ASHA American Speech-Language-Hearing Association

AJA

Tutorial

Implementing Ecological Momentary Assessment in Audiological Research: Opportunities and Challenges

Nadja Schinkel-Bielefeld,[a] Louise Burke,[b] Inga Holube,[c,d] Maria Iankilevitch,[e] Lorienne M. Jenstad,[f] Dina Lelic,[g] Graham Naylor,[b,h] Gurjit Singh,[i,j,k] Karolina Smeds,[l] Petra von Gablenz,[c,d] Florian Wolters,[l] and Yu-Hsiang Wu[m]

[a]WS Audiology, Erlangen, Germany [b]School of Medicine, University of Nottingham, United Kingdom [c]Institute of Hearing Technology and Audiology, Jade University of Applied Sciences, Oldenburg, Germany [d]Cluster of Excellence Hearing4All, Oldenburg, Germany [e]Department of Psychology, University of Victoria, British Columbia, Canada [f]School of Audiology and Speech Sciences, The University of British Columbia, Vancouver, Canada [g]WS Audiology, Lynge, Denmark [h]National Institute for Health and Care Research, Nottingham Biomedical Research Centre, United Kingdom [i]Sonova Canada, Kitchener, Ontario, Canada [j]Department of Speech-Language Pathology, University of Toronto, Ontario, Canada [k]Department of Psychology, Toronto Metropolitan University, Ontario, Canada [l]ORCA Europe, WS Audiology, Stockholm, Sweden [m]Department of Communication Sciences and Disorders, The University of Iowa, Iowa City

ARTICLE INFO

Article History:
Received November 10, 2023
Revision received April 2, 2024
Accepted April 11, 2024

Editor-in-Chief: Peggy B. Nelson
Editor: Marc Brennan

https://doi.org/10.1044/2024_AJA-23-00249

ABSTRACT

Ecological momentary assessment (EMA) is a way to evaluate experiences in everyday life. It is a powerful research tool but can be complex and challenging for beginners. Application of EMA in audiological research brings with it opportunities and challenges that differ from other research disciplines. This tutorial discusses important considerations when conducting EMA studies in hearing care. While more research is needed to develop specific guidelines for the various potential applications of EMA in hearing research, we hope this article can alert hearing researchers new to EMA to pitfalls when using EMA and help strengthen their study design. The current article elaborates study design details, such as choice of participants, representativeness of the study period for participants' lives, and balancing participant burden with data requirements. Mobile devices and sensors to collect objective data on the acoustic situation are reviewed alongside different possibilities for EMA setups ranging from online questionnaires paired with a timer to proprietary apps that also have access to parameters of a hearing device. In addition to considerations for survey design, a list of questionnaire items from previous studies is provided. For each item, an example and a list of references are given. EMA typically provides data sets that are rich but also challenging in that they are noisy, and there is often unequal amount of data between participants. After recommendations on how to check the data for compliance, reactivity, and careless responses, methods for statistical analysis on the individual level and on the group level are discussed including special methods for direct comparison of hearing device programs.

Ecological momentary assessment (EMA) is a method for evaluating experiences in everyday life that alleviates many of the shortcomings of laboratory or clinic approaches assessing only few very specific situations and retrospective approaches that are prone to memory bias (Shiffman et al., 2008). Several studies indicate that EMA is more targeted or sensitive than retrospective assessments (Andersson et al., 2021; Wu et al., 2020) or use it because traditional methods are not sufficiently sensitive (Timmer et al., 2018a). EMA relies on self-reports completed by participants in their usual environments and in the moment, minimizing memory bias. The momentary nature of the responses allows for different situations to be evaluated separately. Nowadays, EMA is often performed on a

Correspondence to Nadja Schinkel-Bielefeld: nadja.schinkel-bielefeld@wsa.com. **Disclosure:** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

smartphone (e.g., Andersson et al., 2021; L. A. Burke & Naylor, 2020; Hasan et al., 2014; Lelic et al., 2022; Schinkel-Bielefeld et al., 2020; von Gablenz et al., 2021), which allows easy prompting and answering of in situ surveys, scalable study designs, and near–real time monitoring of data collection via cloud services.

EMA and related methods have been used in behavioral sciences since the late seventies (Csikszentmihalyi et al., 1977) but have only recently gained popularity in audiology. EMA has served a variety of research purposes in audiology (for a review of recent applications, see Holube et al., 2020). To mention a few of the most common research areas, EMA has been used to examine the auditory ecology (i.e., the environments in which those listeners are required to function; Gatehouse, 1991; Gatehouse et al., 2003; Jensen & Nielsen, 2005) of test participants, for instance, to compare across groups or across individuals (e.g., von Gablenz et al., 2021), or to better understand which situations to emulate in laboratory tests (e.g., Smeds et al., 2020). Other studies monitor how individuals cope with hearing problems or tinnitus (e.g., Henry et al., 2012; Schlee et al., 2016) and how these problems evolve over time or analyze influences of hearing problems on fatigue (L. A. Burke & Naylor, 2020). EMA has also been used to assess the benefit of an intervention such as hearing aid provision (e.g., Andersson et al., 2021; Glista et al., 2009; Jensen et al., 2019; Schinkel-Bielefeld, 2020; Wu et al., 2019) or how participants interact with assistive technology (e.g., Vercammen et al., 2021; Welling et al., 2020).

Apart from the general advantages of EMA outlined above, the application of EMA specifically in audiological research offers some additional benefits. As the acoustic environment can change drastically from one moment to the next, so can potential hearing difficulties. Therefore, EMA's ability to focus respondents' attention on the present moment is a very valuable property. It acquires additional value if collection of pertinent objective data is time-locked to self-report instances and if EMA data collection instances are linked to hearing device status. However, when attempting to exploit these audiology-specific opportunities, one faces a host of challenges, which are also more pronounced in audiology than in many other fields. These include issues of speech privacy, classification of potentially quickly changing auditory environment, and novel sampling biases arising from an elderly, potentially less tech-savvy and more conservative target population and disinclination to complete EMA surveys in some of the situations most relevant to audiology, such as group conversations.

There are several articles with guidelines for implementing EMA studies, both with general recommendations (Shiffman et al., 2008; Stone & Shiffman, 1994) and targeting specific contexts (e.g., work organization; Beal & Weiss, 2003) or populations (e.g., young people; Heron et al., 2017). While such articles are helpful in designing any EMA study, there are aspects specific to audiology not covered by those resources. This tutorial is the outcome of an online workshop series over the course of 2 years where the authors exchanged experiences with EMA and discussed implementation choices. All participants of the workshop series had experience with EMA in audiology and had published research on this topic. While we hope that through this tutorial researchers getting started with EMA can benefit from our experience, this can of course not replace solid scientific evidence that on many design and analysis questions in EMA is still missing. The current work builds on previous review articles of EMA methods in hearing research (Holube et al., 2020; Timmer et al., 2018b). While Holube et al. (2020) summarized the current state and challenges of EMA some years ago and reasonably cited Jacobs and Kaye (2015) stating that EMA in hearing research is still in its infancy, EMA has certainly now reached adolescence. The authors of the present article have been involved with EMA in one way or another. They share their experiences and cautiously provide recommendations for the use of EMA in audiological research to increase interoperability of research activities and to advance the methodology. Finally, the current article provides an update on the extant literature published in the last 4 years.

In the following sections, we describe aspects of EMA research in audiology from the last decade. For each aspect, we describe the opportunities presented by EMA, the challenges to realizing those opportunities, and (where possible) considerations for the reader based on our collective experience. This process also reveals areas in which further methodological research is needed before guidelines can be formulated. Considerations that must be made in all audiological research studies regardless of methodology are not discussed, except where these are of direct relevance to EMA-related considerations.

Some of the recommendations in this article are based on the authors' experience or opinion rather than being solely evidence based. Recommendations may not be universally applicable to all EMA studies. Hence, readers should carefully assess their research questions and decide whether to integrate the suggestions of this article into their studies based on their specific content.

Study Protocol (Design and Execution)

EMA study designs are typically very different from studies using retrospective self-report, even when the research questions are similar. Design factors for which EMA provides new opportunities and challenges include the recruitment and retention of participants, length and timing of the study, study design and sampling scheme, and survey design. Study design should include considerations for obtaining high-quality data, while minimizing participant burden and maximizing compliance. This section considers all these factors except survey design, which is discussed in detail later in a section of its own.

Opportunities

EMA provides the basis for gaining rich insights into people's hearing-related experiences, across time, environments, intents, and activities. The flexibility of design with respect to data collection period, survey timing, and survey content is one of EMA's major advantages, allowing researchers to focus on different aspects of participants' experience as desired.

The sampling scheme determines the timing of survey prompts, and the time windows after prompts within which responses are allowed (i.e., assumed to be valid). The main classes of schemes are (a) time based, in which clock timepoints (random, semi-random, or at fixed intervals) trigger survey prompts; (b) event based, in which survey prompts are triggered when objective metrics (e.g., sound level, location, movement) acquired from coupled devices fulfill specified criteria; and (c) user initiated, in which the participants themselves decide when to complete a survey (see, e.g., Holube et al., 2020, for an overview). The scheme in action in any given EMA study may be a combination of several types and multiple instances of a given type, according to study needs.

The ability to implement complex sampling schemes enables EMA studies to gather data addressing research questions with differing degrees of retrospection in parallel, potentially increasing the validity of the study outcomes. Surveys can address momentary experience ("Right now . . . "), include history ("Since the previous survey . . . " or "In the last five minutes . . . ") or even look forward ("What do you expect . . . "). Enquiring about momentary experience is particularly well suited when the focus is on sensations or attributes that are hard to remember, such as loudness. Event-based prompts can be based on data collected from sensors in hearing devices, smartphones or other devices, to sample specific environmental conditions.

For self-initiated surveys, the research aim will usually necessitate some specific instructions for when to initiate a survey. Studies can aim at covering a variety of listening situations with instructions such as "describe listening conditions that last longer than 10 minutes" (Wu & Bentler, 2012) or "all types of sound environments are relevant" (Jensen & Nielsen, 2005). They can also aim at catching specific situations not accessible to objective detection, such as situations experienced as fatiguing (L. A. Burke & Naylor, 2020) or good/difficult listening experiences (Lelic et al., 2022). In some studies, the test participants are asked to self-initiate a certain number of responses per day (Schinkel-Bielefeld et al., 2020).

EMA studies can be configured to reach large and/or diverse populations, for instance by allowing anyone who downloads an app to participate (Pryss et al., 2019; Schlee et al., 2016). For some research questions, these benefits may outweigh any loss of laboratory-standard data, such as audiograms. Even when selecting participants from a prerecruited pool, "non-contact EMA" (not involving face-to-face meeting with a researcher) can reduce the burden of participation and minimize pressure on physical research facilities (Borschke et al., 2021; L. A. Burke & Naylor, 2022).

In contrast with traditional retrospective self-report studies, EMA studies typically harvest multiple self-reports from each participant on each construct of interest. This generates a very different structure in the resulting data and thereby creates opportunities to gain insights into those constructs that are not possible with more traditional approaches. Novel aspects include temporal variation and dependence on preceding events.

Challenges

The temptation could be strong to maximize data collection period, survey frequency, and survey detail because the cost to the researcher of gathering more data is close to zero. However, the risk is that only the most highly motivated and compliant participants will complete this type of design, leading to biased results. Resisting this temptation and finding a balance between acceptable participant burden and data richness is possibly the most prevalent challenge in EMA study design. It is also recommended to keep the study design simple and to focus on a single research question, rather than trying to use the often large amount of data collected to answer too many research questions at once.

The sociodemographic composition of participant samples is often considered of secondary importance in experimental studies but is likely to be important in EMA studies, which typically probe everyday social behavior (Schinkel-Bielefeld, Lee, et al., 2022). For example, employed individuals have shown both lower EMA compliance and greater withdrawal (L. A. Burke & Naylor, 2020; Wu et al., 2021). Smartphone-based EMA is undeniably easiest when carried out with participants who are already adept smartphone users. However, several of the present authors have successfully collected EMA data with smartphone-naïve participants, although the burden for the researcher and participant might be higher (Maes et al., 2022). Even with smartphone-savvy participants, previous experience of the tasks involved in EMA leads to higher compliance and completion rates (L. A. Burke & Naylor, 2022). In contrast to lab-based procedures, participants in EMA studies are required to independently carry out the tasks, operate the equipment, and judge whether the EMA activity poses a risk to self or others (e.g., while driving or walking). This may discourage certain types of people from participating or cause them to drop out. The same is true in relation to EMA devices that alter the appearance of the wearer (e.g., neck-worn devices, cables).

The timing of an EMA data collection period is important to consider, since people's activities can vary across, for instance, the time of day, the day of week, the season, and due to personal exceptional events. Whether any of these factors present a significant risk will depend on the research question. Situations that are very infrequent (but potentially important for the participant or researcher) may not be captured, even in a long data collection period.

It is sometimes impractical, distracting, prohibited, or even unsafe to respond to an EMA survey immediately and with reference to the present moment, as it requires interrupting the current activity. This effect may be accentuated in situations that demand careful attention—which may well be the types of situations of most interest to the researcher (Schinkel-Bielefeld et al., 2020). For general self-reports, Maniaci and Rogge (2014) found that inattention when responding was associated with low internal motivation.

Recommendations

Participants

Factors that might affect the participants' likelihood of completing the study or providing high-quality data could be used as inclusion or exclusion criteria. Such criteria include tech-savviness, (low) hearing handicap (i.e., for research examining effects of rehabilitation), and issues concerning cognition (i.e., being able to understand instructions) and sensory abilities such as low vision and manual dexterity (i.e., being able to effectively use a smartphone). Alternatively, participants could be assessed on those aspects to later analyze how compliance or data quality covary with each aspect. Participants could also be assessed for personal traits that might have an impact on their self-reporting of the constructs being studied (Steffens et al., 2017).

At the end of the study, it is advised to do a debriefing or exit interview to encourage participants to report on potential problems. This can help to improve the design of subsequent studies and to understand reasons for any unexpected results.

Study Period

When scheduling participants, and if the design depends on representative sampling of "normal" daily life, researchers should ask participants about any planned special events upcoming, such as vacations or hospital stays. Then, if the timeline allows it, the EMA data collection period could be altered for those participants. Researchers should also use debriefing interviews or surveys to determine if any unexpected or exceptional events have occurred and treat these time periods appropriately in data analysis.

Length of field trial should be a balance between the time needed to answer the research question and participant burden. Three weeks with eight random prompts per day plus extra prompts in loud environments and asking the participants to self-initiate surveys was too burdensome in one study (Schinkel-Bielefeld et al., 2020). However, after a 4-week study with only self-initiated surveys in nine speaking situations per week, three out of 16 participants spontaneously offered to extend the study period during the remote exit interview and no one complained about the study length (Schinkel-Bielefeld et al., personal communication). Two weeks with six surveys per day was well accepted (L. A. Burke & Naylor, 2020).

In audiological EMA studies aiming to encompass a representative diversity of auditory ecology for individual participants, informal analyses by two of the present authors' groups have found that a data collection period of 1 week covers people's auditory reality and related everyday life routines well. A second week of data collection did not provide significant additional insights (Naylor et al., personal communication: seven to eight surveys per day; Wolters et al., personal communication: seven to eight surveys per day).

If hearing device functionality under diverse and realistic conditions is the subject of investigation, a short study period with many prompts may suffice (von Gablenz et al., 2021), especially if participants are instructed to fulfill certain tasks (e.g., going to a mall, a restaurant, a busy street) designed to encompass relevant dimensions of diversity. Weekend activities may be different from those during weekdays. Thus, it is recommended to include both if the study is shorter than a week.

If acclimatization to hearing devices or changes in behavior are of interest, or it is important to include uncommon situations, then it is vital that the study spans a sufficient time. When comparing two hearing device conditions, it is recommended to include the same days of the week in the two phases as certain activities may recur on certain weekdays. While this may suggest having a full EMA week for each phase of a study, this can result in high participant burden. It may be helpful to allow participants to choose some EMA-free days (e.g., every Monday), or to build in a break between study phases.

Sampling Scheme

For momentary assessments, it is generally advantageous if the time of the prompt is not predictable to the participants. This prevents them from altering their behavior based on prompt expectation. Prompts at random times aim to get a representative distribution of the participants' listening situations throughout the study period and to avoid sampling a regularly recurring (e.g., same time every day) activity. There can, however, also be research questions for which it would be beneficial to sample specific reoccurring events. The chance of catching uncommon events can be increased by asking participants to self-initiate surveys in those situations or by asking about experiences during the period since the last prompt (Galvez et al., 2012). In the latter case, one should consider how valid retrospective responses may be.

There should be a timeout for surveys, to prevent participants from starting the survey at one point in time and finishing it much later. This reduces the likelihood of the first part of the survey referring to a very different situation than the last part. Longer timeouts of 1–2 hr may be justified if, for example, objective data are collected between the start and finish of a survey and the data collection is supposed to also cover longer conversations (Schinkel-Bielefeld et al., personal communication). If only subjective data are captured, the survey timeout period should be shorter (e.g., 15–30 min). Based on prior experience of how long after group conversations participants typically responded, Schinkel-Bielefeld et al. (2020) selected a 30-min timeout.

Compliance, Burden, and Reactivity

As a general principle, participant compliance (the degree to which all required tasks are correctly carried out) decreases as participant burden increases (Schinkel-Bielefeld, Gotholt Madsen, & Lelic, 2022). Compliance in EMA is influenced by several factors such as carrying the equipment as instructed, the request to respond every time a survey prompt occurs, answering individual survey questions conscientiously, and alerting the researchers when problems occur. Correspondingly, efforts should be made to minimize participant burden, as poor compliance diminishes data quality. There is no clear threshold for acceptable compliance at the level of individual participants, but evidence of systematic noncompliance on some aspect across participants might indicate a threat to validity in the results. Therefore, we recommend that compliance is carefully considered. This is likely to be both quantitative (e.g., percentage of surveys completed) and qualitative (e.g., debriefing interview to identify specific situations in which the participant did not carry the equipment). A discussion about how to calculate compliance can be found in the Analysis section.

Instructions, training, and the availability of ongoing support for the participants play an important role in participant experience, burden, and compliance (Mehl & Conner, 2011; Scollon et al., 2009). T. C. Christensen et al. (2003) suggest that, if possible, all initiation sessions should be delivered by the same researcher or group of (consistently trained) researchers. Some important aspects of the instructions are listed below.

- By the end of instruction and training, participants should understand what is expected of them, as they are more likely to comply with the EMA schedule and provide useful responses if they are invested in the study (Green et al., 2006).
- The schedule, including the number of surveys to expect per day and the number of days in the study period, should be clear to the participant. Depending on study design, this is also the point at which the EMA schedule can be customized to the participant's preference. Setting the schedule in ways that work for individual participants can be helpful for compliance (Bolger et al., 2003).
- If the EMA protocol includes self-initiated surveys, participants need to understand the nature of the events of interest. Clear definition and understanding are essential for obtaining good data (Bolger et al., 2003).
- If the goal of the EMA study is to assess the frequency of particular listening situations, it is important to emphasize that repeated reports on the same type of situation are of interest, as participants may consider repeated reports in similar situations irrelevant and/or be bored by repeating their responses.

Regardless of prior level of technological competence (L. E. Burke et al., 2017), participants should receive practical training on:

- how to use the smartphone and app, including how to respond to EMA surveys;
- how to interpret survey questions and response options;
- how to use, care for, and charge the equipment, which may include a study smartphone, hearing devices, a streaming device, and other sensors; and
- what to do if problems are encountered.

Ideally, multiple modalities should be used for instructions, training, and help, including an interactive session and supplementary resources to take home for the duration of data collection (LoBuono et al., 2020). In-person sessions are ideal but can be substituted by live video-conferencing in noncontact studies. Prerecorded instructional videos may be better than written instructions for practical how-to aspects. In the absence of videos, written instructions (which should always be provided) should include plenty of visual aids to understanding.

Where possible, participants should be able to complete at least one practice survey at initiation. It may be useful for the researcher to use a prototype smartphone and app for this purpose. Otherwise, any practice responses should be removed prior to analysis. Researchers must also consider if they want to include "run-in" days or surveys in their study, as done in Wu et al. (2019), during which participants learn to navigate the EMA methodology with frequent researcher interaction (Wu et al., 2018). This refers to the first *n* days of the study, or the first *n* surveys completed, which would be considered practice days/surveys and discarded from the final data set. Unusual response patterns may occur as participants are getting used to the equipment and may need time to arrive at a stable mapping from the rating scale to experienced situations. In addition, participants' awareness of aspects relevant to the research question might change. The duration of the study and the considerations of the impact of discarding potentially useful data from those initial days should guide this decision. If the study includes a technical intervention under test (e.g., hearing aids), an acclimatization period for getting used to that technology parallel or prior to the run-in days is advisable. With respect to researcher activities to support compliant behavior once data collection has begun, recommendations include:

- Participants must be able to contact the researchers if problems arise, and preferably each participant interacts with the same researcher throughout. Contact details can be provided within the app or phone software, on the back of the phone, and/or on article instruction guides.
- Software on the phone that allows screen sharing during the call may make it easier for study personnel to give instructions, especially to less tech-savvy participants.
- It can be beneficial to proactively check in with participants during the data collection phase, especially early on. Jenstad et al. (2021) checked in with participants daily for the first few days, using the participant's preferred contact method, then weekly until the end of the study. Lelic et al. (2022) checked in with participants at least twice over a 2-week trial. Identifying participants who are unable or unwilling to comply at an early stage minimizes distress and allows researchers to legitimately discard a participant's data and recruit a replacement. Note that checking in with participants is itself a source of participant burden and therefore should not be overused.
- It can be beneficial to monitor ongoing data collection via cloud data services and follow up with participants whose data appear too sparse or implausible, to identify and rectify misunderstandings or technical issues.
- Providing a summary of individual EMA results after the data collection period can motivate participants to comply in their own interest (von Gablenz et al., 2021).

Other recommended means to improve compliance or reduce burden include:

- Surveys should be short. This may be particularly important in situations where survey completion otherwise inconveniences the participant (Schinkel-Bielefeld et al., 2020). Stone and Shiffman (2002) suggested a maximum completion time of 3 min.
- Prorated monetary incentives (i.e., paying a certain amount per response) should be used with caution. While monetary incentives can affect motivation for participating in the study, they do not necessarily ensure study compliance (Henry et al., 2012) and they could lead to false reporting.
- Participants should be able to snooze prompts, so that the prompts are presented again after a predefined duration of time (Jenstad et al., 2021), or to configure do-not-disturb times (Schinkel-Bielefeld et al., 2020). If there is no such function, irritation may cause participants to leave the phone at home or set it to silent, leading to missed surveys.
- Using the participant's own smartphone may increase the motivation to carry it and would ensure that the participant knows how to use the smartphone. However, technical problems or high battery consumption resulting from the EMA app could potentially result in frustration if it affects their own phone. If the participant's own apps drain a lot of battery, it may not last for a full day, impacting the amount of EMA data that can be collected. Participants may also be unwilling to change phone settings to the researcher's requirements. On a study smartphone, a kiosk mode, that is, a mode where only the EMA app is accessible, may reduce the fear of causing harm by pressing any wrong buttons.

- In cases when the ability to use an EMA device may be limited by physical conditions such as fine motor deficits or visual impairment, the EMA device or user interface should be adapted, if possible.

Shiffman et al. (2008, p. 20) defined *reactivity* as "the potential for behavior or experience to be affected by the act of assessing it." In practice, reactivity has not been found to be a significant issue in EMA studies in general (Barrett & Barrett, 2001; Shiffman et al., 2008; Trull & Ebner-Priemer, 2013) or in hearing research (e.g., L. A. Burke & Naylor, 2020; Jenstad et al., 2021; von Gablenz et al., 2021). Although reactivity seems to be a small effect, steps should be taken to reduce the likelihood of it occurring, including a variable prompting schedule for phenomena sensitive to reactivity (Bolger et al., 2003) and minimizing the number of prompted surveys (Liao et al., 2016). We recommend that the risk of reactivity be considered and its occurrence monitored (e.g., via pre–post comparison of a reference measure independent of the study outcome measures and/or during debriefing interview) and reported. Alternative post hoc methods are described later under the Analysis section.

As with any type of participatory study, thorough piloting is essential to resolve inadequacies, starting, for example, with the researchers themselves, then semi-colleagues, and finally a few representative participants. At each stage, the complete protocol should be evaluated.

Technical Platform

In this section, technical recommendations for conducting EMA research in audiology are provided. Given the ubiquity of smartphones and recent advancements in the development of publicly available EMA application solutions, we focus on electronic EMA solutions. EMA studies in the field of audiological research have combined participants' momentary self-reports with the additional collection of objective, often acoustical, data from various sources (Holube et al., 2020). These objective data are typically collected automatically without the need for participants' interaction with the EMA system. Self-reports are an essential part of EMA methodology (Shiffman et al., 2008), and consequently, the recommendations summarized in this section presume that self-reports are a mandatory element of EMA study designs.

Opportunities

Objective acoustic data can be recorded through hearing devices (Aldaz et al., 2016; J. A. Christensen et al., 2019; Jensen et al., 2019; Timmer et al., 2017), smartphone devices (Aldaz et al., 2016; Mehl, 2017), noise dosimeter (Wu & Bentler, 2012; Wu et al., 2015), ear-level microphones (Kowalk et al., 2020; Pohlhausen et al., 2022), or audio recorders (Hasan et al., 2017; Jensen & Nielsen, 2005; Li et al., 2014; Wu et al., 2018).

EMA research involving hearing devices brings the advantage of a device that can monitor objective data without additional equipment. A hearing device's built-in detectors can be used to estimate the sound level, whether any speech is present in the environment, and the sound environment classification (Jensen et al., 2019; Lelic et al., 2022; Schinkel-Bielefeld, 2020; Schinkel-Bielefeld et al., 2020). In addition to microphones, hearing devices may contain other sensors such as accelerometers or gyroscopes. Given the active research on the use of electroencephalogram or electromyography for predicting auditory attention, more novel sensors may become available in hearing devices in the future. Furthermore, data describing the devices' momentary performance (e.g., active features) can also be gained.

Objective data may replace some subjective assessments, or they might be used to classify the environment in which the participants make a self-report. Objective data such as estimated overall sound levels, signal-to-noise ratio (SNR), and/or sound classes can also be used to trigger surveys in specific events (Jenstad et al., 2021; Schinkel-Bielefeld et al., 2020).

For studies in everyday life, where the element of a controlled trial is lost, objective data can be used to verify subjective reports. Inconsistencies are sometimes found. For example, a participant might indicate that she is in a quiet home environment when the objective data point to a loud traffic environment. If such inconsistencies are found, attempts should be made to investigate them, for example, during a debriefing interview. When logging hearing program or hearing aid volume, these parameters can be taken into account during data analysis. This way they can be controlled for without the need to keep them constant for all participants.

Moreover, long-term objective data allow for observing trends in the acoustic environments where interaction with an EMA app is less frequent. Based on objective hearing aid data, Schinkel-Bielefeld et al. (2020) observed that speech-in-noise situations were underrepresented and quiet situations were overrepresented in subjective assessments relative to the entire wear time. Finally, data from hearing devices can be used to gain insight into performance of the device's features during critical listening moments (Lelic et al., 2022).

Using a smartphone app rather than pen-and-paper surveys offers the opportunity for adaptive surveys, where responses to one question determine which questions appear next. Additional opportunities lie in the possibility to change the questions and the sampling scheme remotely for different phases in the study or for troubleshooting. Software-based surveys allow for the expiration of prompts; that is, the survey times out after too long has passed since the prompt. When using a manufacturer-specific app, the app may be programmed to also function as a remote control for hearing devices, allowing for functions that may benefit the participant. Finally, there might be other possibilities, such as gamification, that have not yet been used for audiological research.

Challenges

Many individuals with hearing impairment are older adults who, as a group, have less experience with smartphones and potentially have deficits with vision and motor skills that make smartphone use challenging. Test participants with hearing impairment may not hear the prompt to complete a survey.

If collecting objective data at high sampling rates, the data will require a lot of storage space on the smartphone. The demand for memory will be further increased if photo or video data are stored. Although EMA research can benefit from hearing device parameters, such as sound levels or classification of the situation, the acoustic detectors in the hearing devices are usually designed to steer the functionality of the hearing device and cannot easily be tailored to specific research questions or compared between manufacturers.

Recommendations

Mobile Devices

The EMA app can either be installed on the participant's own smartphone or on a dedicated study phone. As discussed in the section about study design (Compliance, Burden, and Reactivity), both strategies have pros and cons. Other devices, such as smartwatches, can also support prompts and simple responses (Ponnada et al., 2021).

We cannot make recommendations regarding type of phone and operating system for EMA studies. A noncontact EMA study, where participants used their own smartphones, found that technical problems were less frequent for participants who used smartphones with tightly governed operating systems (i.e., iPhone iOS; L. A. Burke & Naylor, 2022). However, if a dedicated study phone is used, it is important to consider the type of phone the participants are used to. More training might be needed if an unfamiliar operating system is used. Finally, the use of a dedicated study phone has the further advantage of minimizing differences in specifications between phones, which could potentially have an impact on objective data measured by the phone (e.g., physical step-counting).

If different types of data are collected, for example, subjective assessments as well as objective recordings, it is possible to combine the use of different apps for EMA responses and objective data collection in the background as long as participants do not have to switch between apps and the apps do not hinder each other to collect data when active simultaneously. Cables and plugs into the smartphone should be avoided if possible.

The device screen size should be sufficiently large to provide readability and allow participants to use the on-screen keyboard to type survey responses. Smaller screens or smartwatches are acceptable if only very simple questions are posed with yes/no answers or a brief rating scale (Manini et al., 2019; Ponnada et al., 2021).

Ideally, the accumulated responses from a participant's test period should be storable on the device's internal memory, even if data are regularly uploaded to a secure server. This is useful in case a participant spends large amounts of the study time in an area with insufficient internet reception. The device memory must handle the expected amount of data, particularly if objective data are being collected. As an alternative, cloud storage should be considered, which might reduce storage demands. Nevertheless, data storage on the smartphone should be possible in case internet access is not available. Since Wi-Fi access cannot be assumed for all participants, SIM cards might be required. In terms of the device's power consumption, it is required that the battery lasts throughout the participants' waking hours and is fully rechargeable when not used.

For studies involving prompted surveys, the output level of the device must be sufficiently high to make prompts audible, even in environments of high sound levels or when the smartphone is worn beneath clothes or in a bag. Haptic feedback (vibration), visual notification, or streaming of the signal to hearing devices becomes essential for participants with severe hearing loss. Notifications streamed to hearing devices additionally allow the notification to go unnoticed by bystanders. This may make it easier to bring the study smartphone and keep alarms enabled in situations where an audible alarm might be considered inappropriate. Incident reports of technical issues experienced by individual participants should be kept, such that future versions of the setup can be improved (L. E. Burke et al., 2017; T. C. Christensen et al., 2003).

Objective Data

Storing audio recordings of everyday situations could be ethically problematic if people who did not consent are audible. In several countries, such recordings are forbidden for privacy reasons. Instead, one can store smoothed or down-sampled features of the input such as sound levels, estimated SNRs, or sound classifications (Bitzer et al., 2016). This can be done with automatic methods that extract relevant parameters. Bitzer et al. (2016) proposed a framework in which a temporal weighting of successive frames with a time constant of 125 ms was applied during calculation of the power spectral density to prohibit any speech intelligibility. Recommendations for how to analyze and report acoustic data can be found in the Analysis section.

The position in the hearing device architecture where the recording was taken from should be reported, that is, whether the objective data were captured before or after processing in the hearing device (e.g., directionality, noise reduction, compression). Furthermore, we recommend verifying the relationship between the objective data from hearing devices and the acoustic input with a variety of stimuli prior to the study.

When audio recordings or features derived from those recordings are used in a study, the sound quality of the recording and placement of the recording device can affect the results. This must also be considered during the training phase. For example, Wu et al. (2018) instructed their participants how to wear the audio recorder, especially regarding the orientation of the microphone (e.g., to always keep the microphone facing outward and not under clothing).

One option to facilitate environment characterization is to take a photo or a short video of the scene when making an assessment. Manual or automatic face smearing may then be necessary for compliance with data privacy regulations.

EMA Applications

EMA applications must account for a variety of factors described in detail throughout this review: the design and setup of the study, data collection, and data analysis. In general, three approaches can be distinguished: (a) using a diary or logbook application paired with a reminder application, (b) using a publicly available or open-source EMA application (e.g., olMEGA, LifeData, Ethica), or (c) using a manufacturer-specific application (e.g., Widex/Sivantos EMA app, Sonova MobEval). Apps that can collect manufacturer-independent objective data allow for the comparison of different devices or studies without hearing devices when objective data are of interest.

Timestamps for each assessment and the objective data should be stored alongside the data themselves. The type of response, either self-initiated or prompted, should be identified. The ability to customize the survey notification modality (haptic, visual, or auditory) may be helpful. Pilot testing of EMA applications for technical and user-related robustness is essential.

Surveys

Aspects of EMA survey design and content are generally tailored to the specific research question(s) of individual studies. Consequently, the comparison of outcomes across EMA studies, even when similar concepts are assessed, is challenging. A compilation of recommendations for design and content may improve the validity of individual studies and facilitate the comparison of findings across studies. The recommendations compiled in this section are based on experiences from many EMA studies executed by the present authors.

Opportunities

The primary opportunities offered by surveys in a smartphone-based EMA context are those of programming flexibility, different trigger options (e.g., self-triggering, triggering at different times of the day, or triggering based on features of the environment observed by sensors), including special prompts in certain acoustic environments (e.g., loud noise), branching of questionnaires, and adaptive behavior (e.g., questions can be omitted when not relevant). Schinkel-Bielefeld et al. (2020) have shown that offering a short mandatory and long optional version of the same survey may offer a good balance between getting detailed information for most situations and not imposing too much burden. This approach would still provide the most essential information in situations where compliance with longer surveys may be difficult, as, for example, speech in noise or group conversations.

Challenges

With smartphones, limited screen space means that items must be brief. Scrolling to see all the response options of a single question or to get to the next button is not intuitive for participants and requires more training. Many different response options are available (see the Survey Item section); however, it is unknown which response options are best suited for different contexts. Finally, a further challenge is the inability of the experimenter to clarify meaning of survey questions and response options once participants are out in the field. This speaks to the development of careful instructions and ensuring that survey items are piloted with naïve participants.

Recommendations

Survey Item Design

Here, we describe different question types and considerations. There are different response types typically employed in EMA. Single-choice, slider questions or scale questions using stars or emojis are often employed for ratings. Multiple-choice questions are often more suitable for context questions, such as the listening situation. While sliders have the advantage that only labels for the end points are necessary, single-choice questions that offer a label for each option might be interpreted more similarly by different participants (Marks & Gescheider, 2002) and are easier to deal with in case of limited dexterity or vision. Ideally, a question and all its response options should fit on one screen.

Free-text responses can be used if qualitative responses are central to the study or to obtain a response that does not fit any of the provided response options in a named category question. This is beneficial in initial studies when not all relevant response options are known in advance. When both options (predefined responses and free text) are provided, participants should be instructed to tick the response option if applicable and only use the "other/further detail" if necessary.

Studies may include comparisons between hearing device programs. This can be done in direct paired comparisons, where participants switch between two listening programs and indicate which program they prefer (Smeds et al., 2021). Using simulations, Leijon et al. (2019) showed that the inclusion of graded/ordinal responses increased the precision of the measurements considerably. Inclusion of "cannot decide" or tied responses also tended to improve precision. Alternatively, comparisons can be made using indirect comparisons, where participants rate one listening program and later the other listening program (for instance, using one program one day and the other program the next day). Afterwards, the researcher can compare these two ratings. Direct paired comparisons have the advantage that both programs are evaluated in the same situations, whereas indirect comparisons are quicker to perform and can be perceived as less burdensome by participants (Schinkel-Bielefeld, Gotholt Madsen, & Lelic, 2022).

If the participant has difficulty with one question in a specific situation, allowing them to skip that question ensures that subsequent questions can be answered instead of aborting the whole survey. However, it is recommended to make at least the central questions of the study mandatory. If questions with slider responses are included as optional questions and the slider cursor has an initial placement, it should be ensured that the researcher can distinguish between a skipped question and a question that is answered with the initial value of the slider. Hence, movement of the slider should be necessary for the answer to be accepted.

Content

We do not yet have sufficient data in our research field to recommend a core set of survey items to be included in all EMA studies. Instead, in the Appendix, we provide examples of survey items that have been implemented in various EMA studies and proven to provide valuable data. Examples have been grouped in three categories: (a) questions useful in most audiological studies, (b) questions useful for studies using hearing aids, and (c) questions for specific research questions. Each example includes a reference to the originating study and a list of references that have used similar question content. If the examples were originally not in English, translated questions and response options from the publications have been used.

Often a last optional question giving participants the opportunity to tell the researcher anything they consider relevant is a good idea. This can contain valuable information from reporting errors, difficulties with EMA equipment or hearing devices, to additional information that may help interpret a particular response.

Almost all EMA studies in audiology require categorization of the listening situation for each survey. In order to make EMA data comparable across studies, we recommend that data on experienced listening situations be organized in a way that is compatible with the Common Sound Scenarios (CoSS) framework (Wolters et al., 2016), which categorizes listening situations based on contextual classification of intention and task. The CoSS framework can be incorporated into the study design, such that options of listening activity reflect those in the framework. Alternatively, the acquired data can be recategorized into CoSS categories in the analysis stage (e.g., von Gablenz et al., 2021).

Given the ongoing burden of EMA studies and the repeated nature of data collection, so-called careless responses may be a bigger problem in EMA than in traditional survey or online survey studies (Scollon et al., 2009). If there are concerns about the possibility of careless responding to surveys, one approach is to explicitly ask participants if they were paying attention to the questions, as done by Eisele et al. (2020), or to introduce trick questions that appear occasionally, such as asking the participant to select Option 1 and if the participant selects anything else, it could be deemed as a careless response. However, including such trick questions should be kept to a minimum as there is a fine line between attention check and policing (and irritating) the participant.

Analysis

In this section, issues to consider when analyzing EMA data are outlined together with what has been done so far and coupled with the authors' recommendations for handling this type of data.

Opportunities

EMA studies typically provide large data sets with multiple sources of data. These rich data sets can often be analyzed for multiple research questions. Due to the large number of responses, they can be analyzed for different situations separately, revealing potentially opposite effects in different situations that may cancel each other out when pooling over all situations. In addition, they allow for the analysis of longitudinal processes. Data can be analyzed at the group level as well as for individual participants.

Challenges

The amount and quality of EMA data can vary drastically between individual participants, and this needs to be addressed during data analysis. Compliance, reactivity, sampling strategies, EMA equipment, and lifestyle itself are among several factors that can influence data quality. A good study design, as discussed in previous sections, can overcome some of these issues, but not everything can be controlled for.

EMA studies typically include both subjective and objective data. Subjective EMA data can be nominal, for instance, categorizing the experienced environment in predefined categories (e.g., Smeds et al., 2020) or ordinal, for instance, rating listening effort in the experienced situation (e.g., von Gablenz et al., 2021). Researchers and statisticians are debating if powerful analysis methods suitable for interval data can be used for ordinal data. Some researchers argue that parametric analyses should not be used for ordinal data (e.g., Liddell & Kruschke, 2018; Svensson, 2001), whereas others argue that parametric analyses can be used when certain assumptions are met (see Oleson et al., 2022).

Because objective data are mostly collected with sensors external to the smartphone, data transmission is not always guaranteed, which creates a risk of large amounts of missing data. Data analysis approaches for EMA studies, particularly those involving objective data, are still in their early stages. Various analysis approaches have been applied, and there is no clear consensus yet.

Recommendations

Data Checking and Cleaning

The amount of analyzable data and data loss should be documented.

Compliance. Calculation of survey compliance should take into account the sampling scheme and instructions given to the participant (Holube et al., 2020). Compliance has typically been based on proportion of reports answered as a function of expected number of prompts. As EMA studies often include a combination of prompted and self-initiated responses, it has been a challenge for the research community to agree on the most appropriate way to calculate compliance. Compliance should be reported as number of completed surveys relative to the expected number, if applicable separately for different kinds of reports (e.g., prompted and self-initiated). The timing of missed surveys (e.g., if they were always during the same time window each day) should be considered in interpretation of compliance.

When the EMA study is based on self-initiated surveys, numerical compliance may be difficult to calculate and should be based on whether the participants performed the task as instructed. For example, if the task was to complete three surveys each day, compliance is easily calculated. If the task was to complete a survey whenever in a noisy environment, it is less straightforward. While subjective data cannot verify the surveys were completed in the correct environments, communication with the participant during debriefing can achieve some level of confidence.

In any case, the assessment of compliance should not only be based on numbers. Reasons for missing data should be investigated and reported too, to cover the multiple dimensions of compliance described above under the Study Protocol section.

Researchers have several options to deal with participants with low compliance. First, the low compliance could be ignored if the research question does not depend on compliance. Second, participants with poor compliance could be excluded from analysis. If this is the case, there should be an a priori decision made about when and why to exclude participants. For participants with poor compliance, it is important to understand whether the available observations represent their true trajectories. If so, their small number of observations could be seen as a random sample of the larger number of observations they could have made. A small number of observations for some participants could mean that their results will not weigh as heavily toward the overall results as the results of a participant with many observations.

Careless responses. Even if participants answer many surveys and seemingly have a high compliance, it may be that they answer those questions carelessly (responding habitually or randomly), leading to poor data quality. According to Welling et al. (2021), carelessness may look like straight lining (i.e., giving the same response over and over), repeated patterns (such as up-down-up-down no matter what the question or condition is), or random answers.

Careless responses may be detected by examining expected response patterns. Welling et al. (2021) suggest two methods for detecting probable careless responses: speed of the response (outliers that are exceptionally fast or slow) and consistency across responses (a lack of consistency under similar conditions could indicate careless responses).

It is important to screen for careless responses and discuss their impact on the results. Debriefing interviews could be used to understand what caused them. However, the study by Welling et al. (2021) showed that careless responses did not occur frequently. One reason could be that it is easier to skip a prompted survey than to respond carelessly.

Reactivity. A way to check for reactivity effects post hoc is to evaluate the EMA data for trends and response changes over time (both within and across participants; Barta et al., 2012). Note that there could be other effects over time such as acclimatization or study fatigue as well. Henry et al. (2012) examined individual tinnitus index scores in their chronological sequence graphically and classified the data as improving, worsening, or consistent trends. Von Gablenz et al. (2021) evaluated individual EMA data graphically on speech understanding, listening effort, and disability in their chronological order to assess trend. Additionally, statistical analyses were performed to estimate trend effects based either on the survey sequence or on the very first and the very last assessments for similar everyday listening tasks.

Descriptive Statistics

Descriptive statistics summarize, for instance, proportions of selected categories, distributions of data, and comparisons of conditions. Factors to consider when reporting descriptive statistics in EMA research are described in the following subsections.

Acoustic data. Objective data can be analyzed across the entire study period (e.g., when analyzing auditory ecology) or in relation to the EMA surveys. When objective data are not recorded simultaneously with subjective data, information about the time delay between them should be taken into consideration when analyzing the momentary objective data. Even when objective and subjective data are recorded simultaneously, the analysis approach should take into consideration that the participant may be responding in relation to a particular portion of the objective data's sampling time window. Often a short period prior to the start of the survey (Andersson et al., 2023; Schinkel-Bielefeld et al., 2020) or at delivery of the prompt (Jenstad et al., 2021; Jorgensen et al., 2023) or the time period during survey completion (Andersson et al., 2021) is used. It is also possible to ask participants to enter the time the rating is referring to (Borschke et al., 2024). Objective data can be reported descriptively, or they can be considered in statistical modeling, for example, in linear mixed-effects models.

To enable comparisons across studies, it is desirable to report the same physical measures for acoustic data. For example, broadband levels should preferably be given with A-weighting, as this is matched to human loudness perception. If that is not possible, the bandwidth and the frequency response of the recording chain should be documented. The length of the averaging time window used for the acoustic measurements or the use of asymmetric time constants (to estimate, e.g., peak values) will impact the outcome; therefore, for metrics derived on fluctuating signals, the bandwidth, weighting, and averaging approach should be reported.

Given recorded audio from an EMA participant's daily environment, it is sometimes desired to estimate time-varying levels of component signals in a mixture (e.g., other people's voices vs. wearer's own voice vs. other sounds). This is very challenging without a highly controlled (i.e., laboratory) recording setup. Wu et al. (2018) extracted (other persons') speech + noise and noise-only segments from their recorded data by listening to the audio signals. Levels and SNRs were then estimated using a power subtraction technique, similar to that used by Smeds et al. (2015). The method suffers from bias at low SNRs. Brungart et al. (2020) had no audio recordings available but relied on time-dependent level measurements of the mixture of environmental sounds for SNR estimations. They estimated levels of target and background sounds by comparing the level recordings from real-world environments to level recordings of mixed signals with known SNRs in the laboratory. This approach cannot eliminate bias from unknown periods of (possibly loud) wearer's own voice (Pohlhausen et al., 2022; Ryherd et al., 2012). Some hearing aids also enable detection of the wearers own voice, such that data with own voice can be excluded when estimating environmental SNR (Schinkel-Bielefeld et al., 2023). All SNR estimates suffer from the problem that the signal of interest to the participant is not necessarily known. Exact algorithms of SNR estimates in hearing aids are usually tied to hearing aid functionality and hence confidential, which makes it difficult to compare estimates across different studies.

Unequal amount of data. A potential imbalance in the number of reports across participants needs to be considered when summarizing the data. One way to do this is to reduce the data into one representative number per participant such as a mean, median, or proportions (e.g., Oleson et al., 2022). This approach reduces complexity and allows a concise presentation but disregards within-participant variability. Another way for compensating imbalance is to use case weights as done by von Gablenz et al. (2021). For both aggregation and weighting, it is important to consider whether a minimum number of responses per participant should be required. If weights are applied, it is recommended to check and report the share of markedly low and high weighting factors to characterize the weighting model.

Variability. Variability of data within and between participants as well as listening situations can give another layer of information about people's lifestyle and should be reported. These types of data can give an understanding of (a) whether the person's day-to-day life is dynamic or consistent and (b) which observations are person specific and which persist on the group level.

Substantive Analysis

In this section, we outline three recommended approaches for analyzing EMA data. The best approach for any given study will depend on the nature of the data available and the specific study design.

Nonoverlap of all pairs. Von Gablenz et al. (2021) examined individual benefit of a hearing aid intervention using the "nonoverlap of all pairs" (NAP; Manolov et al., 2016; Parker & Vannest, 2009; Parker et al., 2011) method, an analysis method developed for single participant research that provides an estimate of the probability that a participant's ratings are different across phases, for example, before and after an intervention. The NAP outcome is a value between 0 and 1, where the two extremes indicate that every response from one phase is higher or lower than any response from the other phase (no overlapping data). A NAP value at 0.5 indicates the chance level, which should be interpreted as "no change" (complete overlap).

The NAP method has several advantages. The measure is applicable with almost any data type and distribution. It provides a statistical value that can be used for further analysis at the group level. Moreover, NAP values can be derived from receiver operating characteristic analysis and thus are easy to calculate with any statistical software and come with confidence intervals, thus indicating the uncertainty of the effect size estimate. One disadvantage with the NAP method is that the degree of change is not considered; that is, the estimated effect will be the same whether the responses changed 0.1 scale point or 10 scale points.

Paired comparisons. Leijon et al. (2019) proposed a Bayesian analysis approach for paired comparisons assessed on ordinal scales. A software package accompanying the article is available (PairedCompCalc, available at the Python Package Index). Outcome of the analysis is the probability of a hypothesized difference between, for instance, two hearing aid programs, ranging from 0% to 100%. A credibility close to 50% indicates that it is not possible to draw a clear conclusion. The software can provide results for the population mean, for a random (not yet seen) individual in the population, and for individual test participants in the study (not described in the accompanying article).

The advantage of this approach is that it can take ordinal data without assumptions of the distribution of the ordinal ratings, and it does not impose requirements on the number of responses for each pair or the number of test participants. However, few test participants and/or responses per test participant will reduce the statistical certainty of the results.

Multilevel modeling. Multilevel modeling is a popular approach for analyzing EMA data, for several reasons. First, it accounts for the variance related to the repeated-measures nature of EMA data. Second, it allows for both categorical and continuous independent variables. Third, it allows researchers to keep data from all participants, even when there is unequal amount of data between participants. Overall, multilevel modeling gives researchers the flexibility to explain variance due to fixed and random effects more accurately. However, multilevel modeling approaches for EMA data also encounter critique. In particular, Ram et al. (2017) argue that the pooling of heterogeneous person-related data into one single model compromises ecological validity.

Multilevel modeling is conventionally performed as a linear mixed model in audiological EMA studies (e.g., Timmer et al., 2018b). Therefore, the assumptions for parametric regression-based analyses should be met: The dependent variable is a ratio or interval scale, the conditional residuals are normally distributed, and variance heterogeneity is not present or is accounted for. Subjective EMA data are often collected on ordinal scales but treated as interval data. This is a controversial topic particularly in multilevel modeling. Bauer and Sterba (2011) criticized this approach, as did Liddell and Kruschke (2018), who suggested utilizing an ordinal outcome using a cumulative link function in a generalized linear mixed model in a well-constructed Bayesian hierarchical model. This method was used by von Gablenz et al. (2021) for studying hearing aid benefit. Alternatively, Leijon et al. (2023) presented and evaluated a ready to use Python software solution for Bayesian analysis of EMA data (EmaCalc, available at the Python Package Index). For binary outcome variables with longitudinal data, Thomas et al. (1998) suggested using mixed-effects logistic regression. Oleson et al. (2022) stressed the high complexity and interpretational difficulties of alternative modeling approaches and recommended the use of parametric linear mixed models when normality of residuals can be assumed.

Concluding Remarks

Research Gaps

So far, the number of participants is rather low in most EMA hearing studies. Therefore, the representativeness of the samples and the level of generalization are still limited. The use of a common set of questions by many researchers could facilitate meta-analyses and, in general, would improve the comparability of results.

More research is needed to understand and predict the burden of a study design better. While many studies evaluate the participant burden, those measures are often not comparable. Hence it would be useful to develop a validated questionnaire about participant burden in EMA to use across different studies and laboratories to show to which extent different components of an EMA study contribute to participant burden.

More research is also needed to understand the relationship between momentary and retrospective responses. This would include identifying which study design and analysis would allow researchers to examine that relationship to sort out which events (and the user's experience of those events) predict long-term outcomes (such as disability, handicap, help seeking, intervention compliance, and benefits).

EMA has great potential not only as a research tool but also in clinical rehabilitation. There, however, burden is likely a bigger issue than for voluntary test participants. The tool would need to be usable also by participants of low techno-readiness and even people with cognitive impairment. Another open question is how to best display the collected data to a health care provider to support the individualized treatment optimally. More research is needed to identify the best design choices for EMA as intervention. This could range from a positive focus intervention by asking the user to report their positive listening experiences (Lelic et al., 2023) to ecological momentary intervention, where subjective and objective data are used to deliver intervention (e.g., offer of a new hearing program) or advice (e.g., to wear hearing aids) at a time or situation when the participant is most likely to act upon it.

Reporting Requirements

Von Elm et al. (2007) published the Strengthening and Reporting of Observational Studies in Epidemiology (STROBE) checklist, which has been widely adapted as a tool for guiding the reporting of observational research. Liao et al. (2016) adapted the STROBE Checklist for Reporting EMA Studies (CREMAS). This checklist provides a useful starting point for reporting audiological EMA research (Timmer et al., 2018b). The CREMAS is relatively detailed, and it is advisable to consult the original article before writing up any EMA article. Categories span the title to the conclusion, and items recommend being reported include:

- rationale
- training
- technology
- wave (referred to in this article as "phase") duration, that is, the number of waves included in the study and the overall time period for their conduct
- monitoring period, that is, the number of days with EMA data collection, separately for weekdays and weekend days
- prompting design
- prompt frequency
- design features
- attrition
- prompt delivery
- latency, that is, the amount of time elapsed between the prompt signal and the response to the prompt
- compliance rate
- missing data
- limitations

Additional recommendations for audiology EMA research by the authors of this article include:

- time and place of data collection
- position of hearing device architecture
- duration of time for acclimatization to hearing aids
- mean wearing time of the recording device per person
- amount of analyzable objective data
- amount of objective data lost
- reasons for missing objective data
- broadband levels in dBA along with smoothed and down-sampled levels in 1/3 octave bands, or at least the bandwidth and the frequency response of the signal chain
- report of bandwidth, weighting, and averaging approach for metrics derived on fluctuating signals

Conclusions

EMA is a well-established research methodology in social and health sciences that has seen considerable growth in parallel with the spread of smartphone ownership. For audiological research, EMA offers attractive opportunities for addressing previously studied questions in more depth and with more ecological validity than traditional methods of investigation (not least when self-report is coupled with objective data from hearing devices), and for addressing new questions. While these opportunities have spawned rapid growth in EMA studies in audiology, it is not yet a matured methodology, and presents a range of possibly unfamiliar challenges. These challenges cover all aspects of research studies, from protocol design to participant recruitment, the technical platform, self-report survey design, and data analysis.

Based on the collective experience of audiological researchers who have applied EMA, recommendations can be made that may assist others who are considering applying EMA methodology to tackle and resolve the challenges and to best exploit the opportunities offered by EMA. Widening adoption of some basic good practices and reporting standards will help the field to move forward by supporting transparency, reproducibility, and rigor.

Data Availability Statement

For this article, no data have been collected that could be made available.

Acknowledgments

Graham Naylor was supported by the Medical Research Council (Grants MR/S003576/1 and MR/X003620/1). Petra von Gablenz was supported by the Research Fund of Jade University of Applied Sciences (SO-EMA) and by the Lower Saxony Ministry for Science and Culture with funds from the governmental funding initiative zukunft.niedersachsen of the Volkswagen Foundation, project “Data-driven health (DEAL).” The authors would like to thank Jacob Oleson for a guest talk in their workshop series and for feedback on an earlier draft of the Analysis section.

References

Aldaz, G., Puria, S., & Leifer, L. J. (2016). Smartphone-based system for learning and inferring hearing aid settings. *Journal of the American Academy of Audiology, 27*(09), 732–749. https://doi.org/10.3766/jaaa.15099

Andersen, L. S., Andersson, K. E., Wu, M., Pontoppidan, N., Bramsløw, L., & Neher, T. (2019). Assessing daily-life benefit from hearing aid noise management: SSQ12 vs. ecological momentary assessment. In *Proceedings of the International Symposium on Auditory and Audiological Research*.

Andersson, K. E., Andersen, L. S., Christensen, J. H., & Neher, T. (2021). Assessing real-life benefit from hearing-aid noise management: SSQ12 Questionnaire versus ecological momentary assessment with acoustic data-logging. *American Journal of Audiology, 30*(1), 93–104. https://doi.org/10.1044/2020_AJA-20-00042

Andersson, K. E., Neher, T., & Christensen, J. H. (2023). Ecological momentary assessments of real-world speech listening are associated with heart rate and acoustic condition. *Frontiers in Audiology and Otology, 1*, Article 127521. https://doi.org/10.3389/fauot.2023.1275210

Barrett, L. F., & Barrett, D. J. (2001). An introduction to computerized experience sampling in psychology. *Social Science Computer Review, 19*(2), 175–185. https://doi.org/10.1177/089443930101900204

Barta, W. D., Tennen, H., & Litt, M. D. (2012). Measurement reactivity in diary research. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 108–123). The Guilford Press.

Bauer, D. J., & Sterba, S. K. (2011). Fitting multilevel models with ordinal outcomes: Performance of alternative specifications and methods of estimation. *Psychological Methods, 16*(4), 373–390. https://doi.org/10.1037/a0025813

Beal, D. J., & Weiss, H. M. (2003). Methods of ecological momentary assessment in organizational research. *Organizational Research Methods, 6*(4), 440–464. https://doi.org/10.1177/1094428103257361

Bitzer, J., Kissner, S., & Holube, I. (2016). Privacy-aware acoustic assessments of everyday life. *Journal of the Audio Engineering Society, 64*(6), 395–404. https://doi.org/10.17743/jaes.2016.0020

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54*(1), 579–616. https://doi.org/10.1146/annurev.psych.54.101601.145030

Borschke, I., Jürgens, T., & Schinkel-Bielefeld, N. (2021). *Remote EMA study about modification and avoidance of difficult listening situations by hearing aid users* [Paper presentation]. The Fifth International Meeting on Internet & Audiology, Chicago, IL, USA. https://osf.io/k7dxm/

Borschke, I., Jürgens, T., & Schinkel-Bielefeld, N. (2024). How individuals shape their acoustic environment: Implications for hearing aid comparison in ecological momentary assessment. *Ear and Hearing*. Advance online publication. https://doi.org/10.1097/AUD.0000000000001490

Bosman, A. J., Christensen, J. H., Rosenbom, T., Patou, F., Janssen, A., & Hol, M. K. (2021). Investigating real-world benefits of high-frequency gain in bone-anchored users with ecological momentary assessment and real-time data logging. *Journal of Clinical Medicine, 10*(17), Article 3923. https://doi.org/10.3390/jcm10173923

Brungart, D. S., Barrett, M. E., Cohen, J. I., Fodor, C., Yancey, C. M., & Gordon-Salant, S. (2020). Objective assessment of speech intelligibility in crowded public spaces. *Ear and Hearing, 41*(Suppl. 1), 68S–78S. https://doi.org/10.1097/AUD.0000000000000943

Burke, L. A., & Naylor, G. (2020). Daily-life fatigue in mild to moderate hearing impairment: An ecological momentary assessment study. *Ear and Hearing, 41*(6), 1518–1532. https://doi.org/10.1097/AUD.0000000000000888

Burke, L. A., & Naylor, G. (2022). Smartphone app–based noncontact ecological momentary assessment with experienced and naïve older participants: Feasibility study. *JMIR Formative Research, 6*(3), Article e27677. https://doi.org/10.2196/27677

Burke, L. E., Shiffman, S., Music, E., Styn, M. A., Kriska, A., Smailagic, A., Siewiorek, D., Ewing, L. J., Chasens, E., & French, B. (2017). Ecological momentary assessment in behavioral research: Addressing technological and human

participant challenges. *Journal of Medical Internet Research, 19*(3), Article e77. https://doi.org/10.2196/jmir.7138
Christensen, J. A., Sis, J., Kulkarni, A. M., & Chatterjee, M. (2019). Effects of age and hearing loss on the recognition of emotions in speech. *Ear and Hearing, 40*(5), 1069–1083. https://doi.org/10.1097/AUD.0000000000000694
Christensen, T. C., Barrett, L. F., Bliss-Moreau, E., Lebo, K., & Kaschub, C. (2003). A practical guide to experience-sampling procedures. *Journal of Happiness Studies, 4*(1), 53–78. https://doi.org/10.1023/A:1023609306024
Csikszentmihalyi, M., Larson, R., & Prescott, S. (1977). The ecology of adolescent activity and experience. *Journal of Youth and Adolescence, 6*(3), 281–294. https://doi.org/10.1007/BF02138940
Dunn, C. C., Stangl, E., Oleson, J., Smith, M., Chipara, O., & Wu, Y.-H. (2021). The influence of forced social isolation on the auditory ecology and psychosocial functions of listeners with cochlear implants during COVID-19 mitigation efforts. *Ear and Hearing, 42*(1), 20–28. https://doi.org/10.1097/AUD.0000000000000991
Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., & Viechtbauer, W. (2020). The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment, 29*(2), 136–151. https://doi.org/10.1177/1073191120957102
Galvez, G., Turbin, M. B., Thielman, E. J., Istvan, J. A., Andrews, J. A., & Henry, J. A. (2012). Feasibility of ecological momentary assessment of hearing difficulties encountered by hearing aid users. *Ear and Hearing, 33*(4), 497–507. https://doi.org/10.1097/AUD.0b013e3182498c41
Gatehouse, S. (1991). The role of non-auditory factors in measured and self-reported disability. *Acta Oto-Laryngologica, 111*(Suppl. 476), 249–256. https://doi.org/10.3109/00016489109127286
Gatehouse, S., Naylor, G., & Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *International Journal of Audiology, 42*(Suppl. 1), 77–85. https://doi.org/10.3109/14992020309074627
Glista, D., O'Hagan, R., Van Eeckhoutte, M., Lai, Y., & Scollie, S. (2021). The use of ecological momentary assessment to evaluate real-world aided outcomes with children. *International Journal of Audiology, 60*(Suppl. 1), S68–S78. https://doi.org/10.1080/14992027.2021.1881629
Glista, D., Scollie, S., Bagatto, M., Seewald, R., Parsa, V., & Johnson, A. (2009). Evaluation of nonlinear frequency compression: Clinical outcomes. *International Journal of Audiology, 48*(9), 632–644. https://doi.org/10.1080/14992020902971349
Green, A. S., Rafaeli, E., Bolger, N., Shrout, P. E., & Reis, H. T. (2006). Paper or plastic? Data equivalence in paper and electronic diaries. *Psychological Methods, 11*(1), 87–105. https://doi.org/10.1037/1082-989X.11.1.87
Hasan, S. S., Brummet, R., Chipara, O., & Wu, Y.-H. (2017). Assessing the performance of hearing aids using surveys and audio data collected in situ. In *2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. https://doi.org/10.1109/INFCOMW.2017.8116390
Hasan, S. S., Chipara, O., Wu, Y.-H., & Aksan, N. (2014). Evaluating auditory contexts and their impacts on hearing aid outcomes with mobile phones. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*. https://doi.org/10.4108/icst.pervasivehealth.2014.254952
Henry, J. A., Galvez, G., Turbin, M. B., Thielman, E. J., McMillan, G. P., & Istvan, J. A. (2012). Pilot study to evaluate ecological momentary assessment of tinnitus. *Ear and Hearing, 33*(2), 179–290. https://doi.org/10.1097/AUD.0b013e31822f6740
Heron, K. E., Everhart, R. S., McHale, S. M., & Smyth, J. M. (2017). Using mobile-technology-based ecological momentary assessment (EMA) methods with youth: A systematic review and recommendations. *Journal of Pediatric Psychology, 42*(10), 1087–1107. https://doi.org/10.1093/jpepsy/jsx078
Holube, I., von Gablenz, P., & Bitzer, J. (2020). Ecological momentary assessment in hearing research: Current state, challenges, and future directions. *Ear and Hearing, 41*(Suppl. 1), 79S–90S. https://doi.org/10.1097/AUD.0000000000000934
Jacobs, P. G., & Kaye J. A. (2015). Ubiquitous real-world sensing and audiology-based health informatics. *Journal of the American Academy of Audiology, 26*(9), 777–783. https://doi.org/10.3766/jaaa.15010
Jensen, N. S., Hau, O., Lelic, D., Herrlin, P., Wolters, F., & Smeds, K. (2019, September 9-13). Evaluation of auditory reality and hearing aids using an ecological momentary assessment (EMA) approach. In *Proceedings of the 23rd International Congress on Acoustics (ICA), Aachen, Germany*. https://publications.rwth-aachen.de/record/769547/files/769547.pdf [PDF]
Jensen, N. S., & Nielsen, C. (2005). Auditory ecology in a group of experienced hearing-aid users: Can knowledge about hearing-aid users' auditory ecology improve their rehabilitation? In A. N. Rasmussen, T. Poulsen, T. Andersen, & C. B. Larsen (Eds.), *21st Danavox Symposium* (pp. 235–258). Danavox Jubilee Foundation. http://www.audiological-library.gnresound.dk/
Jenstad, L. M., Singh, G., Boretzki, M., DeLongis, A., Fichtl, E., Ho, R., Huen, M., Meyer, V., Pang, F., & Stephenson, E. (2021). Ecological momentary assessment: A field evaluation of subjective ratings of speech in noise. *Ear and Hearing, 42*(6), 1770–1781. https://doi.org/10.1097/AUD.0000000000001071
Jorgensen, E. J., Stangl, E., Chipara, O., Hernandez, H., Oleson, J., & Wu, Y.-H. (2021). GPS predicts stability of listening environment characteristics in one location over time among older hearing aid users. *International Journal of Audiology, 60*(5), 328–340. https://doi.org/10.1080/14992027.2020.1831083
Jorgensen, E. J., Xu, J., Chipara, O., Oleson, J., Galster, J., & Wu, Y.-H. (2023). Auditory environments and hearing aid feature activation among younger and older listeners in an urban and rural area. *Ear and Hearing, 44*(3), 603–618. https://doi.org/10.1097/aud.0000000000001308
Kowalk, U., Franz, S., Groenewold, H., Holube, I., von Gablenz, P., & Bitzer, J. (2020). olMEGA: An open source android solution for ecological momentary assessment. *GMS Zeitschrift für Audiologie – Audiological Acoustics, 2*(Doc08), 1–9.
Leijon, A., Dahlquist, M., & Smeds, K. (2019). Bayesian analysis of paired-comparison sound quality ratings. *The Journal of the Acoustical Society of America, 146*(5), 3174–3183. https://doi.org/10.1121/1.5131024
Leijon, A., von Gablenz, P., Holube, I., Taghia, J., & Smeds, K. (2023). Bayesian analysis of ecological momentary assessment (EMA) data collected in adults before and after hearing rehabilitation. *Frontiers in Digital Health, 5,* Article 16. https://doi.org/10.3389/fdgth.2023.1100705
Lelic, D., Nielsen, J., Parker, D., & Marchman Rønne, F. (2022). Critical hearing experiences manifest differently across individuals: Insights from hearing aid data captured in real-life moments. *International Journal of Audiology, 61*(5), 428–436. https://doi.org/10.1080/14992027.2021.1933621
Lelic, D., Parker, D., Herrlin, P., Wolters, F., & Smeds, K. (2023). Focusing on positive listening experiences improves hearing aid outcomes in experienced hearing aid users. *International Journal of Audiology*. Advance online publication. https://doi.org/10.1080/14992027.2023.2190006

Lelic, D., Wolters, F., & Schinkel-Bielefeld, N. (2024). Measuring hearing aid satisfaction in everyday listening situations: Retrospective and in-situ assessments complement each other. *Journal of the American Academy of Audiology*. Advance online publication. https://doi.org/10.1055/a-2265-9418

Li, L., Vikani, A. R., Harris, G. C., & Lin, F. R. (2014). Feasibility study to quantify the auditory and social environment of older adults using a digital language processor. *Otology and Neurotology, 35*(8), 1301–1305. https://doi.org/10.1097/MAO.0000000000000489

Liao, Y., Skelton, K., Dunton, G., & Bruening, M. (2016). A systematic review of methods and procedures used in ecological momentary assessments of diet and physical activity research in youth: An adapted STROBE checklist for reporting EMA studies (CREMAS). *Journal of Medical Internet Research, 18*(6), Article e151. https://doi.org/10.2196/jmir.4954

Liddell, T. M., & Kruschke, J. K. (2018). Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology, 79*(Nov 2018), 328–348. https://doi.org/10.1016/j.jesp.2018.08.009

LoBuono, D. L., Leedahl, S. N., & Maiocco, E. (2020). Teaching technology to older adults: Modalities used by student mentors and reasons for continued program participation. *Journal of Gerontological Nursing, 46*(1), 14–20. https://doi.org/10.3928/00989134-20191118-02

Maes, I., Mertens, L., Poppe, L., Crombez, G., Vetrovsky, T., & Van Dyck, D. (2022). The variability of emotions, physical complaints, intention, and self-efficacy: An ecological momentary assessment study in older adults. *PeerJ, 10,* Article e13234. https://doi.org/10.7717/peerj.13234

Maniaci, M. R., & Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality, 48,* 61–83. https://doi.org/10.1016/j.jrp.2013.09.008

Manini, T. M., Mendoza, T., Battula, M., Davoudi, A., Kheirkhahan, M., Young, M. E., Weber, E., Fillingim, R. B., & Rashidi, P. (2019). Perception of older adults toward smartwatch technology for assessing pain and related patient-reported outcomes: Pilot study. *JMIR mHealth and uHealth, 7*(3), Article e10044. https://doi.org/10.2196/10044

Manolov, R., Losada, J. L., Chacón-Moscoso, S., & Sanduvete-Chaves, S. (2016). Analyzing two-phase single-case data with non-overlap and mean difference indices: Illustration, software tools, and alternatives. *Frontiers in Psychology, 7,* Article 32. https://doi.org/10.3389/fpsyg.2016.00032

Mansour, N., Westermann, A., Marschall, M., May, T., Dau, T., & Buchholz, J. (2021). Guided ecological momentary assessment in real and virtual sound environments. *The Journal of the Acoustical Society of America, 150*(4), 2695–2704. https://doi.org/10.1121/10.0006568

Marks, L. E., & Gescheider, G. A. (2002). Psychophysical scaling. In *Stevens' handbook of experimental psychology* (3rd ed.). Wiley. https://doi.org/10.1002/0471214426.pas0403

Mehl, M. R. (2017). The electronically activated recorder (EAR): A method for the naturalistic observation of daily social behavior. *Current Directions in Psychological Science, 26*(2), 184–190. https://doi.org/10.1177/0963721416680611

Mehl, M. R., & Conner, T. S. (2011). *Handbook of research methods for studying daily life*. The Guilford Press.

Oleson, J. J., Jones, M. A., Jorgensen, E. J., & Wu, Y. H. (2022). Statistical considerations for analyzing ecological momentary assessment data. *Journal of Speech, Language, and Hearing Research, 65*(1), 344–360. https://doi.org/10.1044/2021_JSLHR-21-00081

Parker, R. I., & Vannest, K. J. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*(4), 357–367. https://doi.org/10.1016/j.beth.2008.10.006

Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification, 35*(4), 303–322. https://doi.org/10.1177/0145445511399147

Pohlhausen, J., Holube, I., & Bitzer, J. (2022). Near-ear sound pressure level distribution in everyday life considering the user's own voice and privacy. *Acta Acustica, 6,* Article 40. https://doi.org/10.1051/aacus/2022035

Ponnada, A., Thapa-Chhetry, B., Manjourides, J., & Intille, S. (2021). Measuring criterion validity of microinteraction ecological momentary assessment (micro-ema): Exploratory pilot study with physical activity measurement. *JMIR mHealth and uHealth, 9*(3), Article e23391. https://doi.org/10.2196/23391

Pryss, R., Schlee, W., Reichert, M., Kurthen, I., Giroud, N., Jagoda, L., Neuschwander, P., Meyer, M., Neff, P., & Schobel, J. (2019). Ecological momentary assessment based differences between Android and iOS users of the TrackYourHearing mHealth crowdsensing platform. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. https://doi.org/10.1109/EMBC.2019.8857854

Ram, N., Brinberg, M., Pincus, A. L., & Conroy, D. E. (2017). The questionable ecological validity of ecological momentary assessment: Considerations for design and analysis. *Research in Human Development, 14*(3), 253–270. https://doi.org/10.1080/15427609.2017.1340052

Ryherd, S., Kleiner, M., Waye, K. P., & Ryherd, E. E. (2012). Influence of a wearer's voice on noise dosimeter measurements. *The Journal of the Acoustical Society of America, 131*(2), 1183–1193. https://doi.org/10.1121/1.3675941

Schinkel-Bielefeld, N. (2020). Laboratory experiments versus ecological momentary assessment? The quest to evaluate real life hearing aid performance. In *Forum Acusticum* (pp. 91–98). https://hal.archives-ouvertes.fr/hal-03234226

Schinkel-Bielefeld, N., Gotholt Madsen, N., & Lelic, D. (2022). Comparing hearing aid programs using ecological momentary assessment: Direct versus indirect comparison. In *19th International Symposium on Hearing: Psychoacoustics, Physiology of Hearing, and Auditory Modelling, from the Ear to the Brain (ISH2022)*. https://doi.org/10.5281/zenodo.6576911

Schinkel-Bielefeld, N., Kunz, P., Zutz, A., & Buder, B. (2020). Evaluation of hearing aids in everyday life using ecological momentary assessment: What situations are we missing? *American Journal of Audiology, 29*(3S), 591–609. https://doi.org/10.1044/2020_AJA-19-00075

Schinkel-Bielefeld, N., Lee, Z. X., Cho Cho Thein, K., Martin, J., & Stasiak, S. (2022). On the feasibility of assessing social situations using ecological momentary assessment–A comparison between Singapore and German test participants. *Vibrations in Physical Systems, 33*(1), Article 2022109. https://doi.org/10.21008/j.0860-6897.2022.1.09

Schinkel-Bielefeld, N., Ritslev, J., & Lelic, D. (2023). Reasons for ceiling ratings in real-life evaluations of hearing aids: The relationship between SNR and hearing aid ratings. *Frontiers in Digital Health, 5*. https://doi.org/10.3389/fdgth.2023.1134490

Schlee, W., Pryss, R. C., Probst, T., Schobel, J., Bachmeier, A., Reichert, M., & Langguth, B. (2016). Measuring the moment-to-moment variability of tinnitus: The TrackYourTinnitus smart phone app. *Frontiers in Aging Neuroscience, 8,* Article 294. https://doi.org/10.3389/fnagi.2016.00294

Scollon, C. N., Prieto, C.-K., & Diener, E. (2009). Experience sampling: promises and pitfalls, strength and weaknesses. In

E. Diener (Ed.), *Assessing well-being: The collected works of Ed Diener* (pp. 157–180). Springer. https://doi.org/10.1007/978-90-481-2354-4_8

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Smeds, K., Gotowiec, S., Wolters, F., Herrlin, P., Larsson, J., & Dahlquist, M. (2020). Selecting scenarios for hearing-related laboratory testing. *Ear and Hearing, 41*(Suppl. 1), 20S–30S. https://doi.org/10.1097/AUD.0000000000000930

Smeds, K., Larsson, J., Dahlquist, M., Wolters, F., & Herrlin, P. (2021). Live evaluation of auditory preference, a laboratory test for evaluating auditory preference. *Journal of the American Academy of Audiology, 32*(08), 487–500. https://doi.org/10.1055/s-0041-1735213

Smeds, K., Wolters, F., & Rung, M. (2015). Estimation of signal-to-noise ratios in realistic sound scenarios. *Journal of the American Academy of Audiology, 26*(2), 183–196. https://doi.org/10.3766/jaaa.26.2.7

Steffens, J., Steele, D., & Guastavino, C. (2017). Situational and person-related factors influencing momentary and retrospective soundscape evaluations in day-to-day life. *The Journal of the Acoustical Society of America, 141*(3), 1414–1425. https://doi.org/10.1121/1.4976627

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine, 16*(3), 199–202. https://doi.org/10.1093/abm/16.3.199

Stone, A. A., & Shiffman, S. (2002). Capturing momentary, self-report data: A proposal for reporting guidelines. *Annals of Behavioral Medicine, 24*(3), 236–243. https://doi.org/10.1207/S15324796ABM2403_09

Svensson, E. (2001). Guidelines to statistical evaluation of data from rating scales and questionnaires. *Journal of Rehabilitation Medicine, 33*(1), 47–48. https://doi.org/10.1080/165019701300006542

Thomas, R., Have, T., Kunselman, A. R., Pulkstenis, E. P., & Landis, J. R. (1998). Mixed effects logistic regression models for longitudinal binary response data with informative drop-out. *Biometrics, 54*(1), 367–383. https://doi.org/10.2307/2534023

Timmer, B. H. B., Hickson, L., & Launer, S. (2017). Ecological momentary assessment: Feasibility, construct validity, and future applications. *American Journal of Audiology, 26*(3S), 436–442. https://doi.org/10.1044/2017_AJA-16-0126

Timmer, B. H. B., Hickson, L., & Launer, S. (2018a). Do hearing aids address real-world hearing difficulties for adults with mild hearing impairment? Results from a pilot study using ecological momentary assessment. *Trends in Hearing, 22*. https://doi.org/10.1177/2331216518783608

Timmer, B. H. B., Hickson, L., & Launer, S. (2018b). The use of ecological momentary assessment in hearing research and future clinical applications. *Hearing Research, 369*, 24–28. https://doi.org/10.1016/j.heares.2018.06.012

Trull, T. J., & Ebner-Priemer, U. (2013). Ambulatory assessment. *Annual Review of Clinical Psychology, 9*(1), 151–176. https://doi.org/10.1146/annurev-clinpsy-050212-185510

Vercammen, C., Bott, A., & Saunders, G. H. (2021). Hearing health in the broader context of healthy living and well-being: Changing the narrative. *International Journal of Audiology, 60*(Suppl. 2), 1–3. https://doi.org/10.1080/14992027.2021.1905893

Von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., Vandenbroucke, J. P., & Initiative, S. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Annals of Internal Medicine, 147*(8), 573–577. https://doi.org/10.7326/0003-4819-147-8-200710160-00010

von Gablenz, P., Kowalk, U., Bitzer, J., Meis, M., & Holube, I. (2021). Individual hearing aid benefit in real life evaluated using ecological momentary assessment. *Trends in Hearing, 25*. https://doi.org/10.1177/2331216521990288

Welling, J., Fischer, R.-L., & Schinkel-Bielefeld, N. (2021). Is it possible to identify careless responses with post-hoc analysis in EMA studies? In *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '21)*.

Welling, J., Krug, A.-E., Fischer, R.-L., & Schinkel-Bielefeld, N. (2020). What helps or hinders hearing aid wearers to change hearing aid programs? In *Jahrestagung der Deutschen Gesellschaft für Audiologie*. https://dga.cloud/s/ndk2idMLbZifWJJ

Wolters, F., Smeds, K., Schmidt, E., Christensen, E. K., & Norup, C. (2016). Common sound scenarios: A context-driven categorization of everyday sound environments for application in hearing-device research. *Journal of the American Academy of Audiology, 27*(7), 527–540. https://doi.org/10.3766/jaaa.15105

Wu, Y.-H., Stangl, E., Chipara, O., Gudjonsdottir, A., Oleson, J., & Bentler, R. (2020). Comparison of in-situ and retrospective self-reports on assessing hearing aid outcomes. *Journal of the American Academy of Audiology, 31*(10), 746–762. https://doi.org/10.1055/s-0040-1719133

Wu, Y.-H., Xu, J., Stangl, E., Pentony, S., Vyas, D., Chipara, O., Gudjonsdottir, A., Oleson, J., & Galster, J. (2021). Why ecological momentary assessment surveys go incomplete: When it happens and how it impacts data. *Journal of the American Academy of Audiology, 32*(01), 016–026. https://doi.org/10.1055/s-0040-1719135

Wu, Y.-H., & Bentler, R. A. (2012). Do older adults have social lifestyles that place fewer demands on hearing? *Journal of the American Academy of Audiology, 23*(9), 697–711. https://doi.org/10.3766/jaaa.23.9.4

Wu, Y.-H., Stangl, E., Chipara, O., Hasan, S. S., DeVries, S., & Oleson, J. (2019). Efficacy and effectiveness of advanced hearing aid directional and noise reduction technologies for older adults with mild to moderate hearing loss. *Ear and Hearing, 40*(4), 805–822. https://doi.org/10.1097/AUD.0000000000000672

Wu, Y.-H., Stangl, E., Chipara, O., Hasan, S. S., Welhaven, A., & Oleson, J. (2018). Characteristics of real-world signal to noise ratios and speech listening situations of older adults with mild to moderate hearing loss. *Ear and Hearing, 39*(2), 293–304. https://doi.org/10.1097/AUD.0000000000000486

Wu, Y.-H., Stangl, E., Zhang, X., & Bentler, R. A. (2015). Construct validity of the ecological momentary assessment in audiology research. *Journal of the American Academy of Audiology, 26*(10), 872–884. https://doi.org/10.3766/jaaa.15034

Appendix (p. 1 of 8)

Survey Item Examples

The table below gives an overview of many types of ecological momentary assessment (EMA) questionnaire items that can be relevant in audiological research. For each item (first column), a description is given (second column), references to research that used the item (third column) are given, and an example is included (fourth column). The description in the second column is included when appropriate. The bold citation in the third column signifies the source of the example. The examples are meant as inspiration, but no recommendations on exact wording are given. Therefore, the response options and the scales used vary from item to item. However, when designing an EMA study, it is recommended to be consistent with the response options and scales. References are sorted by year and then alphabetically. The example reference is marked in bold. An asterisk (*) denotes that the example is translated to English and the original language is given in parentheses. Since these translations were usually performed only for publication purposes and were not actually used in the studies, gold standards for translations have usually not been followed. The original response options may work better than the translated response options.

The table is divided into three sections. First, items that are useful in most audiological research contexts are presented. Second, specific items related to hearing device research are given. Lastly, items related to other research questions are given. Despite the long list of EMA items, the list is not complete, and new research topics will require new EMA items.

Questionnaire items applicable in most research contexts in audiology			
Item	**Description**	**References**	**Example**
Location	***Reasons to include*** The location can indirectly say something about familiarity, typical activities, and physical features of the place (reverberation, wind noise, etc.). ***Things to consider***	Henry et al. (2012), Timmer et al. (2018a), Jensen et al. (2019), **Burke & Naylor (2020)**,[a] Smeds et al. (2020), Dunn et al. (2021), von Gablenz et al. (2021)	Where are you right now? • In my home • Restaurant/bar/cafe • Outdoors • Shops • Work • In transit • Other
Listening activity	***Reasons to include*** The listening activity (task) will likely affect the experience and the benefit of hearing solutions. By asking the participant to state one main listening activity, a complex situation will be easier to understand for the researcher and the data will be easier to interpret. ***Things to consider*** If the research is focused only on situations with speech, the response alternatives can be tailored to those situations. However, it is a *recommendation* to use the CoSS categorization (Wolters et al. 2016), as illustrated in the example in the fourth column, or response options that can, afterward, be grouped into the CoSS categories (von Gablenz et al. 2021).	Galvez et al. (2012), Timmer et al. (2017, 2018a), Jensen et al. (2019), Burke & Naylor (2020), Jorgensen et al. (2021), Schinkel-Bielefeld et al. (2020), **Smeds et al. (2020)**,[b] Bosman et al. (2021), Jenstad et al. (2021), von Gablenz (2021), Lelic et al. (2023)	Type of situation (*Swedish) • Conversation with one person • Conversation with more than one person • Conversation through telephone or other communication devices • Focused listening without ability to control the sound source (e.g., lecture, concert) • Focused listening with ability to control the sound source (e.g., TV, radio, or other media) • Situation without conversation or focused listening

(*table continues*)

Appendix (p. 2 of 8)

Survey Item Examples

Questionnaire items applicable in most research contexts in audiology			
Item	**Description**	**References**	**Example**
Background sounds	***Reasons to include*** The presence of background noise will likely affect the experience and the benefit of hearing solutions. Some studies ask for specific background sound sources, others if background sounds are present. ***Things to consider***	Burke & Naylor (2020), **Schinkel-Bielefeld et al. (2020)**,[b] Smeds et al. (2020), von Gablenz et al. (2021)	What sounds are audible in the background? (*German) • Voices/other people • Traffic noise • Household noise • Music/television • Engines/machinery/ ventilation • Wind • Silence • Other/further detail If "Other/further details" is selected: Please describe which sounds are audible in the background (free text).
Importance of hearing well	***Reasons to include*** Asking for importance (of hearing well) gives the opportunity to focus the attention on the important situations in the analysis. In addition, this item could potentially be used as a proxy for motivation. ***Things to consider*** In some studies, researchers have asked about the general importance of the situation, which is different from asking about the importance of hearing well.	Galvez et al. (2012), Timmer et al. (2017), **Smeds et al. (2020)**,[b] von Gablenz et al. (2021), Lelic et al. (2024)	How important is it for you to hear well in this situation? (*Swedish) • Little or not at all important • Moderately important • Very important
Hearing difficulty	***Reasons to include*** Asking for difficulty gives the opportunity to focus the attention on the easy or difficult situations in the analysis. ***Things to consider*** The concept could be related to listening effort (see below) but is easier to understand and does not change if the participant gives up in a very difficult situation.	Galvez et al. (2012), Andersson et al. (2021), Mansour et al. (2021), Smeds et al. (2020), **Lelic et al. (2023)**[b]	How difficult is it for you to hear/listen in this situation? (*Danish) • Not difficult • Slightly difficult • Moderately difficult • Very difficult • Extremely difficult
Questionnaire items for studies using hearing devices			
Item	**Description**	**References**	**Example**
Hearing device use	***Reasons to include*** In hearing device research, it might be important to know if a self-report reflects aided or unaided listening. Also, if the EMA questionnaire is adaptive, hearing device–related questions can be skipped if the participant is not wearing hearing devices when responding. ***Things to consider*** Objective hearing device data can replace or supplement the self-report. Information about hearing device use can also be included in other items.	Galvez et al. (2012), **Timmer et al. (2018a)**,[c] Wu et al. (2019)	Are you currently wearing hearing aids? • Yes • No

(*table continues*)

Appendix (p. 3 of 8)

Survey Item Examples

Questionnaire items for studies using hearing devices			
Item	**Description**	**References**	**Example**
Satisfaction with hearing devices	***Reasons to include*** ***Things to consider*** Test participants might interpret satisfaction differently since it could be related to, for instance, sound quality, aesthetics, physical fit or the benefit they provide.	Galvez et al. (2012), Dunn et al. (2021), **Jenstad et al. (2021)**,[d] Schinkel-Bielefeld et al. (2020, 2023), Lelic et al. (2023)	In this situation, how satisfied are you with the hearing aids? • 5: Very satisfied • 4 • 3 • 2 • 1: Very dissatisfied (Only anchor words for extreme values used.)
Benefit of hearing devices	***Reasons to include*** ***Things to consider*** Hearing device benefit might be a difficult concept to evaluate using EMA. Test participants could be asked to do the following: • compare current performance with how they think they would do in the situation without hearing devices; • compare current performance with how they think their own hearing devices would work in the situation; and • make direct paired comparisons of benefit for two hearing-aid programs (see below).	Galvez et al. (2012), Wu et al. (2020), **Jenstad et al. (2021)**[d]	In this situation, how much are your hearing aid helping you (compared to no hearing aids)? • 5: Helping very much • 4 • 3 • 2 • 1: Not helping at all (Only anchor words for extreme values used.)
Residual activity limitation	***Reasons to include*** Even when hearing devices are helping, there may be difficulties remaining. ***Things to consider***	Wu et al. (2020), **Jenstad et al. (2021)**[d]	In this situation, how much difficulty do you still have? • 5: No difficulty • 4 • 3 • 2 • 1: Very much difficulty (Only anchor words for extreme values used.)
Quality of processed sound	***Reasons to include*** ***Things to consider*** Sound quality could be a difficult concept for test participants to understand. Sometimes, dimensions of sound quality with descriptive adjectives could be easier to use. Examples: loudness (very loud), clarity (very clear), brightness (muffled).	Schinkel-Bielefeld et al. (2020), Andersen et al. (2019), Bosman et al. (2021), **Schinkel-Bielefeld, Gotholt Madsen & Lelic (2022)**[c]	How is the sound quality in your hearing aids right now? (*Danish) • Very good • Rather good • Neither good nor bad • Rather bad • Very bad

(*table continues*)

Appendix (p. 4 of 8)

Survey Item Examples

Questionnaire items for studies using hearing devices			
Item	**Description**	**References**	**Example**
Program preference	***Reasons to include*** When preference for two or more hearing device programs in specific situations is of interest. ***Things to consider*** Program preference can be investigated either using direct paired comparisons (A/B) in each experienced situation or indirectly, for instance, by using alternating programs each day and rate some attribute that, afterward, can be compared across days. Alternatively, a crossover design with one program per study phase can be used. If the EMA app allows automatic program switches, the direct paired comparisons can be done in a randomized and blind way. For the indirect comparisons, it is beneficial if the app can automatically change the program each day.	Jensen et al. (2019), **Smeds et al. (2021)**,[e] Schinkel-Bielefeld et al. (2022)	Which hearing aid program do you prefer? (*Swedish) • Program 1 • Program 2 • Hear a difference but have no preference • Do not hear a difference
Hearing aid program performance	***Reasons to include*** ***Things to consider***	Welling et al. (2020), **Andersson et al. (2021)**,[a] Bosman et al. (2021)	How would you rate the current hearing aid program performance in this particular situation? (*Danish) Slider: *Poor* (0) to *High* (10)
Additional questionnaire items to consider depending on research context			
Item	**Description**	**References**	**Example**
Time since event	***Reasons to include*** Asking informants to indicate the time that has elapsed between an event and the report allows the analysis to be limited to really momentary reports on a case-by-case basis. This item is particularly important if objective data are collected since it is crucial to match objective and subjective data. ***Things to consider*** If the item is used for matching of subjective and objective data, the item should be mandatory and placed at the beginning of the survey so that the time estimate is not biased by the time it takes to fill out the survey.	Timmer et al. (2018a), Schinkel-Bielefeld et al. (2020), **von Gablenz et al. (2021)**[a]	How many minutes have elapsed since the event? (*German) • Now • < 2–3 • < 5 • < 10 • < 15 • < 20 • < 30
Frequency of occurrence	***Reasons to include*** In theory, random time-based prompts should reveal the occurrence of various situations. However, especially for rare or short events, this is only true if the trial is very long. Also, there are situations where it is inconvenient to respond to a survey, and these situations might be underrepresented. Therefore, this item could be instructive. ***Things to consider*** It can be easier to answer if the response alternatives are concrete (as in column 4) rather than if they are more qualitative (often, rarely, etc.).	**Smeds et al. (2020)**,[b] Galvez et al. (2012), Lelic et al. (2023)	How often are you in this situation? (*Swedish) • Daily • Weekly • Monthly or less

(*table continues*)

Appendix (p. 5 of 8)

Survey Item Examples

Additional questionnaire items to consider depending on research context			
Item	**Description**	**References**	**Example**
Activity (not hearing related)	***Reasons to include*** Activities that are not directly connected to hearing can be important and might indicate the intent of the participant. ***Things to consider*** If the EMA survey also contains a question on the location, answer options to the activity can be adapted to the location.	Henry et al. (2012), **Burke & Naylor (2020)**,[a] von Gablenz et al. (2021)	What are you doing right now? • Watching/listening to entertainment • Conversation • Using the internet • Reading • Eating • Household task • Other *Here, only answer options for the home environment are shown.*
Activity limitation	***Reasons to include*** The consequences of hearing problems can be a limitation of activities or their enjoyment. ***Things to consider***	Timmer et al. (2018a), **Dunn et al. (2021)**,[d] von Gablenz et al. (2021)	My hearing difficulties limited what I wanted to do or say • Strongly agree • Agree • Neutral • Disagree • Strongly disagree
Location of sound source	***Reasons to include*** Sound from nearby locations (or streamed sound) and sound sources in the field of view may be easier to understand than more distant or out of view sources. Also, if hearing aids are worn, signal processing might differ depending on the direction of the sound. Some studies also ask for the position of distractors. ***Things to consider***	Timmer et al. (2018a), Jorgensen et al. (2021), **Schinkel-Bielefeld et al. (2020)**[b]	Where is the person/sound source you are listening to located? (*German) • Constantly directly in front of me • In my field of view, but not directly in the front • Behind me or to the side (out of view) • At varying positions (e.g., alternating talkers, in motion yourself) • Other
Loudness/ noisiness	***Reasons to include*** ***Things to consider*** While loudness and noisiness are different constructs, in some languages, the word for noise is ambiguous. Some studies ask for the overall loudness, the loudness of the background, and/or the loudness of the talker.	Galvez et al. (2012), Timmer et al. (2017), Wu et al. (2019), Burke & Naylor (2020), Jorgensen et al. (2021), Smeds et al. (2020), Bosman et al. (2021), **Dunn et al. (2021)**,[d] Glista et al. (2021), Jenstad et al. (2021), Mansour et al. (2021), von Gablenz et al. (2021)	Overall, how loud were the background/environmental sounds? • Very loud • Loud • Medium • Soft • Very soft
Speaker characteristics	***Reasons to include*** When analyzing speech situations, speaker characteristics may be of importance. These include familiarity of speakers (including voices from TV and radio) or features such as gender and age. ***Things to consider***	Timmer et al. (2017; 2018a), Jorgensen et al. (2021), Dunn et al. (2021), **von Gablenz et al. (2021)**[c]	Are you familiar with the voices? (*German) • Familiar, • Unfamiliar • Both familiar and unfamiliar

(*table continues*)

Survey Item Examples

Additional questionnaire items to consider depending on research context			
Item	**Description**	**References**	**Example**
Visual cues	***Reasons to include*** Lip reading can facilitate communication. Nonverbal information provided by the talker could give additional cues. ***Things to consider***	**Timmer et al. (2018a)**,[c] Jorgensen et al. (2021), Dunn et al. (2021)	Could you see the talker's face? • No • Yes, but only sometimes • Almost always
Sound localization	***Reasons to include*** The ability to localize sound sources enhance communication and is important for understanding of the environment. ***Things to consider***	Wu et al. (2019), Schinkel-Bielefeld et al. (2020), **von Gablenz et al. (2021)**[c]	How well do you hear where individual sounds come from? (*German) • Perfectly • Very well • Rather well • Okay • Rather poorly • Very poorly • Not at all
Perceived level of speech of interest relative to background sounds	***Reasons to include*** SNR is central to perception. The level of sound of interest relative to background sounds can be asked subjectively. ***Things to consider*** SNR is hard to estimate objectively, as one can never be certain what the participant is paying attention to.	**Dunn et al. (2021)**[d]	The speech of interest was ______ when compared to all other sounds. • Much louder • Somewhat louder • Equally loud • Somewhat softer • Much softer
Speech understanding	***Reasons to include*** ***Things to consider*** Response options for this question vary a lot in character. Some are quantitative (for instance, percentage of words understood); others use subjective descriptions (like the example given here).	Timmer et al. (2017), Timmer et al. (2018a), Wu et al. (2019), Schinkel-Bielefeld et al. (2020), Andersen et al. (2019), Bosman et al. (2021), Dunn et al. (2021), Glista et al. (2021), Jenstad et al. (2021), Mansour et al. (2021), **von Gablenz et al. (2021)**[c]	How well or poorly do you understand? (*German) • Perfectly • Very well • Rather well • Fair • Rather poorly • Very poorly • Not at all

(*table continues*)

Appendix (p. 7 of 8)

Survey Item Examples

Additional questionnaire items to consider depending on research context			
Item	**Description**	**References**	**Example**
Listening effort	***Reasons to include*** In situations where speech understanding is good, listening effort might show more differentiation between test conditions. ***Things to consider*** Widely used, but with mixed experiences. If participants stop exerting effort and do not try to understand anymore, effort may be low despite the difficulty of the situation. May be misinterpreted as a trick question. Asking about (hearing) difficulty is an alternative (see above).	**Timmer et al. (2017, 2018a)**,[c] Wu et al. (2019), Schinkel-Bielefeld et al. (2020), Andersson et al. (2021), Dunn et al. (2021), Jenstad et al. (2021), Mansour et al. (2021), von Gablenz et al. (2021)	On average, how much effort did you have to put in to listen effectively? • No effort • Some effort • Moderate effort • Quite a bit of effort • A lot of effort
Fatigue	***Reasons to include*** ***Things to consider*** Hearing problems and compensatory strategies can increase fatigue. This is a consequence of hearing problems that is not necessarily attributed to hearing problems by the participants. Therefore, it may be advantageous to measure general fatigue rather than listening fatigue. However, the sensitivity of a generic fatigue item to measure listening fatigue must be considered.	**Burke & Naylor (2020)**,[a] Schinkel-Bielefeld et al. (2020)	Please rate your fatigue (weariness, tiredness) by selecting the one number that best describes your fatigue right now. Slider: *No fatigue* = 0 to *As bad as you can imagine* = 10
Annoyance	***Reasons to include*** ***Things to consider*** Some studies ask for the annoyance of hearing problems; others ask specifically for the annoyance of (background) sounds.	Galvez et al. (2012), Smeds et al. (2020), **Mansour et al. (2021)**[a]	Are you annoyed with certain sounds in this environment? • Not at all annoyed • Not that annoyed • Somewhat annoyed • Very annoyed • Extremely annoyed
Pleasantness	***Reasons to include*** ***Things to consider*** Several studies ask for sound pleasantness or the sound quality. While this is often associated with technical reproduction of sound and respective artifacts, hearing impairment can also lead to speech sounding more muffled or sound being perceived as too soft.	Andersson et al. (2021), Bosman et al. (2021), **Mansour et al. (2021)**,[a] von Gablenz et al. (2021)	Does this environment sound pleasant to you? • Not at all pleasant • Not that pleasant • Somewhat pleasant • Very pleasant • Extremely pleasant
Mood	***Reasons to include*** ***Things to consider*** Mood can be influenced by hearing problems. Conversely, mood may influence the hearing performance or the willingness to answer an EMA questionnaire. There are multiple mood states of possible interest (e.g., stress and anxiety) and different dimensions of mood that can be assessed (e.g., valence and intensity).	Schinkel-Bielefeld et al. (2020), **von Gablenz et al. (2021)**,[c] Schinkel-Bielefeld et al. (2022)	How do you feel right now? (*German) 😀 🙂 😐 😕 ☹️

(*table continues*)

Appendix (p. 8 of 8)

Survey Item Examples

Additional questionnaire items to consider depending on research context			
Item	**Description**	**References**	**Example**
Feelings related to hearing difficulties	***Reasons to include*** Hearing loss can lead to negative feelings and an impact on mental health. Various psychosocial consequences of hearing problems can be assessed. Dunn et al. (2021) investigated sadness, loneliness, and nervousness, but other feelings can be investigated. ***Things to consider*** Sometimes feelings arise as a consequence of the hearing loss without the person attributing this to hearing loss. Hence, it may also make sense to ask for feelings in general and not only in conjunction with hearing loss.	**Dunn et al. (2021)**[d]	My hearing difficulties made me feel . . . (Sad or depressed/lonely or isolated/nervous or anxious) • Strongly agree • Agree • Neutral • Disagree • Strongly disagree
Impact on others	***Reasons to include*** When hearing difficulties are reflected in reactions of conversation partners, the situation could be particularly stressful. ***Things to consider***	Galvez et al. (2012), **von Gablenz et al. (2021)**[c]	Do you think your conversation partner finds the conversation difficult? (*German) • Extremely difficult • Very difficult • Considerably difficult • Moderately difficult • Slightly difficult • Very slightly difficult • Not difficult at all

Note. CoSS = Common Sound Scenarios; EMA = ecological momentary assessment; SNR = signal-to-noise ratio. [a]Reprinted under Creative Commons CC BY 4.0. [b]Reprinted under Creative Commons CC BY-NC-ND 4.0. [c]Reprinted under Creative Commons CC BY-NC 4.0. [d]Reprinted with permission of Wolters Kluwer Health, Inc. [e]Reprinted with permission of Georg Thieme Verlag KG.