Published online: 2020-12-15

Why Ecological Momentary Assessment Surveys Go Incomplete: When It Happens and How It Impacts Data

Yu-Hsiang Wu, MD, PhD[1] Jingjing Xu, PhD[2] Elizabeth Stangl, AuD[1] Shareka Pentony, MS[2] Dhruv Vyas, MS[3] Octav Chipara, PhD[3] Anna Gudjonsdottir, MS[4] Jacob Oleson, PhD[4] Jason Galster, PhD[2]

[1] Department of Communication Sciences and Disorders, The University of Iowa, Iowa City, Iowa
[2] Starkey Hearing Technologies, Eden Prairie, Minnesota
[3] Department of Computer Science, The University of Iowa, Iowa City, Iowa
[4] Department of Biostatistics, The University of Iowa, Iowa City, Iowa

Address for correspondence Yu-Hsiang Wu, MD, PhD, yu-hsiang-wu@uiowa.edu

J Am Acad Audiol 2021;32:16–26.

Abstract

Background Ecological momentary assessment (EMA) often requires respondents to complete surveys in the moment to report real-time experiences. Because EMA may seem disruptive or intrusive, respondents may not complete surveys as directed in certain circumstances.

Purpose This article aims to determine the effect of environmental characteristics on the likelihood of instances where respondents do not complete EMA surveys (referred to as survey incompletion), and to estimate the impact of survey incompletion on EMA self-report data.

Research Design An observational study.

Study Sample Ten adults hearing aid (HA) users.

Data Collection and Analysis Experienced, bilateral HA users were recruited and fit with study HAs. The study HAs were equipped with real-time data loggers, an algorithm that logged the data generated by HAs (e.g., overall sound level, environment classification, and feature status including microphone mode and amount of gain reduction). The study HAs were also connected via Bluetooth to a smartphone app, which collected the real-time data logging data as well as presented the participants with EMA surveys about their listening environments and experiences. The participants were sent out to wear the HAs and complete surveys for 1 week. Real-time data logging was triggered when participants completed surveys and when participants ignored or snoozed surveys. Data logging data were used to estimate the effect of environmental characteristics on the likelihood of survey incompletion, and to predict participants' responses to survey questions in the instances of survey incompletion.

Results Across the 10 participants, 715 surveys were completed and survey incompletion occurred 228 times. Mixed effects logistic regression models indicated that survey incompletion was more likely to happen in the environments that were less quiet and contained more speech, noise, and machine sounds, and in the environments

Keywords
- ecological momentary assessment
- hearing aid

received
April 21, 2020
accepted after revision
June 1, 2020
published online
December 15, 2020

Thieme Medical Publishers, Inc., 333 Seventh Avenue, 18th Floor, New York, NY 10001, USA

DOI https://doi.org/10.1055/s-0040-1719135.
ISSN 1050-0545.

wherein directional microphones and noise reduction algorithms were enabled. The results of survey response prediction further indicated that the participants could have reported more challenging environments and more listening difficulty in the instances of survey incompletion. However, the difference in the distribution of survey responses between the observed responses and the combined observed and predicted responses was small.

Conclusion The present study indicates that EMA survey incompletion occurs systematically. Although survey incompletion could bias EMA self-report data, the impact is likely to be small.

In real-world research a way to collect information on an individual's experience as it happens is ecological momentary assessment (EMA), which is also known as experience sampling or ambulatory assessment.[1] This methodology has respondents report on what happens during or shortly after they experience it in their lives. EMA can provide a wealth of information on moments in a respondent's lives without the distortions caused by recalling memories and delayed evaluation of experiences. EMA can also collect detailed contextual information with self-reported experience, allowing researchers to examine the relationship between experience and context. EMA has been validated as a research methodology in audiology,[2,3] has test–retest reliability comparable to retrospective questionnaires,[4] and may have more sensitivity to detect differences in outcomes between different hearing aid (HA) technologies than retrospective self-reports.[5] Use of the EMA methodology in audiology research has become more popular over the past decades. EMA has been employed in multiple ways to assess real-world listening difficulty or HA outcomes, including paper-and-pencil journals,[6–9] daily diaries,[10] portable computers,[11] and smartphones.[12,13]

Although the use of EMA in audiology research is increasing, EMA is not a tool without limitations. For example, the burden of EMA on respondents is high. Because (1) it is impossible to strictly control real-world conditions and (2) respondents' experience could vary from time to time even in the same conditions, EMA relies on a large amount of data from each respondent to derive a clear pattern of human experiences. Therefore, respondents often are asked to complete as many EMA surveys as possible over a long period of time (e.g., days or weeks), contributing to this high burden. The repeated assessment nature of EMA could be fatiguing, which in turn could reduce the response rate of EMA. Furthermore, because respondents are often asked to complete surveys in the moment (e.g., during a conversation) to report real-time experience, respondents could feel that EMA is disruptive or intrusive, lowering the response rate. In audiology literature, the method used to compute the response rate varies. In general, the response rate ranges from 68 to 93%.[2,9,11,14] Although some studies reported high response rates, a perfect response rate (100%) was rarely observed. Therefore, respondents do not always complete EMA surveys as they are directed.

When do respondents not complete a survey? Pentony et al[15] interviewed 19 adults after participating in a study in which EMA was used. In the interviews 12 participants indicated that the repetitive surveys interrupted or somewhat interrupted their daily activities. Surveys were particularly disruptive when the participants were working, in a class, driving, or engaging in activities (e.g., swimming, social events). Similar qualitative data have been reported by Galvez et al.[11] Therefore, it seems that, when EMA surveys are not completed, it is in a systematic manner: Respondents are less likely to complete EMA surveys in more engaging or social activities. In the present study, the circumstances in which respondents completed EMA surveys are referred to as survey completion. In contrast, the circumstances in which respondents do not complete surveys or do not want to take surveys (e.g., snooze surveys) are referred to as survey incompletion.

If survey incompletion happens systematically, how would this impact the EMA self-reported data? The impact could depend on the sampling strategy of EMA. In general, there are two main sampling strategies: time-based and event-based.[16] For time-based EMA, surveys are scheduled at regular or irregular intervals. Respondents are prompted to complete a survey by, for example, the ringing or vibration of a smartphone. Because time-based EMA is often used to obtain momentary data that would be representative of the experience across days or weeks, it is important for time-based EMA to provide unbiased samples of experience across all kinds of daily situations. In this case, survey incompletion could bias the EMA data. For example, previous research has used time-based EMA to estimate the degree and distribution of listening difficulty.[11] If respondents tend not to complete surveys in social events that are often noisy, the situations in which listening is difficult would be undersampled and therefore listening difficulty would be underestimated. Furthermore, if HA features such as noise reduction algorithms tend to be enabled in the instances of survey incompletion, EMA research (e.g., Wu et al[13]) may underestimate the effect of these features.

The effect of survey incompletion on data bias is likely to be smaller in event-based EMA. In this sampling strategy, researchers define the event of interest and respondents are asked to complete EMA surveys when the predefined event occurs. For example, Wu and Bentler[8] used event-based EMA to investigate the effect of visual cues on HA directional microphone benefit. The event of interest was the listening situation in which (1) the primary talker is in front of the listener, (2) the talker-listener distance is less than 10 feet, (3) the noise source(s) is(are) not in front of the listener, and (4) the listener can see the talker's face. If the event of interest is very specific and is clearly defined, survey incompletion would bias the data less because respondents would conduct surveys in the predefined events (although the number of responses would be lower than expected). However, if the event of interest is loosely defined and includes a wide range of situations (e.g., "Conduct a survey when you are listening to speech"), respondents may be less likely to complete surveys in certain situations and therefore bias EMA data.

To date, most data in the literature about EMA survey incompletion are qualitative interview data.[11,15] No previous research has used objective data to quantify the relationship between environmental characteristics and EMA survey incompletion. Also, although it is possible that survey incompletion could cause EMA research to underestimate the effect of HA features designed to reduce noise (e.g., directional microphones and noise reduction algorithms), no prior research has examined if these features behave differently in the instances of survey completion and incompletion. Furthermore, it is unknown how survey incompletion would affect EMA self-report data.

The first objective of the present study was to determine the effect of environmental characteristics on the likelihood of EMA survey incompletion. It was hypothesized that respondents would be less likely to complete surveys in the environments that had higher sound level (i.e., noisier) and when speech was present (e.g., social events), and in the environments in which HA noise reduction features (directional microphones and noise reduction algorithms) were enabled. The second objective of the present study was to estimate how survey incompletion would impact EMA self-report data.

Methods

The present study was part of a larger study investigating the social and auditory lifestyles of adults living in urban and rural areas. Experienced, bilateral HA users were recruited from the community. Study HAs were equipped with real-time data loggers, an algorithm that logged the data generated by HAs (e.g., overall sound level, environment classification, and feature status), were fit to the study participants. The study HAs were also connected via Bluetooth to a smartphone app, which collected the real-time data logging data as well as presented the study participants with surveys about their listening environments and experiences. The participants were sent out to wear the HAs and complete surveys as they went about their daily lives for 1 week.

Participants

Ten participants (6 males, 4 females) recruited from the area around Iowa City, IA ($n = 7$) and Berkley, CA ($n = 3$) took part in the study. Their ages ranged from 50 to 76 years with a mean of 68.2 years (standard deviation [SD] = 8.0). Subjects were eligible to participate if they were experienced HA users (at least 1 year) with bilateral sensorineural hearing loss (air-bone gap ≤ 10 dB hearing level [HL]) and no thresholds poorer than 70 dB HL between 500 and 4,000 Hz. The mean pure-tone thresholds are shown in ►**Fig. 1**. Participants also needed to be native speakers of English because speech recognition tests

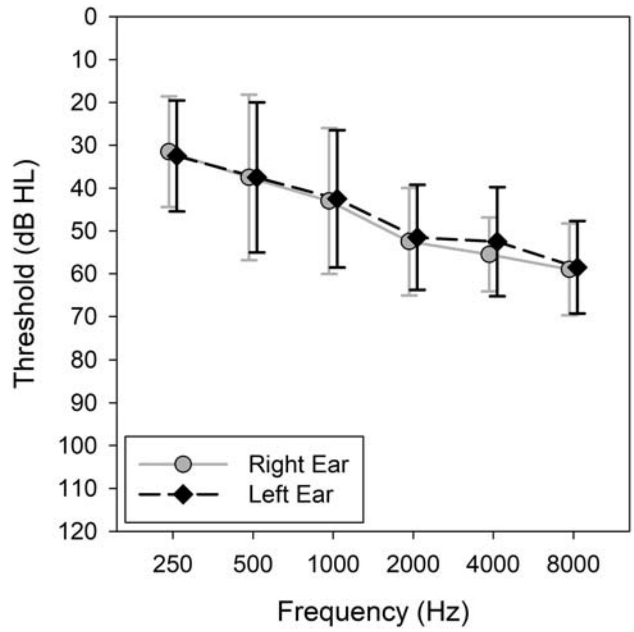

Fig. 1 Average audiograms for left and right ears of study participants. Error bars = 1 standard deviation (SD).

used in the larger study were in English. No prior smartphone experience was needed for inclusion in the study. The data of these 10 participants were initially used to conduct power analysis for the present study. Because of the large statistical effect sizes found in these data, no more participants were recruited for the present study.

HAs and Fitting

Starkey Halo 2 i2400 receiver-in-the-canal HAs were used in this study. The HAs had 24 channels and were equipped with features including wide dynamic range compression, environment classifier, directional microphone, noise reduction algorithms, and Bluetooth technology. The HAs had research firmware that could log the real-time data generated by the HAs and stream the data to smartphones via Bluetooth. The HAs were coupled to the participants' ears using acoustically appropriate, noncustom earpieces. Gain settings for the study HAs were set to match the participants' personal HAs. To do this, the real ear aided responses of the participants' own HAs were first measured with speech at 55, 65, and 75 dB sound pressure level (SPL) using a HA probe-microphone analyzer (Audioscan Verifit; Dorchester, Ontario, Canada). Additionally, the real ear aided response for an 85-dB SPL swept pure-tone (i.e., the real ear saturation response)[17] was also measured. The personal settings were documented and the study aids were set to match the real ear aided responses of each speech level and the real ear saturation response as closely as possible. The study HAs were configured to have one memory in which an automatic adaptive directional microphone and noise reduction algorithms were enabled. The manual button on each HA was set to control volume.

EMA and Smartphone App

EMA was implemented using Samsung (Seoul, South Korea) Galaxy S6 smartphones. The smartphone EMA app was developed specifically for this project and had two main functions. The first function was to deliver surveys to the participants and collect the participants' responses. In the present study, the app was configured to notify participants to complete surveys (by audio alarm or vibration) approximately every 45 minutes during a time-window specified by the participant (i.e., the time-based sampling strategy). When notified, the participants were asked to complete the survey whenever possible. The app also labeled four types of survey incompletion: Ignore, when the participants did not respond to the app notification within 5 minutes after the notification has been delivered; Exit, when participants responded to the notification and opened the app but pressed the EXIT button of the app to terminate the survey; Timeout: when participants started the survey but were unable to complete it in 5 minutes; and Snooze, when participants pressed the SNOOZE button of the app to prevent surveys in the next 30 minutes.

The second function of the app was to enable the HA's real-time data logger. When this function was triggered, the app first connected to the study HAs through a Bluetooth connection. Next, the real-time data from the two HAs were streamed to and recorded in the smartphone with a time stamp. The data sampling rate was 2 Hz (i.e., sampling twice per second per HA). The data logging data used in the present study included the overall sound level at HA's microphone (referred to as Broadband Input), environment classification (labeled as Quiet, Speech, Noise, Machine, Music, and Wind), microphone mode (directional or omnidirectional; referred to as Directivity in this article), and the amount of gain reduction in each of the 24 channels (referred to as Gain Reduction). See the "Data Processing" section below for more information about the real-time data logging data. To answer the research questions of the present study, the real-time data logger was enabled in the instances of survey completion and incompletion. For the survey completion instances, data logging started 5 minutes before a survey was scheduled and ended 5 minutes after the survey was completed. If the participants did not complete the survey (e.g., because of Ignore, Exit, or Timeout), the prenotification logged data were deleted from the app. When the app determined that the instance of Ignore, Exit, Timeout, or Snooze occurred, data logging started and lasted for 5 minutes.

Of note, is that the EMA notifications were only delivered when the HAs and the app were connected. If the HAs were turned off or if the HAs and phone were too far apart, the app would not send a notification or log HA data. Also note that the event-based sampling strategy was not used in the present study. That is, the participants were unable to initiate surveys. If they missed a survey, they needed to wait for the next survey notification.

The EMA survey was designed for the larger study. The questions in the survey focused on characterizing participants' listening environments and experience. See ►**Appendix 1** for the entire survey. ►**Table 1** shows the four survey questions used in the present study to examine the impact of survey incompletion on EMA self-report data: noisiness, signal-to-noise ratio (SNR), speech understanding, and listening effort. These four questions were selected because they were closely related to the acoustic characteristics of the environments. The survey questions were presented adaptively such that the SNR, speech understanding, and listening effort questions would be presented only when participants indicated that they were listening to speech in the beginning of the survey. Participants were instructed to answer the survey questions based on what they had experienced during the previous 5 minutes. Participants tapped a button on the smartphone app screen to indicate their responses.

Table 1 EMA survey questions and responses used in the present study

Questions	Responses	
[Noisiness] Overall, how loud were the background environmental sounds?	☐ ☐ ☐ ☐ ☐	Very loud Loud Medium Soft Very soft
[Signal-to-noise ratio] (If listening to speech) The speech of interest was _______ when compared with all other sounds.	☐ ☐ ☐ ☐ ☐	Much louder Somewhat louder Equally loud Somewhat softer Much softer
[Speech understanding] (If listening to speech) You could follow the conversation/speech.	☐ ☐ ☐ ☐ ☐	Strongly agree Agree Neutral Disagree Strongly disagree
[Listening effort] (If listening to speech) You had to strain to understand the conversation/speech.	☐ ☐ ☐ ☐ ☐	Strongly agree Agree Neutral Disagree Strongly disagree

Abbreviation: EMA, ecological momentary assessment.
Note: Square brackets show the name of the question and parentheses show the logic of the presentation of survey questions.

Procedures

The data were collected at the University of Iowa and the Starkey Hearing Research Center in Berkeley. The same protocol was used at both sites. The study consisted of two visits to the research laboratory. At the first visit, once the consent form was signed, the participants' hearing thresholds were obtained using pure-tone audiometry. Next, measurements of the participants' own HAs were made and the gain of the study HAs was matched to the participants' user settings. The participants were then instructed and trained on how to use the study HAs and the EMA smartphone app. They were familiarized with the survey questions, the laboratory smartphone, and how to take and snooze the surveys. They were also quizzed on how to use the HAs and the EMA app. All questions about using the devices were answered at this time as well. Once the participants were comfortable with the study devices and the researchers were confident in the participants' abilities, the participants were sent out to use the study HAs and take surveys as they went about their lives for the next week.

A week later the participants returned to the research laboratory. The EMA data was downloaded from the laboratory smartphones along with HA's real-time data logging information. HAs were returned to the laboratory. Study subjects were paid for their participation.

Data Processing

Before data analysis, HA data logging data were first processed. For the instance of survey completion, the data recorded in the 5 minutes before the survey notification were used. For the instance of survey incompletion, the entire 5-minute data recordings were used. In each 5-minute data recording, there were approximately 600 sets of data from each HA (i.e., two sets per second per device). Each set of data contained the momentary data of Broadband Input, environment classification label, Directivity, and Gain Reduction. Broadband Input was a continuous variable (in dB FS; decibels relative to full scale) and each set of data contained one Broadband Input value. Across the 600 Broadband Input values (i.e., across the 5-minute recording), the median was calculated, converted to dB SPL, and then used in analysis. Each set of data also contained one to three environment classification labels (more than one labels could be assigned by the HA environment classifier; e.g., Speech plus Noise). Across the 600 sets of data, the probability of a label being assigned was calculated. The calculation was conducted for each environment classification label separately, generating six variables (Quiet, Speech, Noise, Machine, Music, and Wind). Directivity was a continuous variable ranging from 0 to 1. The median of the Directivity value across the 5-minute recording was first determined. The median value was then rounded to the nearest integer such that 0 and 1 represented omnidirectional and directional modes, respectively. Gain Reduction was a continuous variable (in dB). Each data set contains 24 Gain Reduction values (24 channels). The median of the Gain Reduction values across the 5-minute recording was first calculated in each channel. The 24 median values were then averaged. The data processing was conducted for each HA separately. Therefore, in each instance of survey completion and survey incompletion, there were 18 data logging variables (9 variables for each HA) for analysis.

Results

Across the 10 participants, 715 surveys were completed and survey incompletion occurred 228 times (►**Table 2**). The average time to compete a survey was 75.2 seconds (SD = 10.3). The incompletion rate, which was defined as the count of incompletion instances divided by the count of total instances, ranged from 7.3 to 44.6% across the participants (mean = 24.2%, SD = 12.6%). Among the four types of incompletion, Ignore was the most frequent one (70.6%), followed by Snooze (20.2%). Because Exit and Timeout rarely happened, the data of the four types of incompletion were pooled. Therefore, two data sets were used in the present study for analysis: (1) the survey completion data set, which contained 715 instances and consisted of data logging data

Table 2 The counts of the instances of survey completion and survey incompletion, and the incompletion rate of each participant

Subject	Completion	Incompletion				Incompletion rate
		Ignore	Exit	Timeout	Snooze	
1201	98	6	0	0	15	17.6%
1203	56	44	0	0	1	44.6%
1301	106	4	1	4	0	7.8%
1302	80	13	0	6	4	22.3%
1304	78	15	0	0	7	22.0%
1305	26	8	0	0	1	25.7%
1309	76	4	0	0	2	7.3%
2307	54	34	1	2	5	43.8%
2308	69	15	0	1	9	26.6%
2309	72	18	5	1	2	26.5%

Table 3 Mean and standard deviation of each data logging variable of the instances of survey completion and incompletion, and the statistics of mixed effects logistic regression model examining the effect of data logging variable on the likelihood of survey incompletion

	Mean (SD)		Statistics		
	Completion	Incompletion	Odds estimation	*z*-Value	*p*-Value
Broadband Input (dB SPL)	51.9 (4.09)	59.5 (5.07)	1.08	12.19	< 0.0001
Quiet label (%)	42.1 (19.3)	20.7 (11.6)	0.14	-9.67	< 0.0001
Speech label (%)	24.1 (8.11)	33.7 (16.1)	3.45	6.68	< 0.0001
Noise label (%)	41.2 (16.9)	50.7 (13.8)	3.03	5.72	< 0.0001
Machine label (%)	3.73 (1.89)	10.2 (10.3)	8.33	5.99	< 0.0001
Music label (%)	1.73 (1.76)	4.38 (5.04)	NA	NA	NA
Wind label (%)	0.03 (0.06)	0.05 (0.10)	NA	NA	NA
Directivity	0.07 (0.06)	0.21 (0.18)	3.70	7.80	< 0.0001
Gain Reduction (dB)	0.41 (0.26)	0.91 (0.51)	0.63	–8.29	< 0.0001

Abbreviations: NA, not available; SD, standard deviation; SPL, sound pressure level.

and EMA self-report data, and (2) the survey incompletion data set, which contained 228 instances and consisted of only data logging data (i.e., no EMA self-report data).

Likelihood of Survey Incompletion

The mean and SD of each data logging variable of the survey completion and incompletion instances are shown in ►**Table 3**. To determine if the participants were less likely to complete surveys in certain types of environments or when HA noise reduction features were enabled, a mixed effects logistic regression model with random intercept for subject was employed using the entire data set (survey completion data set plus survey incompletion data set). The dependent variable was survey completion (completion vs. incompletion). The independent variables were HA (left vs. right) and data logging variable. Separate models were created for each data logging variable, except for the Music and Wind variables. The Music and Wind environments occurred so rarely that these two variables were not included in the analysis.

Results first indicated that, for all models, the effect of HA (left vs. right) was not significant (i.e., $p > 0.05$). Results further indicated that the effect of all data logging variables was significant (►**Table 3**). The odds estimation suggested the direction of the effect. For example, when the Broadband Input value increased by 1 dB, the odds of survey incompletion increased by 8% (odds ratio = 1.08). When the probability of the Speech label occurring in the 5-minute data recording increased by 1 (from 0 to 100%), the odds of survey incompletion increased by 345% (odds ratio = 3.45). These results suggested that the participants were less likely to complete surveys when the overall sound level was higher, in environments that were not quiet, in environments that contained speech, noise, and machine sounds, when directional mode was enabled, and when the noise reduction algorithm reduced more gain. ►**Fig. 2** shows the estimated survey incompletion rate as a function of each data logging variable.

Effect of Survey Incompletion on EMA Data

To determine how survey incompletion could impact EMA self-report data, the survey responses that could have been collected (if there was no survey incompletion) were estimated. Four questions shown in ►**Table 1** were used in the analysis. For the noisiness question, a linear mixed model with random intercepts for subjects was created using the survey completion data set. The dependent variable was question response (coded as 1 to 5, see ►**Table 1**) and the independent variables were the data logging variables shown in ►**Table 2** excluding the Music and Wind variables. The created model first predicted the question response (rounded to the nearest integer) for the survey completion data set. The comparison between the predicted responses and observed responses indicated that the model predicted the correct responses 46% of the time and was within one response category (e.g., "Very soft" vs. "Soft," see ►**Table 1**) 95% of the time. The model was then used to predict the question response for the survey incompletion data set.

Similar linear mixed models were also created to predict the responses of the SNR, speech understanding, and listening effort questions using the survey completion data set. The models predicted the correct responses 42 to 51% of the time and were within one response category 92 to 96% of the time. Because these three questions were presented only when the participants were listening to speech, before using the predicted response data it is necessary to estimate if the three questions would be presented to the participants in the instances of survey incompletion. To do this, a mixed effects logistic regression model with random intercepts for subjects was created using the survey completion data set. The dependent variable was the presentation of the questions (yes or no) and the predictor was the Speech variable—one of the data logging variables. The model was found to predict question presentation correctly 70% of the time. The model then predicted question presentation for the survey incompletion data set. When the model predicted that the three questions would not be

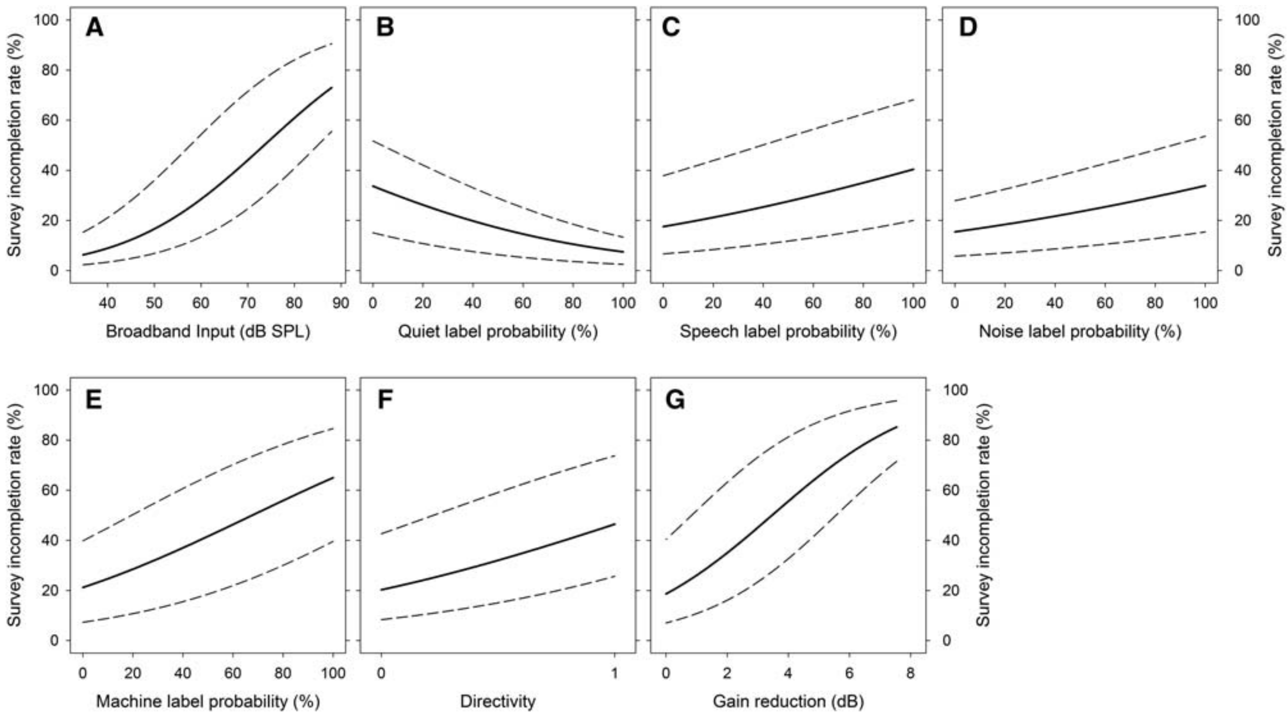

Fig. 2 Estimated survey incompletion rate (solid line) and the 95% confidence interval (dashed line) as a function of each hearing aid data logging variable. Directivity: 0 = omnidirectional mode, 1 = directional mode.

presented in an instance, the predicated responses of that instance were discarded.

►**Fig. 3** shows the results. Note that the *x*-axis of the figure is arranged such that the left side of the figure represents less challenging environments or lower listening difficulty. ►**Fig. 3A–D** show the counts of observed responses (the survey completion instances) and the counts of predicted responses (the survey incompletion instances) of each of the four survey questions. The predicted responses of the noisiness question in the incompletion instances were either "Soft," "Medium," or "Loud" (►**Fig. 3A**). For the SNR, speech understanding, and listening effort questions (►**Fig. 3B, C**), the predicted responses were mainly the second from the left response shown on the *x*-axis of the figures (e.g., "Somewhat

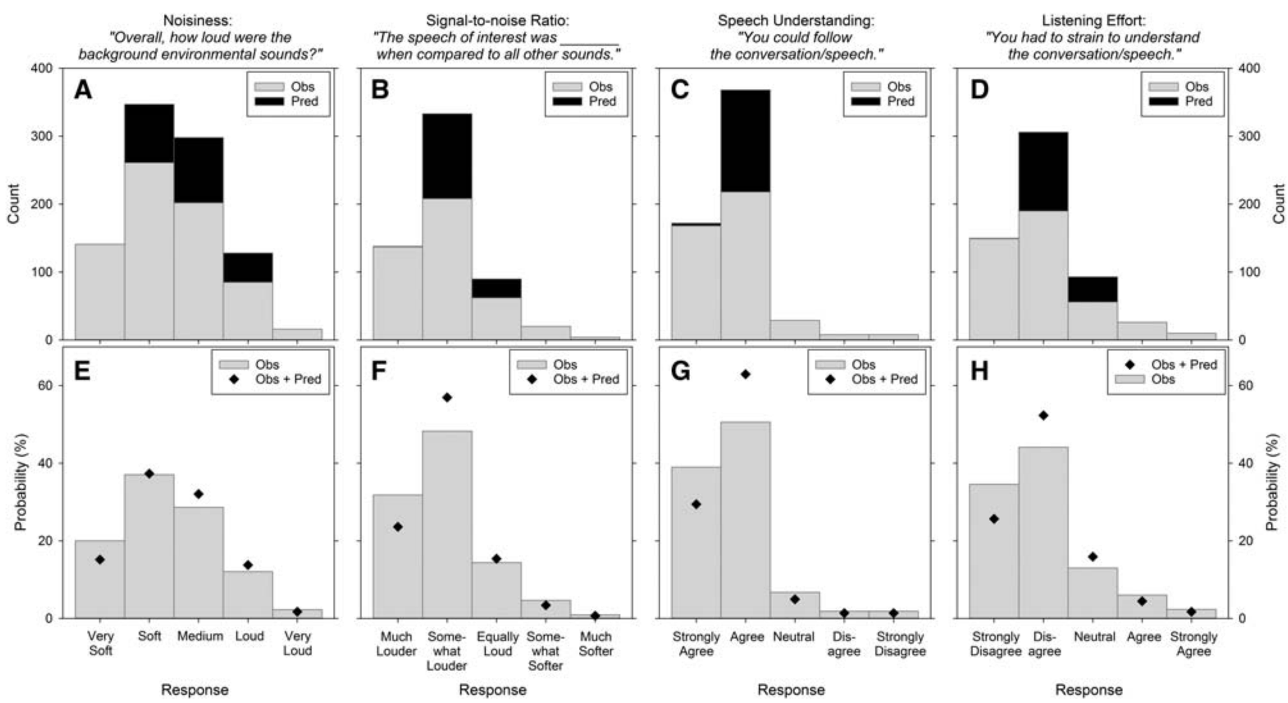

Fig. 3 Counts of observed (Obs) and predicted (Pred) responses (**A–D**) and normalized histograms of observed responses and combined observed and predicted (Obs + Pred) responses (**E–H**) of each survey question.

louder" of the SNR question in ►Fig. 3B). ►Fig. 3E–H show the normalized histograms of the observed responses (gray bars in the figures) and the normalized histograms the combined observed and predicted responses (black diamond symbols). For the noisiness question (►Fig. 3E), the participants would be less likely to select the "Very soft" response and more likely to select the "Medium" and "Loud" response if there was no survey incompletion. For the SNR, speech understanding, and listening effort questions (►Fig. 3F–H), the participants would be less likely to select the responses that represented the least challenging environments or the least listening difficulty (i.e., the furthest left response in the figure) and more likely to select the options that represented slightly challenging environments or slight listening difficulty (the response second from the left in the figure) if there was no survey incompletion. For the remaining three response options that represented more challenging environments or more listening difficulty, the difference in probability between the observed responses and combined observed and predicted responses was relatively small.

Discussion

The objectives of the present study were to determine how environmental characteristics would affect the likelihood of EMA survey incompletion and to estimate how survey incompletion would impact EMA self-report data.

Likelihood of Survey Incompletion

The participants of the present study were less likely to complete surveys in environments that had higher overall sound level (less quiet) and contained speech, noise, and machine sounds (►Table 3 and ►Fig. 2). For example, when the overall sound level increased from 40 to 80 dB SPL, the survey incompletion rate could increase from 8 to 60% (►Fig. 2A). When the probability of the Speech label increased from 0 to 100%, the survey incompletion rate was estimated to increase from 17 to 41% (►Fig. 2C). These results are consistent with the interview data[11,15] showing that respondents feel that EMA is disruptive in situations such as while working, driving, and other social events. In these situations, respondents may intentionally ignore survey notifications or snooze surveys (e.g., when driving or in a class). It is also harder for respondents to hear or feel the survey notifications in these situations (e.g., noisy social events).

Because the environments in which survey incompletion tended to occur were noisier, it is not surprising that HAs were more likely to switch to directional microphones and reduce more gain in these environments (►Table 3 and ►Fig. 2). For example, the survey incompletion rate in the environments in which directional microphones were enabled (52%) was estimated to be much higher than the rate in the environments in which omnidirectional microphones were used (20%) (►Fig. 2F). Therefore, if EMA, especially time-based EMA, is used to examine the effect of HA noise reduction features,[13] the situations in which the features would work could be undersampled.

Impact on EMA Data

Four questions about listening environments and listening difficulty were used to estimate the impact of survey incompletion on EMA self-report data. The participants' responses were in line with the literature[13,18] showing that noisy environments in which listening was difficult did not occur very often (►Fig. 3). For example, the participants reported that the background sounds were loud or very loud only 12 and 2% of the time, respectively (►Fig. 3E), and reported that they agreed or strongly agreed that listening was effortful only 6 and 2% of the time, respectively (►Fig. 3H).

The observed responses, as well as the data logging data from HAs, were used to predict the participants' responses to the survey questions in the instances of survey incompletion. Because survey incompletion tended to occur in environments that were noisier, it is not surprising that very few predicted responses were those that represented very quiet environments (e.g., "Very soft" of the noisiness question) or very easy listening (e.g., "Strongly disagree" of the listening effort question) (►Fig. 3–D). However, the responses were rarely predicted to be those that represented very challenging environments (e.g., "Loud" and Very loud" of the noisiness question) or very high listening difficulty (e.g., "Agree" and "Strongly Agree" of the listening effort question) either. This finding suggested that, even though survey incompletion tended to occur in noisier or more challenging situations, very noisy environments in which listening was very difficult still did not occur very often in the instances of survey incompletion. ►Fig. 3E–H further indicated that the histograms of observed responses and the histograms of combined observed and predicted responses did not substantially differ to each other, especially at the right tails of the histograms that represented more difficulty listening. Therefore, although survey incompletion could bias EMA self-report data, it seems that, in terms of obtaining momentary data about environmental characteristics and listening difficulty to provide a representative picture of a day or week for adult HA users, the impact of survey incompletion on EMA data are not large. The bias is more likely to impact the situations that are less challenging.

Implications

Although survey incompletion may not substantially impact EMA self-report data, it still needs to be minimized in EMA research. Providing respondents with training and instructions that emphasize the importance of reducing survey incompletion could be helpful. Offering incentives based on the counts of completed surveys could also help.

Another approach to minimize the impact of survey incompletion is to use context-weighted, time-based EMA. For example, the present study's statistical models that used HA data logging data to predict the probability of survey incompletion could be implemented to smartphone EMA app. When the app is about to deliver a scheduled survey notification, the probability of survey incompletion will be calculated and used to weight the probability of notification delivery, such that the higher the incompletion probability, the higher the probability that the notification will be delivered. This approach will not eliminate survey incompletion, but it may reduce the likelihood of undersampling certain situations. This approach, however, requires respondents to wear sensors (like the HAs used in the present study) in EMA research.

Limitations

The first limitation of the present study is that the EMA self-report data in the instances of survey incompletion were predicted by statistical models and the predictions were not perfect. Although model's predictions were within one response category approximately 95% of the time, the models predicted the correct responses only around 45% of the time. The prediction errors could affect the estimated impact of survey incompletion on EMA data. Second, only four survey questions that were more related to environmental characteristics were included in the present study to estimate the impact of survey incompletion. It is unknown how survey incompletion would affect other EMA self-report data. Third, the participants of the present study generally had high motivation to complete EMA surveys. The mean incompletion rate was 24.2% (►**Table 1**), which could roughly translate to a response rate of 75%. If respondents are less motivated and have higher incompletion rates, the impact of survey incompletion on EMA data could be much larger than what is estimated by the present study.

Further, there is room for the EMA app to improve. In the present study, the real-time data logging data generated by the study HAs were recorded after the app determined that Ignore, Exit, Timeout, or Snooze occurred. However, the data collected before these instances would carry more relevant information on the environmental characteristics leading up a survey incompletion. This is especially the case in the situations that the environmental characteristics fluctuate rapidly. Therefore, the present study may not accurately estimate the effect of environmental characteristics on survey incompletion. Finally, the environmental characteristics used in the present study were generated by the study HAs. Because HA manufacturers use different environment classification algorithms and implements features in different ways, the results of the current study (e.g., survey incompletion rate shown in ►**Fig. 2**) may not be replicable if different HA devices are used.

Conclusion

The results of the present study indicate that survey incompletion is more likely to happen in the environments that are less quiet and contain more speech, noise, and machine sounds, and in the environments wherein HA noise reduction features are enabled, suggesting that survey incompletion of time-based EMA occurs systematically. The present study further indicates that, although survey incompletion could bias EMA self-report data in terms of obtaining momentary data about environmental characteristics and listening difficulty to provide a representative picture of a day or week, the impact is likely to be small.

Funding

The present study was supported by the grants from Starkey Hearing Technologies, the National Institute on Deafness and Other Communication Disorders (R01DC015997), the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR 90REGE0013), and National Science Foundation (SCH 1838830).

Conflict of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper. The contents of this paper were partially developed under a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant number 90REGE0013). NIDILRR is a Center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this paper do not necessarily represent the policy of NIDILRR, ACL, HHS, and you should not assume endorsement by the Federal Government.

Acknowledgments

None.

References

1 Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1–32
2 Timmer BHB, Hickson L, Launer S. Ecological momentary assessment: feasibility, construct validity, and future applications. Am J Audiol 2017;26(3S):436–442
3 Wu YH, Stangl E, Zhang X, Bentler RA. Construct validity of the ecological momentary assessment in audiology research. J Am Acad Audiol 2015;26(10):872–884
4 Wu YH, Stangl E, Chipara O, Zhang X. Test-retest reliability of ecological momentary assessment in audiology research. J Am Acad Audiol, In press
5 Wu YH, Stangl E, Chipara O, Gudjonsdottir A, Oleson J, Bentler RA. Comparison of in-situ and retrospective self-reports on assessing hearing aid outcomes. J Am Acad Audiol, In press
6 Preminger JE, Cunningham DR. Case-study analysis of various field study measures. J Am Acad Audiol 2003;14(01):39–55
7 Walden BE, Surr RK, Cord MT, Dyrlund O. Predicting hearing aid microphone preference in everyday listening. J Am Acad Audiol 2004;15(05):365–396
8 Wu YH, Bentler RA. Impact of visual cues on directional benefit and preference: part II–field tests. Ear Hear 2010;31(01):35–46
9 Wu YH, Bentler RA. Do older adults have social lifestyles that place fewer demands on hearing? J Am Acad Audiol 2012;23(09):697–711
10 Bentler R, Wu YH, Kettel J, Hurtig R. Digital noise reduction: outcomes from laboratory and field studies. Int J Audiol 2008;47(08):447–460
11 Galvez G, Turbin MB, Thielman EJ, Istvan JA, Andrews JA, Henry JA. Feasibility of ecological momentary assessment of hearing difficulties encountered by hearing aid users. Ear Hear 2012;33(04):497–507
12 Timmer BHB, Hickson L, Launer S. Do hearing aids address real-world hearing difficulties for adults with mild hearing impairment? Results from a pilot study using ecological momentary assessment. Trends Hear 2018;22:2331216518783608

13 Wu YH, Stangl E, Chipara O, Hasan SS, DeVries S, Oleson J. Efficacy and effectiveness of advanced hearing aid directional and noise reduction technologies for older adults with mild to moderate hearing loss. Ear Hear 2019;40(04):805–822

14 Henry JA, Galvez G, Turbin MB, Thielman EJ, McMillan GP, Istvan JA. Pilot study to evaluate ecological momentary assessment of tinnitus. Ear Hear 2012;33(02):179–290

15 Pentony S, Xu J, Wu YH, Galster J. Using ecological momentary assessment in audiology research: the participants' perspective. Paper presented at the American Auditory Society, Scottsdale, AZ; 2018. Available at: https://haar.lab.uiowa.edu/sites/haar.lab.uiowa.edu/files/wysiwyg_uploads/pentony_et_al_aas_2018_2010.pdf

16 Shiffman S. Designing protocols for ecological momentary assessment. In: Stone A, Shiffman S, Atienza A, Nebeling L, eds. The Science of Real-Time Data Capture: Self-Reports in Health Research. New York, NY: Oxford University Press; 2007:27–53

17 Ricketts TA, Bentler RA, Mueller HG. Pre-fitting tests using frequency-specific measures. In: Essentials of Modern Hearing Aids: Selection, Fitting, and Verification. San Diego, CA: Plural Publishing; 2018:101–134

18 Humes LE, Rogers SE, Main AK, Kinney DL. The acoustic environments in which older adults wear their hearing aids: insights from datalogging sound environment classification. Am J Audiol 2018; 27(04):594–603

Appendix 1 EMA survey questions and responses

Questions	Responses	
1. Were you actively listening most of the time?	☐ ☐	Yes No
2. [If actively listening] What did your active listening involve? (select all that apply)	☐ ☐ ☐ ☐ ☐	Conversation, live Conversation, via electronic device Speech/music listening, live Speech/music listening, media Environmental sounds listening
3. [If in a conversation] Were you talking with more than one person?	☐ ☐	Yes No
4. [If listening to speech or music] What kind of sounds were you listening to? (select all that apply)	☐ ☐	Speech Music
5. Were you in wind?	☐ ☐	Yes No
6. Was there music in the background environment?	☐ ☐	Yes No
7. Were there people around you talking in the background environment?	☐ ☐	Yes No
8. Overall, how loud were the background environmental sounds?	☐ ☐ ☐ ☐ ☐	Very loud Loud Medium Soft Very soft
9. [If listening to speech] The speech of interest was ________ when compared with all other sounds.	☐ ☐ ☐ ☐ ☐	Much louder Somewhat louder Equally loud Somewhat softer Much softer
10. [If listening to speech] You could follow the conversation/speech.	☐ ☐ ☐ ☐ ☐	Strongly agree Agree Neutral Disagree Strongly disagree
11. [If listening to speech] You had to strain to understand the conversation/speech.	☐ ☐ ☐ ☐ ☐	Strongly agree Agree Neutral Disagree Strongly disagree
12. [If listening to speech] Where was the speech of interest (relative to you head)?	☐ ☐ ☐ ☐ ☐	In front In back To the right To the left Not applicable
13. [If listening to speech] Could you see the talker's face?	☐ ☐ ☐ ☐ ☐	Never/Not at all About ¼ of the time About ½ of the time About ¾ of the time All the time
14. [If listening to speech] Where were the background environmental sounds (relative to you head)?	☐ ☐ ☐ ☐ ☐ ☐	In front In back To the right To the left All around Not applicable
15. [If listening to speech] You were frustrated in this listening situation.	☐ ☐ ☐ ☐ ☐	Strongly agree Agree Neutral Disagree Strongly disagree
16. Did you make any adjustments to your hearing aids (HAs) to improve your listening experience?	☐ ☐ ☐ ☐ ☐	HAs volume up HAs volume down Remove HAs Other No action taken

Note: Square brackets show the logic of the presentation of survey questions.

