (full model). There were no gender differences in the last survey completed, however. Rather, two sets of circumstances were associated with higher longitudinal compliance: availability of Internet service at home, and family socioeconomic status, as measured by mother's educational attainment at the birth of the teen. Teens with home Internet service completed four more diaries, on average, than their peers lacking convenient home access to the Web. Results for the last survey completed were lower

[13] Only 19 teens (0.8 percent) requested to withdraw from the study; of these, 3 (15.8 percent of withdrawals) were delayed enrollees.

[14] The sole exception obtains for access to text-enabled devices, which is associated with 2.5 higher odds of completing 17 or more diaries; however, this association is rendered statistically insignificant after youth and family characteristics are modeled.

by half a survey, on average. Finally, respondents whose mothers completed high school or college averaged two to three more surveys, on average, compared with their peers whose mothers did not complete high school.

Discussion

It is rather remarkable that, conditional on enrolling in the *mDiary* study, adolescents answered 67 percent of the 13,275 possible diaries, and that fully two-thirds completed the last survey. That roughly one-third of teens who completed the final diary did so with interim non-response raises questions about how they differ from respondents who missed diaries and eventually attrited and those who attrited without skipping diaries. The unique design of *mDiary*, which samples from a birth cohort of "digital natives" (Prensky 2001; Hargittai 2010), provides several important advantages for understanding response behavior which include extensive background data since birth collected over a 15-year period; a known sampling frame that permits statistical inference; and because the diaries were administered digitally, also rich paradata about the process of data collection. Teens media habits are well aligned for web-administration of high frequency surveys, provided youth actually enrolled.

The paradata provide important insights for improving longitudinal participation in ways that build on teens' digital communication preferences. Two aspects of recruitment experiences are key to understanding enrollment, namely the difficulty of obtaining assent and communication lags between obtaining assent and inviting teens to enroll. The inverse association between enrollment likelihood and the number of calls to obtain assent suggests low interest in the study. Teens who proved difficult to assent also enrolled with delay at higher rates and completed significantly fewer diaries than their peers who enrolled after only one invitation.

Decisions about the amount of team effort expended in obtaining teen assent are under the control of the research team, as is the management of time lags between assent and the invitation to register for the study. Delays in excess of 15 days dampened interest in the study and resulted in lower enrollment rates as well higher odds of delayed enrollment. For digital natives accustomed to quick if not immediate follow-up, communication delays are costly for participation.

Some aspects of recruitment carry over to longitudinal compliance, notably mode of assent and enrollment with delay. In addition, and unsurprising given teens communication preferences, smartphone access is virtually a *sine qua non* for teen participation in web-administered surveys (Tienda et al. 2018). Conditional on assent, teens with access to text-enabled devices were over 6 times as likely to enroll as their peers lacking texting ability. Although most teens are considered to be digitally proficient to varying degrees, skills needed for texting differ from those needed to establish and manage email accounts and retrieve gift cards electronically (Hargittai 2010). Teens who required assistance during the registration process also were more likely to enroll with delay compared with their more proficient peers. Recruitment experiences that result in delay ultimately undermine longitudinal compliance, resulting in 4 fewer completed surveys, on average, and higher attrition. Although access to text-enabled devices was not associated with overall longitudinal compliance, access to in-home Internet service boosted the number of completed diaries. Nevertheless, because email is not popular among youth, requiring email addresses to receive enrollment instructions proved limiting. For teens, delivering instructions via text could potentially increase enrollment; this survey delivery experiment is worth further investigation.

The burgeoning field of human-computer interaction indicates that mobile technology is best suited for short surveys and that "apps" are preferable to Internet browsers for administering repeated surveys (Buskirk and Andres 2012; Link et al. 2014). Non-compliance is costly, particularly for studies about youth because of the extra protections required to conduct studies targeting minors. Although there is growing agreement that rising nonresponse rates do not necessarily compromise data quality or introduce systematic bias, most of the evidence is based on adults and longitudinal surveys administered annually or bi-annually (Czajka and Beyler 2016; Groves and Peytcheva 2008). These inferences do not necessarily apply to youth, nor to high frequency studies that span several months. Nearly 5 percent of teens who registered for *mDiary* did not complete a single diary beyond the enrollment survey. It is possible that the length of the enrollment survey, which required a median completion of 7.3 minutes, dampened interest in participating. That roughly 10 percent of delayed enrollees only completed the enrollment survey compared with 4 percent of on-time enrollees further questions the wisdom of recruiter persistence because the difficulty of obtaining assent is inversely related with completion diaries, hence no substantive information is obtained. It is unclear whether early attrition would have been lower had the enrollment survey been shorter, but this is an important question for further empirical research that also should consider whether topic salience contributed to early opt-out of the study.

The opportunity to sample from an ongoing study with extensive background data about participants kept the enrollment survey shorter than it might be otherwise, and also permits a close examination of the demography of attrition. Consistent with studies of adults, boys were more likely than girls to attrite after the enrollment survey, and conditional on enrolling, girls completed more diaries than their male peers. The high frequency data coupled with diary-

specific paradata, such as interim contacts with the research team, change in Internet providers or cell phone numbers, potentially can shed new light on the origins of gender differences in attrition by clarifying under what circumstances intermittent nonresponse results in attrition. Our focus on overall compliance is but a first step toward understanding youth response behavior in high frequency surveys. Several puzzles remain, however, including the lack of associations between overall compliance and both incentive frequency and topic salience. Both outcomes may reflect the extensive heterogeneity that undergirds overall compliance measures. Nearly one quarter of teens who experienced varying levels of interim missingness took the final survey, which was rewarded with a $10 incentive, which suggests that incentives may not be inconsequential for adolescents' longitudinal compliance.

Our study has several limitations, including the narrow age range, urban bias, and focus on overall compliance despite acknowledging extensive heterogeneity in longitudinal response behavior and attrition. Summary outcomes such as total number of completed diaries underutilize the rich measurement afforded by high frequency data. Heterogeneity of interim missingness can only be understood using methods suitable for time-varying outcomes. Future research will use time-varying methods to evaluate the diverse response/nonresponse patterns and potentially clarify why topic salience was unrelated to overall longitudinal compliance; whether incentive frequency is associated with wave-to-wave nonresponse; and whether extenuating circumstances or major life events such as mortality or sickness of a family member or residential moves are associated with intermittent response behavior (Kocar 2019). It is possible that the incentive contrast (awarded after three or four consecutive surveys) was insufficient to detect associations in overall compliance, but may be clarifying about the timing of attrition in response to the reset of the incentive cycle.

References

Anderson, Monica. 2015. "How Having Smartphones (or Not) Shapes the Way Teens Communicate." *Pew Research Center*. Retrieved September 17, 2019 (https://www.pewresearch.org/fact-tank/2015/08/20/how-having-smartphones-or-not-shapes-the-way-teens-communicate/).

Anderson, Monica. 2016. "How Parents Monitor Their Teen's Digital Activities." *Pew Research Center*. (https://www.pewresearch.org/fact-tank/2016/01/07/parents-teens-digital-monitoring/).

Anderson, Monica and Jingjing Jiang. 2018. *Teens, Social Media & Technology 2018*. Pew Research Center.

Barber, Jennifer, Yasamin Kusunoki, Heather Gatny, and Paul Schulz. 2016. "Participation in an Intensive Longitudinal Study with Weekly Web Surveys Over 2.5 Years." *Journal of Medical Internet Research* 18(6):e105.

Barber, Jennifer S., Yasamin Kusunoki, and Heather H. Gatny. 2011. "Design and Implementation of an Online Weekly Journal to Study Unintended Pregnancies." *Vienna Yearbook of Population Research* 9(1):327–34.

Bergdall, Anna R., Joan Marie Kraft, Karen Andes, Marion Carter, Kendra Hatfield-Timajchy, and Linda Hock-Long. 2012. "Love and Hooking Up in the New Millennium: Communication Technology and Relationships among Urban African American and Puerto Rican Young Adults." *The Journal of Sex Research* 49(6):570–82.

Boys, Annabel, John Marsden, Garry Stillwell, Kevin Hatchings, Paul Griffiths, and Michael Farrell. 2003. "Minimizing Respondent Attrition in Longitudinal Research: Practical Implications from a Cohort Study of Adolescent Drinking." *Journal of Adolescence* 26(3):363–73.

Bristle, Johanna, Martina Celidoni, Chiara Dal Bianco, and Guglielmo Weber. 2014. *The Contribution of Paradata to Panel Cooperation in SHARE*. 19–2014. Survey of Health, Ageing and Retirement in Europe.

Buskirk, Trent D. and Charles Andres. 2012. "Smart Surveys for Smart Phones: Exploring Various Approaches for Conducting Online Mobile Surveys via Smartphones." *Survey Practice* 5(1):3072.

Callegaro, Mario. 2013. "Do You Know Which Device Your Respondent Has Used to Take Your Online Survey?" *Survey Practice* 3(6).

Coyne, Sarah M., Laura M. Padilla-Walker, and Hailey G. Holmgren. 2017. "A Six-Year Longitudinal Study of Texting Trajectories During Adolescence." *Child Development* 89(1):58–65.

Czajka, John L. and Amy Beyler. 2016. "Declining Response Rates in Federal Surveys: Trends and Implications (Background Paper)." *Mathematica*. (https://www.mathematica-mpr.com/our-publications-and-findings/publications/declining-response-rates-in-federal-surveys-trends-and-implications-background-paper).

Durrant, Gabriele and Frauke Kreuter. 2013. "Editorial: The Use of Paradata in Social Survey Research." *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 176(1):1–3.

Goldberg, Rachel E., Dawn Koffman, and Marta Tienda. 2019. "Using Bi-Weekly Surveys to Assess Adolescent Relationship Flux: Lessons from a Mobile Diary Study." *Journal of Research on Adolescence*.

Goldberg, Rachel E. and Marta Tienda. 2017. "Adolescent Romantic Relationships in the Digital Age." in *Emerging Trends in the Social and Behavioral Sciences*, edited by R. A. Scott and S. Kosslyn.

Groves, Robert M., Don A. Dillman, John L. Eltinge, and Roderick J. A. Little, eds. 2002. *Survey Nonresponse*. Vol. 51. New York, NY: John Wiley & Sons, Ltd.

Groves, Robert M. and Emilia Peytcheva. 2008. "The Impact of Nonresponse Rates on Nonresponse BiasA Meta-Analysis." *Public Opinion Quarterly* 72(2):167–89.

Groves, Robert M., Stanley Presser, and Sarah Dipko. 2004. "The Role of Topic Interest in Survey Participation Decisions." *The Public Opinion Quarterly* 68(1):2–31.

Halpern, Carolyn Tucker, J. Richard Udry, and Chirayath Suchindran. 1994. "Effects of Repeated Questionnaire Administration in Longitudinal Studies of Adolescent Males' Sexual Behavior." *Archives of Sexual Behavior* 23(1):41–57.

Hargittai, Eszter. 2010. "Digital Na(t)Ives? Variation in Internet Skills and Uses among Members of the 'Net Generation'*." Sociological Inquiry 80(1):92–113.

Hensel, Devon J., James D. Fortenberry, Jaroslaw Harezlak, and Dorothy Craig. 2012. "The Feasibility of Cell Phone Based Electronic Diaries for STI/HIV Research." *BMC Medical Research Methodology* 12:75.

Jaccard, James, Robert McDonald, Choi K. Wan, Vincent Guilamo-Ramos, Patricia Dittus, and Shannon Quinlan. 2004. "Recalling Sexual Partners: The Accuracy of Self-Reports." *Journal of Health Psychology* 9(6):699–712.

Kocar, Sebastian. 2019. "The Power of Online Panel Paradata to Predict Non-Response and Attrition."

Laurie, Heather and Peter Lynn. 2009. "The Use of Respondent Incentives on Longitudinal Surveys." Pp. 205–33 in *Methodology of Longitudinal Surveys*, edited by P. Lynn. John Wiley & Sons, Ltd.

Lenhart, Amanda. 2015. “Teens, Social Media & Technology Overview 2015.” *Pew Research Center: Internet, Science & Tech*. (http://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015/).

Link, Michael W., Joe Murphy, Michael F. Schober, Trent D. Buskirk, Jennifer Hunter Childs, and Casey Langer Tesfaye. 2014. “Mobile Technologies for Conducting, Augmenting and Potentially Replacing SurveysExecutive Summary of the AAPOR Task Force on Emerging Technologies in Public Opinion Research.” *Public Opinion Quarterly* 78(4):779–87.

Lugtig, Peter and Annelies Blom. 2018. “Using Paradata to Explain Attrition in the German Internet Panel.”

Lugtig, Peter. 2014. “Panel Attrition: Separating Stayers, Fast Attriters, Gradual Attriters, and Lurkers.” *Sociological Methods & Research* 43(4):699–723.

Post, Ann, Hans Gilljam, Sven Bremberg, and Maria Rosaria Galanti. 2012. “Psychosocial Determinants of Attrition in a Longitudinal Study of Tobacco Use in Youth.” *The Scientific World Journal* 2012:e654030.

Powers, Jennifer and Deborah Loxton. 2010. “The Impact of Attrition in an 11-Year Prospective Longitudinal Study of Younger Women.” *Annals of Epidemiology* 20(4):318–21.

Prensky, Marc. 2001. “Digital Natives, Digital Immigrants Part 1.” *On the Horizon* 9(5):1–6.

Raento, Mika, Antti Oulasvirta, and Nathan Eagle. 2009. “Smartphones.” *Sociological Methods & Research* 37(3):426–54.

Reichman, Nancy E., Julien O. Teitler, Irwin Garfinkel, and Sara S. McLanahan. 2001. “Fragile Families: Sample and Design.” *Children and Youth Services Review* 23(4):303–26.

Rideout, Vicky. 2015. *The Common Sense Census: Media Use by Tweens and Teens*. Common Sense Media.

Runyan, Jason D., Timothy A. Steenbergh, Charles Bainbridge, Douglas A. Daugherty, Lorne Oke, and Brian N. Fry. 2013. “A Smartphone Ecological Momentary Assessment/Intervention ‘App’ for Collecting Real-Time Data and Promoting Self-Awareness.” *PLOS ONE* 8(8):e71325.

Ryan, Camille and Jamie Lewis. 2017. *Computer and Internet Use in the United States: 2015*. ACS-37. United States Census Bureau.

Schoeni, Robert F., Frank Stafford, Katherine A. McGonagle, and Patricia Andreski. 2013. “Response Rates in National Panel Surveys.” *The Annals of the American Academy of Political and Social Science* 645(1):60–87.

Singer, Eleanor, Robert M. Groves, and Amy D. Corning. 1999. "Differential Incentives: Beliefs About Practices, Perceptions of Equity, and Effects on Survey Participation." *The Public Opinion Quarterly* 63(2):251–60.

Singer, Eleanor and Cong Ye. 2013. "The Use and Effects of Incentives in Surveys." *The ANNALS of the American Academy of Political and Social Science* 645(1):112–41.

Tienda, Marta, Goldberg, Rachel, Dawn Koffman, and Naila Rahman. 2018. "Opportunities and Challenges of Using Digital Technology for Intensive Longitudinal Study of Teens' Romantic Relationships." Presented at the annual meetings of the American Association for Public Opinion Research, Denver, CO.

Turner, C. F., L. Ku, S. M. Rogers, L. D. Lindberg, J. H. Pleck, and F. L. Sonenstein. 1998. "Adolescent Sexual Behavior, Drug Use, and Violence: Increased Reporting with Computer Survey Technology." *Science* 280(5365):867.

Wagner, James, Mick P. Couper, William G. Axinn, and Heather Gatny. 2019. "The Utility of a Follow-up Interview for Respondents to a Longitudinal Survey with Frequent Measurement." *Social Science Research* 82:113–25.

Watson, Nicole and Mark Wooden. 2009. "Identifying Factors Affecting Longitudinal Survey Response." Pp. 157–81 in *Methodology of Longitudinal Surveys*. John Wiley & Sons, Ltd.

Table 1. mDiary Recruitment and Enrollment Outcomes

(Medians or Percentages)

	Assented				
	Total (N=689)	Not Enrolled (N=158)	Enrolled (N=531)	On-Time Enrollment (N=484)	Delayed Enrollment (N=47)
Recruitment Experience					
Median # contact attempts post welcome package	4	5	4	4	4
[range]*	[0-48]	[0-48]	[0-35]	[0-31]	[0-35]
Assents by paper or online	23.1	19.6	24.1	23.1	34.0
Enrollment Timing					
Median # days: assent to first enrollment invite	11	12	11	11	12
[range]**	[3-143]	[4-90]	[3-143]	[3-69]	[4-143]
1st enrollment invite during summer	14.4	12.7	14.9	15.5	8.5
Teen's digital access/proficiency					
Receives text messages	92.2	79.1	96.1	96.3	93.6
Uses school email	5.8	5.7	5.8	5.6	8.5
Contacts team about security image	7.0	4.4	7.7	6.4	21.3
Enrolled using smartphone	na	na	73.1	73.8	66.0

Source: mDiary paradata

*Zero contact attempts indicates that family provided consent and assent via mailed welcome invitation

** 143 day lag reflects an protracted period to obtain email address after dropped call post-assent

Table 2. Odds of Enrollment (vs. nonenrollment)

(Reference group in Parentheses)

	Baseline		Full Model	
	Odds Ratio	S.E.	Odds Ratio	S.E.
Recruitment Experience (a)				
# Contact attempts post welcome	0.962	0.015	0.969	0.016
Assents by paper or online (phone)	1.330	0.347	1.199	0.333
Assent/enrollment invite lag (0-7 days)				
8-14 days	0.789	0.194	0.836	0.213
15+ days	0.574	0.159	0.555	0.161
1st enrollment invite during summer	1.221	0.355	1.074	0.333
Digital access/proficiency (a)				
Uses school email	0.946	0.383	0.771	0.325
Receives text messages	6.369	1.921	7.034	2.323
Contacts team about security image	1.546	0.658	1.916	0.855
Respondent attributes (b)				
Female (male)			2.247	0.467
Ever dated **(c)**			1.295	0.270
Home Environment				
Lives with both bio parents **(d)**			0.879	0.068
Has home Internet service **(d)**			1.202	0.424
Home often chaotic **(c)**			1.647	0.701
Often spends time alone **(c)**			0.933	0.271
Mother's Characteristics (b)				
Race (nonHispanic white)				
Black			0.598	0.170
Hispanic			1.523	0.476
Other			0.903	0.484
Education (< HS)				
HS or equivalent			1.492	0.386
Some college			2.087	0.581
College graduate (%)			2.871	1.209
Marital status (married)				
Living with youth father			0.935	0.273
Not living with youth father			1.446	0.444
Constant	0.750	0.319	0.316	0.232

Sources: (a) mDiary paradata; (b) FFCWS baseline; (c) FFCWS Y15-teen; (d) FFCWS Y15-Parent;

Notes: Controls for first enrollment invite quarter

Table 2a. Odds of Delayed or Nonenrollment (vs.enrollment)

(Reference group in Parentheses)

	Baseline				Full Model			
	Delayed		Nonenrollment		Delayed		Nonenrollment	
	Odds Ratio	S.E.	Odds Ratio	S.E.	Odds Ratio	S.E.	Odds Ratio	S.E
Recruitment Experience (a)								
# Contact attempts post welcome	1.047	0.026	1.045	0.016	1.041	0.028	1.037	0.017
Assents by paper or online (phone)	2.047	0.789	0.813	0.215	2.457	1.006	0.930	0.263
Assent/enrollment invite lag (0-7 days)								
8-14 days	1.050	0.427	1.280	0.318	1.107	0.460	1.218	0.315
15+ days	1.066	0.493	1.752	0.489	1.100	0.531	1.838	0.543
1st enrollment invite during summer	0.447	0.248	0.768	0.225	0.544	0.311	0.878	0.275
Digital access/proficiency (a)								
Uses school email	1.671	0.953	1.115	0.459	1.997	1.184	1.405	0.605
Receives text messages	0.493	0.323	0.145	0.046	0.428	0.288	0.128	0.044
Contacts team about security image	4.215	1.730	0.809	0.352	3.792	1.704	0.676	0.309
Respondent attributes (b)								
Female (male)					0.514	0.170	0.415	0.088
Ever dated **(c)**					0.480	0.145	0.692	0.152
Home Environment								
Lives with both bio parents **(d)**					1.162	0.148	1.158	0.091
Has home Internet service **(d)**					2.128	1.687	0.908	0.326
Home often chaotic **(c)**					1.359	0.736	0.633	0.274
Often spends time alone **(c)**					1.342	0.555	1.105	0.326
Mother's Characteristics (b)								
Race (nonHispanic white)								
Black					2.354	1.127	1.826	0.525
Hispanic					1.401	0.724	0.685	0.217
Other					1.665	1.386	1.156	0.627
Education (< HS)								
HS or equivalent					1.243	0.567	0.686	0.181
Some college					1.196	0.566	0.487	0.138
College graduate (%)					1.129	0.748	0.357	0.153
Marital status (married)								
Living with youth father					0.877	0.406	1.056	0.313
Not living with youth father					0.665	0.326	0.664	0.207
Constant	0.152	0.126	1.555	0.681	0.085	0.1171	3.640	2.741

Sources: (a) mDiary paradata; (b) FFCWS baseline; (c) FFCWS Y15-teen; (d) FFCWS Y15-Parent

Notes: Controls for forst enrollment invite quarter

Table 3. Response Behavior by Enrollment Timing

(Means or Column Percentages)

	Enrolled (N=531)	On Time Enrollment (N=484)	Delayed Enrollment (N=47)
Overall Compliance*			
# diaries completed post enrollment survey	17.1	18.2	11.2
(s.d.)	(9.3)	(9.0)	(9.7)
Last diary completed post enrollment survey *	19.4	19.9	14.1
(s.d.)	(9.1)	(8.7)	(10.9)
Response Patterns			
Completed 25 diaries (N=234)	44.1	46.9	14.9
Highly Engaged: completed 22-24 diaries (N=75)	14.1	14.5	10.6
Interim Missing: completed <22 diaries (N=51)	9.6	8.9	17.0
Attrition + Interim Missing: completed <22 diaries (N=70)	13.2	12.2	23.4
Attrition: completed < 22 diaries (N=101)	19.0	17.6	34.0
Only completed enrollment survey (N=26)**	4.9	4.3	10.6

Sources: mDiary enrollment and diary surveys

*There were 25 diary surveys and one enrollment survey

**These respondents are a subset of the attrition group

Table 4. Respondent Characteristics and Recruitment Experience by Response Behavior

(Means, Medians or Percentages)

	Fullly Engaged Completed 25 diaries	Highly Engaged: Completed 22-24 diaries	Interim Missing: Completed <22 diaries	Attrition + Interim Missing: Completed <22 diaries	Attrition: Completed <22 diaries	Only completed enrollment survey ***
	(N=234)	(N=75)	(N=51)	(N=70)	(N=101)	(N=26)
Overall Compliance (c)						
# Diaries completed post enrollment survey	25.0	23.3	14.5	9.2	3.5	na
(s.d.)	(0.0)	(0.8)	(5.7)	(5.6)	(3.9)	
Last diary completed post enrollment survey	25.0	24.9	25.0	13.4	3.5	na
(s.d.)	(0.0)	(0.5)	(0.0)	(6.3)	(3.9)	
Recruitment Experience (a)						
Median # contact attempts post welcome package	3	4	4	5	3	3
[range]*	[0-27]	[0-31]	[1-35]	[0-35]	[0-31]	[0-19]
Assents by paper or online	29.1	24.0	29.4	11.4	18.8	23.1
Enrollment Timing (a)						
Median # days: assent to first enrollment invite	11	10	11	10	11	8.5
[range]**	[4-69]	[4-25]	[3-55]	[3-34]	[4-143]	[4-39]
1st enrollment invite in summer	15.0	12.0	7.8	12.9	21.8	15.4
Digital access/proficiency (a)						
Receives text messages	97.9	96.0	92.2	94.3	95.1	100.0
Uses school email	3.4	9.3	3.9	8.6	7.9	7.7
Contacts team about security image	6.0	5.3	15.7	4.3	11.9	15.4
Enrolled using smartphone	76.1	60.0	74.5	67.1	79.2	73.1
Ever dated **(b)**	70.9	66.7	82.4	77.1	76.2	73.1

mDiary Sources: (a) paradata; (b) enrollment survey; (c) diary surveys

*Zero contact attempts indicates that family provided consent and assent via mailed welcome invitation.

**143 day lag for dropped call post-assent and before email address obtained.

***These respondents are a subset of the attrition group, and completed 0 diary surveys.

Table 5 Correlates of Completed Diaries (regression)

(Reference group in Parentheses)

	Baseline		Full Model	
	coef	**S.E.**	**coef**	**S.E.**
Recruitment Experience (a)				
# Contact attempts post welcome	-0.079	0.077	-0.040	0.077
Assents by paper or online (phone)	2.017	1.042	2.059	1.056
Assent/enrollment invite lag (0-7 days)				
8-14 days	-0.047	0.981	0.006	0.971
15+ days	1.080	1.176	0.772	1.173
1st enrollment invite during summer	-0.576	1.600	-0.588	1.592
Incentive group 4 (group 3)	1.017	0.801	1.139	0.797
Digital access/proficiency (a)				
Uses school email	-2.158	1.687	-2.701	1.675
Receives text messages	3.254	2.045	2.542	2.059
Contacts team about security image	-1.555	1.498	-0.802	1.503
Delayed enrollment	-7.051	1.417	-6.806	1.421
Respondent attributes (b)				
Female (male)			1.484	0.806
Ever dated **(c)**			1.006	0.793
High Grit			-0.415	0.344
Home Environment				
Lives with both bio parents **(d)**			-0.510	0.326
Has home Internet service **(d)**			4.002	1.651
Home often chaotic **(c)**			-0.301	1.537
Often spends time alone **(c)**			-0.904	1.119
Mother's Characteristics (b)				
Race (nonHispanic white)				
Black			-0.398	1.115
Hispanic			-0.088	1.122
Other			1.406	2.024
Education (< HS)				
HS or equivalent			2.139	1.102
Some college			2.777	1.114
College graduate (%)			3.024	1.525
Marital status (married)				
Living with youth father			0.070	1.145
Not living with youth father			-0.277	1.190
Constant	13.086	2.428	8.098	3.483
R-Sq				

Sources: Sources: (a) FFCWS baseline; (b) FFCWS Y15-teen; (c) mDiary enrollment survey; (d) FFCWS Y15-parent; (e) FFCWS Y-15-teen

Notes: Includes conrol for quarter of recruitment

Table 5a. Odds of Completing 17+ Diaries

(Reference group in Parentheses)

	Baseline		Full Model	
	Odds Ratio	**S.E.**	**Odds Ratio**	**S.E.**
Recruitment Experience (a)				
# Contact attempts post welcome	0.974	0.018	0.983	0.019
Assents by paper or online (phone)	2.037	0.557	2.295	0.679
Assent/enrollment invite lag (0-7 days)				
8-14 days	1.030	0.239	1.016	0.249
15+ days	1.178	0.336	1.069	0.323
1st enrollment invite during summer	0.955	0.363	0.944	0.377
Incentive Group 4 (group 3)	1.201	0.234	1.280	0.263
Digital access/proficiency (a)				
Uses school email	0.544	0.212	0.438	0.179
Receives text messages	2.488	1.180	2.352	1.241
Contacts team about security image	0.727	0.257	0.898	0.340
Delayed enrollment	0.252	0.085	0.235	0.085
Respondent attributes (b)				
Female (male)			1.423	0.295
Ever dated **(c)**			1.324	0.281
High grit			0.908	0.819
Home Environment				
Lives with both bio parents **(d)**			0.836	0.069
Has home Internet service **(d)**			2.088	0.852
Home often chaotic **(e)**			1.078	0.445
Often spends time alone **(e)**			0.749	0.212
Mother's Characteristics (b)				
Race (nonHispanic white)				
Black			0.931	0.269
Hispanic			0.935	0.271
Other			1.694	1.014
Education (< HS)				
HS or equivalent			1.839	0.501
Some college			2.697	0.762
College graduate (%)			2.501	1.009
Marital status (married)				
Living with youth father			0.897	0.266
Not living with youth father			1.083	0.337

Sources: (a) paradata; (b) FFCWS baseline; (c) mDiary enrollment; (d) FFCWS Y15-Parent;

Table 6 Correlates of Last Diary Completed (regression)

(Reference group in Parentheses)

	Baseline		Full Model	
	coef	**S.E.**	**coef**	**S.E.**
Recruitment Experience (a)				
# Contact attempts post welcome	-0.055	0.075	-0.038	0.076
Assents by paper or online (phone)	2.046	1.024	2.125	1.048
Assent/enrollment invite lag (0-7 days)				
8-14 days	0.912	0.964	0.991	0.964
15+ days	1.388	1.156	1.061	1.164
1st enrollment invite during summer	-0.584	1.573	-0.343	1.581
Incentive group 4 (group 3)	1.302	0.787	1.456	0.791
Digital access/proficiency (a)				
Uses school email	-1.002	1.658	-1.274	1.662
Receives text messages	0.894	2.010	0.349	2.044
Contacts team about security image	-0.894	1.472	-0.215	1.492
Delayed enrollment	-6.090	1.393	-5.876	1.411
Respondent attributes (b)				
Female (male)			0.808	0.800
Ever dated **(c)**			1.006	0.787
High grit			-0.112	0.342
Home Environment				
Lives with both bio parents **(d)**			-0.671	0.324
Has home Internet service **(d)**			3.521	1.638
Home often chaotic **(c)**			-0.677	1.526
Often spends time alone **(c)**			-0.977	1.111
Mother's Characteristics (b)				
Race (nonHispanic white)				
Black			0.887	1.107
Hispanic			0.563	1.113
Other			1.406	2.009
Education (< HS)				
HS or equivalent			1.645	1.094
Some college			2.028	1.105
College graduate (%)			2.837	1.514
Marital status (married)				
Living with youth father			0.717	1.136
Not living with youth father			0.225	1.181
Constant	16.35	2.386	11.253	3.458
R-Sq				

Sources: Sources: (a) FFCWS baseline; (b) FFCWS Y15-teen; (c) mDiary enrollment survey; (d) FFCWS Y15-parent; (e) FFCWS Y-15-teen

Notes: Includes conrol for quarter of recruitment

Table 6a. Odds of Last Diary Completed 19+

(Reference group in Parentheses)

	Baseline		Full Model	
	Odds Ratio	**S.E.**	**Odds Ratio**	**S.E.**
Recruitment Experience (a)				
# Contact attempts post welcome	0.984	0.019	0.990	0.020
Assents by paper or online (phone)	2.053	0.599	2.284	0.705
Assent/enrollment invite lag (0-7 days)				
8-14 days	1.321	0.316	1.344	0.335
15+ days	1.337	0.393	1.242	0.382
1st enrollment invite during summer	0.900	0.353	0.948	0.386
Incentive Group 4 (group 3)	1.461	0.298	1.575	0.334
Digital access/proficiency (a)				
Uses school email	0.668	0.266	0.591	0.245
Receives text messages	1.325	0.636	1.161	0.612
Contacts team about security image	0.856	0.312	1.063	0.412
Delayed enrollment	0.311	0.103	0.296	0.104
Respondent attributes (b)				
Female (male)			1.186	0.253
Ever dated **(c)**			1.280	0.275
High grit			0.951	0.088
Home Environment				
Lives with both bio parents **(d)**			0.834	0.070
Has home Internet service **(d)**			2.147	0.863
Home often chaotic **(c)**			0.929	0.385
Often spends time alone **(c)**			0.819	0.237
Mother's Characteristics (b)				
Race (nonHispanic white)				
Black			1.211	0.358
Hispanic			1.057	0.313
Other			1.968	1.269
Education (< HS)				
HS or equivalent			1.620	0.450
Some college			2.142	0.621
College graduate (%)			2.226	0.925
Marital status (married)				
Living with youth father			1.052	0.319
Not living with youth father			1.233	0.394

Sources: Sources: (a) FFCWS baseline; (b) FFCWS Y15-teen; (c) mDiary enrollment survey; (d) FFCWS Y15-parent; (e) FFCWS Y-15-teen

Notes: Includes conrol for quarter of recruitment

Appendix A. Characteristics of Recruited Teens by Enrollment Status and Timing

(Percentages)

	Assented				
	Total (N=689)	Not Enrolled (N=158)	Enrolled (N=531)	On-time Enrollment (N=484)	Delayed Enrollment (N=47)
Respondent/parent Characteristics (a)					
Female	51.4	38.6	55.2	56.4	42.6
Mother's Race					
White	29.3	21.5	31.6	32.6	21.3
Black	39.9	55.1	35.4	34.1	48.9
Hispanic	26.6	19.6	28.6	28.9	25.5
Other	4.2	3.8	4.3	4.3	4.3
Mother's Education					
Less than HS	26.1	34.8	23.5	23.8	21.3
HS or equivalent	28.9	32.3	27.9	27.3	34.0
Some college	29.8	24.1	31.5	31.4	31.9
College graduate	15.2	8.9	17.1	17.6	12.8
Mother's marital status					
Married	32.1	25.3	34.1	34.5	29.8
Living with youth father	32.1	34.8	31.3	30.8	36.2
Not living with youth father	35.8	39.9	34.6	34.7	34.0
Topic Salience/Teen Grit (b)					
Ever dated	71.0	78.5	68.7	67.4	83.0
Keeps at schoolwork until done	38.9	42.4	37.9	38.3	27.6
Sticks with plan to get something done	41.2	46.2	39.7	41.3	23.4
Finishes whatever begins	43.5	50.6	41.4	42.4	31.9
Ever dated **(c)**	na	na	73.3	72.9	76.6
Home Environment (b)					
Lives with both bio parents	35.7	24.7	39.0	39.3	36.0
Home often chaotic	6.8	5.7	7.2	6.8	10.6
Often spends time alone	14.5	13.9	14.7	14.1	21.3
Often spends time alone **(c)**	na	na	13.9	13.8	14.9
Has home Internet service **(d)**	92.4	88.0	93.8	93.6	95.7

Sources: (a) FFCWS Baseline survey; (b) FFCWS Y15-teen survey; (c) mDiary enrollment survey; (d) FFCWS Y15-parent survey

Appendix B. Characteristics of Recruited Teens by Compliance Behavior

(Percentages)

	Completed 25 diaries	Highly Engaged: completed 22-24 diaries	Interim Missing and completed < 22 diaries	Attrition + Interim Missing and completed < 22 diaries	Attrition: completed <22 diaries	Only completed Enrollment survey*
	(N=234)	(N=75)	(N=51)	(N= 70)	(N=101)	(N=26)
Respondent/Parent Characteristics (a)						
Female	60.3	52.0	45.1	57.1	49.5	38.5
Mother's Race						
White, Non-Hispanic	38.5	29.3	17.7	18.6	33.7	26.9
Black	29.5	32.0	49.0	50.0	34.7	53.9
Hispanic	26.5	33.3	29.4	31.4	27.7	11.5
Other	5.6	5.3	3.9	0.0	4.0	7.7
Mother's Education						
Less than HS	17.1	18.7	31.4	34.3	30.7	23.1
HS or equivalent	26.9	26.7	25.5	31.4	29.7	26.9
Some college	34.2	34.7	29.4	25.1	27.7	34.6
College graduate	21.8	20.0	13.7	8.6	11.9	15.4
Mother's Marital Status						
Married	40.6	34.7	27.5	25.7	27.7	34.6
Living with youth father	31.2	25.3	27.5	42.9	29.7	23.1
Not living with youth father	28.2	40.0	45.1	31.4	42.6	42.3
Teen Grit Indicators (b)						
Keeps at schoolwork until done	36.3	26.7	45.1	48.6	38.6	42.3
Sticks with plan to get something done	39.3	32.0	43.1	45.7	40.6	38.5
Finishes whatever begins	43.2	29.3	52.9	47.1	36.6	38.5
Home Environment (b)						
Lives with both bio parents	44.9	44.0	43.1	30.1	25.7	30.8
Home often chaotic	6.8	2.7	7.8	11.4	7.9	7.7
Often spends time alone **(c)**	17.1	8.0	9.8	14.3	12.9	19.2
Has home Internet service **(d)**	97.0	97.3	86.3	94.3	87.1	92.3

Sources: (a) FFCWS baseline survey; (b) FFCWS Y15-teen survey; (c) mDiary enrollment survey; (d) FFCWS Y15-parent survey

*Subset of Attrition group, column 3

Figure 1. Subject Recruitment and Compliance Behavior: mDiary Study

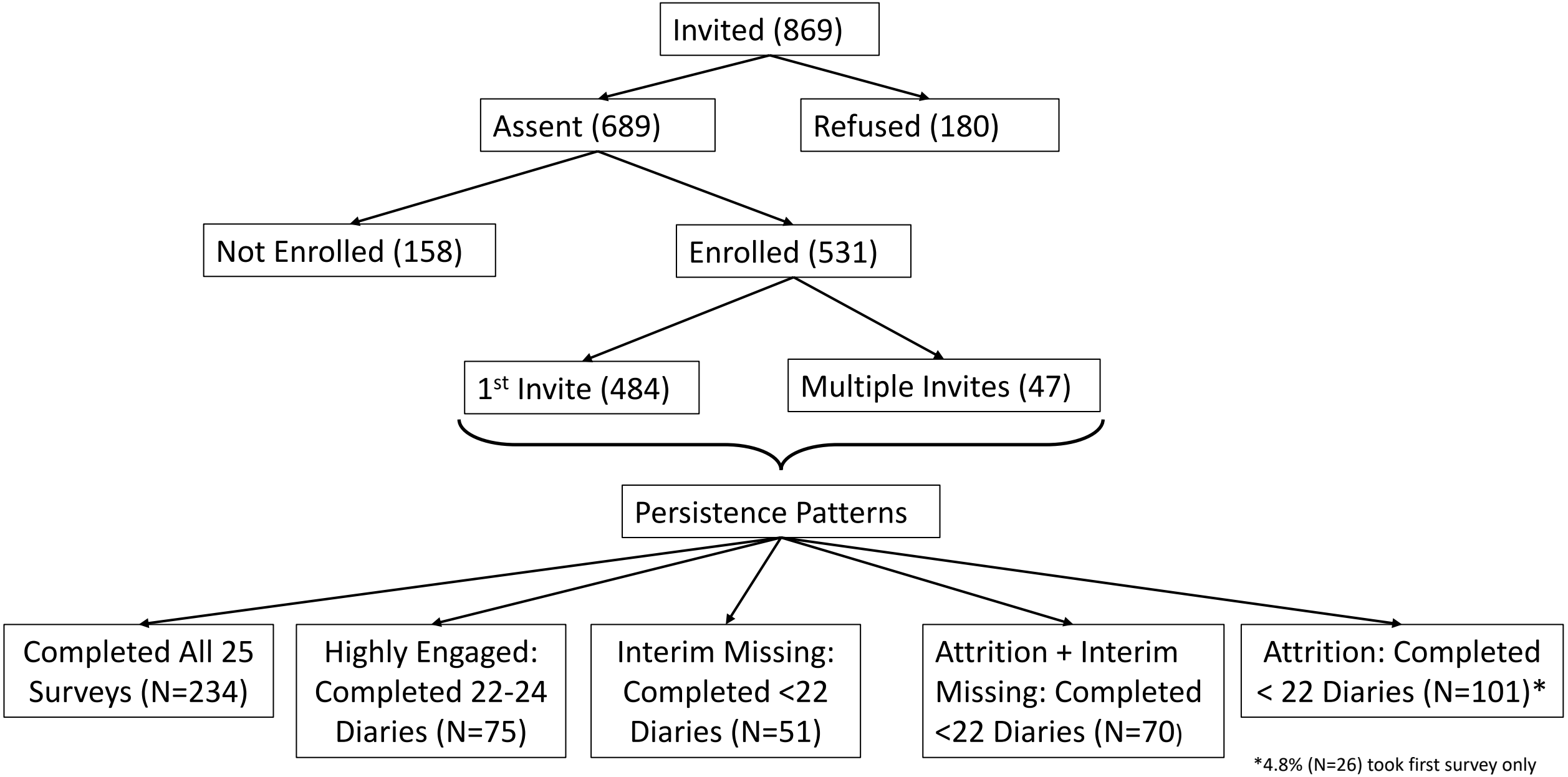

Figure 2. mDiary Retention by Enrollment Behavior: Percentages Completing at least "N" Surveys