VPS VIBRATIONS IN PHYSICAL SYSTEMS

DOI: 10.21008/j.0860-6897.2022.1.09
Vibrations in Physical Systems, 2022, 33(1), 2022109

On the Feasibility of Assessing Social Situations Using Ecological Momentary Assessment – A Comparison Between Singapore and German Test Participants

Nadja SCHINKEL-BIELEFELD[1], Zu Xuan LEE[2], Khine CHO CHO THEIN[3], Jennifer MARTIN[2], Sonia STASIAK[3]

[1] WS Audiology, Henri-Dunant-Str. 100, 91058 Erlangen, Germany
[2] National University of Singapore, 21 Lower Kent Ridge Road, Singapore 119077
[3] WS Audiology, 18 Tai Seng Street, #08-08, Singapore 539775

Corresponding author: Nadja SCHINKEL-BIELEFELD, email: nadja.schinkel-bielefeld@wsa.com

Abstract Ecological Momentary Assessment (EMA) is a method comprising repeated self-reports in the participant's natural environment which can be used to evaluate hearing aids in real life. Social situations are particularly important for such evaluations as these are situations where listening is critical and often difficult. However, as shown for a German subject sample by Schinkel-Bielefeld et al., these are also situations where test participants may skip a questionnaire as it could be perceived as impolite to answer a questionnaire on the smartphone. This leads to the underrepresentation of speech in noise situations in EMA assessment. However, the acceptance of smartphone use in public depends on cultural norms and may be larger in Asian countries. Repeating the previous study with 10 Singaporean hearing-impaired test participants, we observed that self-reports are often answered with some delay but not skipped in social situations. Also, contrary to the German study, speech in noise situations were not underrepresented in questionnaires in the Singapore study.

Keywords: ecological momentary assessment, cultural differences, compliance, bias, social situations, phubbing.

1. Introduction

Ecological Momentary Assessment (EMA) is a method to evaluate the test participants' experiences in everyday life in real time. It consists of repeated self-reports, where participants answer questions on their current intention, perception, or emotion in their usual environment [1, 2].

Nowadays EMA is often performed with a smartphone which also collects objective data in addition to the self-reports. However, sometimes it is inconvenient, regarded impolite, or even dangerous to answer a questionnaire in the moment. Prompting participants at random times to fill out a questionnaire should result in a distribution of situations that resembles those that participants experience in daily life. However, this is no longer true if participants skip or postpone questionnaires in certain situation or do not carry the study phone with them in selected situations, resulting in selection bias.

For many research questions it may not be important if the distribution between reported and experienced situations differs. However, if participants skip questionnaires in group conversations, experiments may miss out on some of the most interesting and relevant situations for assessing hearing problems and the potential benefit of hearing solutions, such as hearing aids.

Especially group conversations are situations which are crucial for the evaluation of hearing problems and/or technical solutions to those problems, as speech in noise situations are often difficult for individuals with hearing impairment. However, depending on the cultural background it can be perceived as impolite to use the smartphone to answer questionnaires during a conversation.

In a study with 20 German test participants [3] it was shown that the perception that answering a phone-based questionnaire would be inappropriate and impolite is an important reason for not filling out questionnaires. Also, when asked for situations in which test participants did not initiate a questionnaire despite having something to report and situations where they would not bring their phone, social situations are the most common answer. As a result, questionnaires answered in speech in noise situations are underrepresented compared to the total fraction of speech in noise situations experienced.

Smartphone adoption rates differ between countries. With 94% of all Singapore residents owning a smartphone in 2019 the smartphone adoption rate in Singapore was the highest in the Asia pacific region

[4]. In contrast, only 78% of Germans own a smartphone [5]. Also, the smartphone usage rate for elderly is higher in Singapore: 76% of people aged 60 and older use a smartphone in Singapore, but only 64% aged 50 and older use a smartphone in Germany [5,6].

Also smartphone usage patterns can differ greatly between different cultural contexts. For example, using the Smartphone Addiction Scale-Short Version [7], several studies find a higher prevalence of 'problematic smartphone usage' in the investigated southeast asian countries than in many European countries [8] with a rather low rate for German students [9].

There is a growing body of literature analysing the effect of phone usage in the copresence of other people, also called 'phubbing'. While it has detrimental consequences for the quality of the conversations [10] it is not uncommon. According to Leuppert and Geber [11] phubbing behaviour is dominated by the perception that those around us are also doing it (descriptive norm) than the feeling that it is socially accepted (injunctive norm) especially in group conversations. Hence, we hypothesized that it may be easier to obtain responses in social situations for EMA performed in Singapore than in Germany.

In order to test this hypothesis and to analyse if social situations are similarly underrepresented in EMA in a Singapore subject population than in Germany we replicated the previous study in Singapore and compared results between the two sample populations.

2. Methods

2.1. Test Participants

Ten Singaporean Chinese experienced hearing aid wearers (mean age 66.9 years, range: 47-73 years) participated in the study. Inclusion criteria were native English or Mandarin Chinese speakers who reported to be comfortable using a smartphone with at least one year experience of wearing hearing aids and a bilateral average hearing loss between 35 and 75 dB HL (frequencies 0.5, 1, 2, and 4 kHz). The recruited participants had a mean pure tone average of 59.4 dB HL (range: 38.8 -75.0 dB HL).

The National University of Singapore Institutional Review Board approved the study (NUS-IRB 2020-43) on 26 October 2020. Participants filled out informed consent prior to participating in the study.

2.2. EMA App

The EMA app (version 4.1.6) administered up to 12 notifications to fill out a questionnaire per day. 8 of those were time based and the remaining ones were only issued in environments louder than 65dB SPL. Notifications were not issued if no Bluetooth connection between phone and hearing aids was available or if a questionnaire has been filled out within the last 15 minutes. Participants also had the possibility to self-initiate a questionnaire anytime. If a prompted questionnaire was not answered within 15 minutes after the notification, another questionnaire was issued that politely asked for the reason for missing the first questionnaire. Furthermore, participants filled out an end of day-questionnaire (EoD-questionnaire) each evening to reflect on situations where they did not bring the study phone, did not trigger a user-initiated questionnaire or observed connection problems.

In addition, classifier data and the mean level was collected by the hearing aid for each minute. Classes comprised Quiet, Noise, Speech in Quiet (SiQ), Speech in Noise (SiN), Car and Music. Sometimes it is not possible or not convenient to answer the questionnaire immediately. To ensure that the objective data can be matched to the subjective data, participants were asked to at least start the questionnaire in such situations. They then have 30 minutes to complete it.

More detailed information on the EMA app and the questionnaires used can be found in [3] (One question enquiring about the preference of the hearing program was omitted here as there was only one hearing program in this study). Questionnaires were forward-backwards translated from German into English and from English into Mandarin Chinese. As the classifier information is saved for 24 hours in newer hearing aids, the app was also able to collect it retrospectively if the Bluetooth collection was interrupted temporarily. However, for the comparison to the previous study no retrospectively collected data was taken into account.

2.3. Experimental Procedure

At the first appointment test participants were fitted with Pure 312 7X hearing aids using the proprietary X-fit formula with experience level 'New'. An own voice training was performed, and fine tuning was added if required. Test participants had one program and could control the volume of the hearing aids.

Participants were provided with a Samsung Galaxy S7 smartphone with preinstalled EMA app along with a written manual. There was an oral introduction into the app and participants filled out current situation questionnaires for three different situations as a training.

In the following two weeks participants were asked to wear the hearing aids, keep the study phone close by and respond to the questionnaire triggers. In addition, they should fill out 5 user-initiated triggers per day. They had the possibility to configure do not disturb-times (DND-times) where no triggers were issued.

The number of answered questionnaires was monitored through a web interface and any potential issues were discussed during a follow up phone call one week after the start of the study.

Upon returning the equipment, participants filled out a questionnaire enquiring about the perceived burden of the study. They were paid up to 50 SGD, including a bonus of 10 SGD for not missing any EoD-questionnaires and a bonus of 5 SGD per week when at least 8 current situation questionnaires per day were answered.

2.4. Data Analysis

Data analysis has been performed in Matlab. Analogous to [3] we first computed distributions of environments of ratios of questionnaires for each participant and then averaged over all participants to ensure that each participant was weighted the same, independent of total wearing time and number of questionnaires. All significance tests are one-sided, paired sample permutation tests using bootstrapping with 100000 repetitions [12]. Effect sizes are computed with the effect size toolbox [13].

3. Results

3.1. Data collected

Over the course of two weeks, we obtained 126 EoD-questionnaires and 1244 questionnaires about the current situation resulting in an average of 8.2 questionnaires per day. Questionnaires contained a short mandatory part with 4-7 questions and a long optional part with up to 37 questions. In 68.7% of all cases the long questionnaire was filled out, which was less than in the German study [3] where 86.7% of collected questionnaires contained the long optional part. With 61.0% self-initiated questionnaires, 26.2% random trigger questionnaires and 12.8% loud environment triggers, there was a larger percentage of self-initiated questionnaires than in the German study where only 32.7% of questionnaires were self-initiated.

The average daily wearing time of the hearing aids ranged from 9 to 15 hours (mean 12.5 hours). The app collected on average 81% of the available data of the classifier. However, when discarding all the data that was collected retrospectively and would not have been collected with the equipment of the previous study, 40% of the theoretically available data are left.

3.2. Missing subjective data

As we wanted to find out in which situations triggered questionnaires were skipped, there was a follow up question if the original questionnaire was not started within 15 minutes after the notification. In this follow-up questionnaire we politely asked participants for the reason of skipping. If they responded that they considered answering inappropriate in that situation, they were asked in what way it was inappropriate. While in both study populations not hearing the alarm was a common reason for not answering the questionnaire, German subjects indicated that answering the questionnaire was inappropriate in that situation twice as often as Singaporean subjects (Fig. 1a). While in the Singapore study cohort 'inappropriate' referred mostly to safety issues, Germans reported impoliteness as the main reason answering a questionnaire was inappropriate (Fig. 1b).

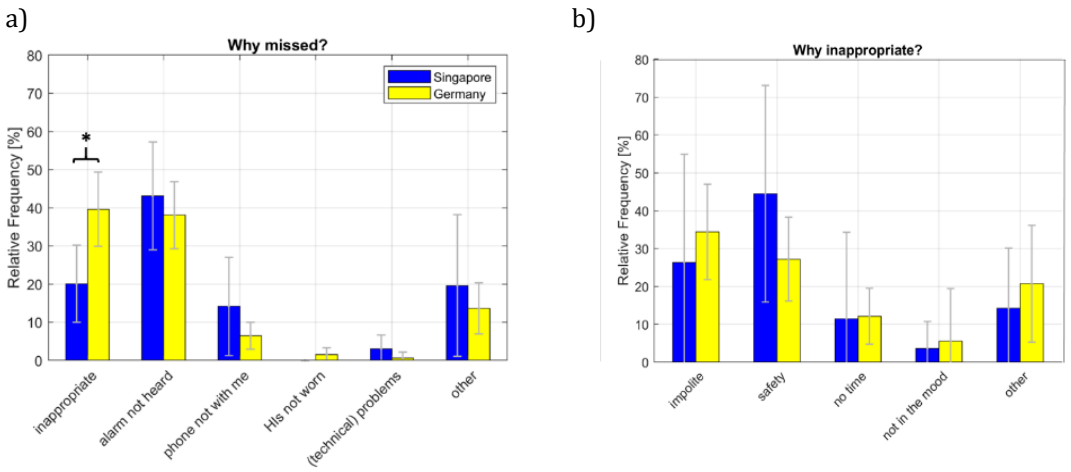

Figure 1. (a) Responses to the question why a questionnaire was missed. (b) If participants indicated that responding would have been inappropriate, a follow up question asked in what way it was inappropriate.

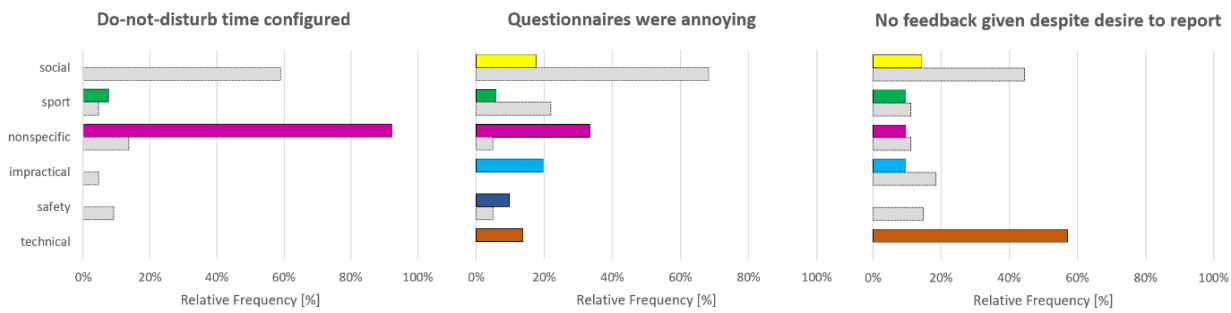

Figure 2. Situations in which participants configured DND-times, were annoyed by questionnaires or did not self-initiate a questionnaire despite the desire to report something, as reported in the EoD-questionnaire. Results for Singaporean participants are in color. For reference, results of German participants are in grey.

While for German test participants social situations were by far the most common reason for missing subjective feedback, Singapore test participants only mentioned them in less than 20% of the cases for not reporting despite desiring to give feedback and questionnaires being annoying. They did not report to have configured DND-times for social situations (Fig. 2).

For the German participants the tendency to skip questionnaires in social situations was also visible in the objective data. However, comparing the 3 minutes before a random trigger questionnaire for Singapore test participants to the entire wearing time data collected by the EMA-app, yields no significant difference in the fraction of SiN situations (Fig. 3). Car classed situations are significantly underrepresented in random trigger questionnaires ($p < 0.05$, Hedges' $g = 0.91$), and SiQ situations are overrepresented ($p < 0.05$, Hedges' $g = 0.64$).

If not skipping questionnaires, it could still be that in certain situations test participants may fill out only the short questionnaire leading to less complete information in these situations. There was one subject who did not fill out any long questionnaire. Analysis of the remaining nine test participants (Fig. 4a) revealed a trend towards larger fractions of long than of short questionnaires for conversations with more than one person ($p = 0.06$, Hedges $g = 0.73$) as well as for radio/TV ($p = 0.05$, Hedges' $g = 0.50$).

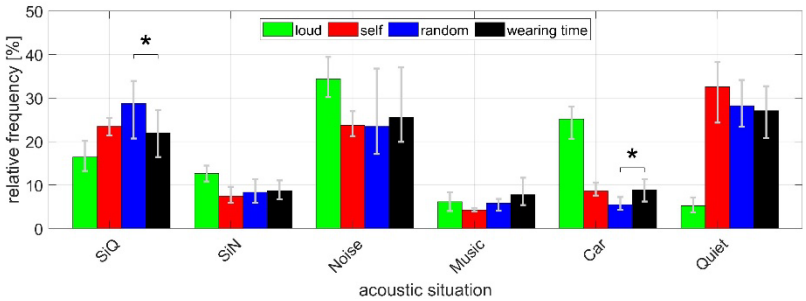

Figure 3. Distribution of the classification of the acoustic situations three minutes before the start of the questionnaire and throughout the data for the entire wearing time collected by the EMA app. Stars mark a significant difference between the distribution of the acoustic classes for the random triggers and the entire wearing time.

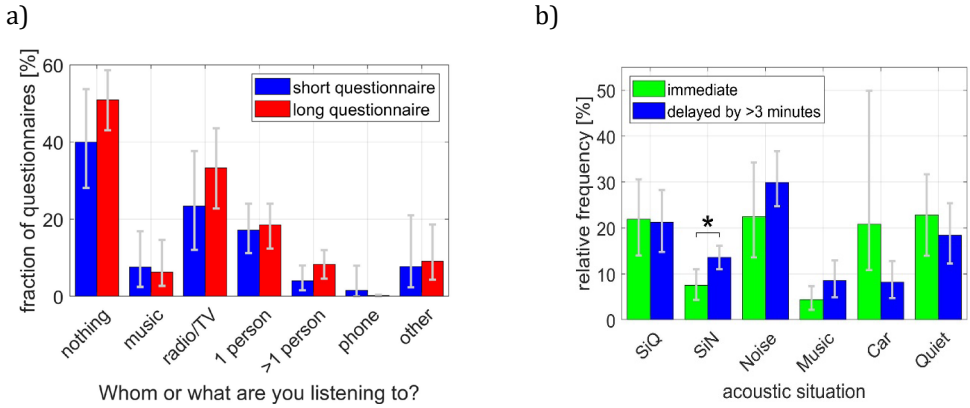

Figure 4. a) Fraction of short and long questionnaires filled out for different listening intents. b) Distribution of acoustic situations in the three minutes before questionnaires that were answered immediately after the prompt and those that were delayed by at least three minutes.

Including all prompted current situation questionnaires, we compared the distribution of acoustic situations for questionnaires that were answered immediately after the prompt and those that were at least 3 minutes delayed (Fig. 4b). This revealed that in SiN situations responses are significantly more often delayed ($p > 0.01$, Hedges' $g = 1.14$).

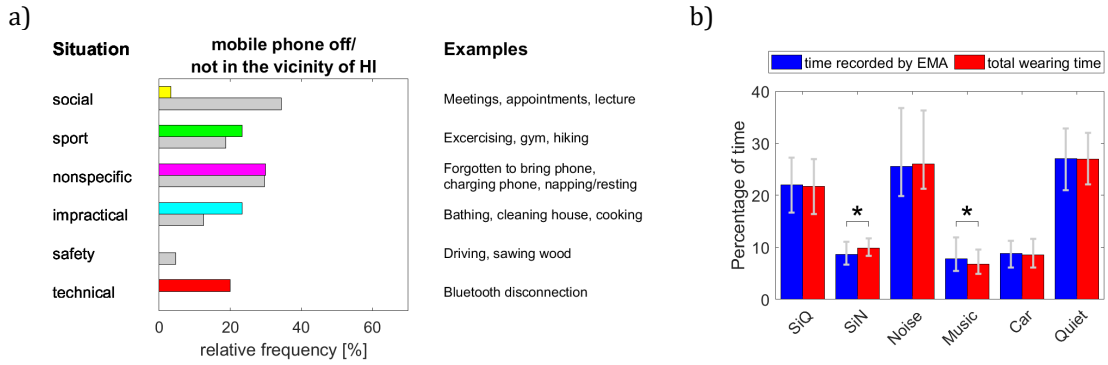

Figure 5. a) Reasons for the study phones being switched off or not in the vicinity. b) Distribution of acoustic situations in the data collected by the EMA app (dependent on Bluetooth connection) and during the entire wearing time.

3.3. Missing objective data

Answering questionnaires requires to carry along the study phone. This is also required if any objective data from the hearing aids should be collected that is not saved in the hearing aids for a prolonged time. While German participants often did not carry along their study phones in social situations, social situations only were responsible for 3.3% of all instances where Singaporeans did not carry along their study phones (Fig. 5a). Comparing the data collected by the EMA app (dependent on Bluetooth connection) and the data collected by the hearing aid read out at the end of the study which comprised the entire wearing time independent of Bluetooth connection reveals a small, but significant underrepresentation of SiN (Hedges' g = 0.36) and overrepresentation of Music (Hedges' g = 0.22) in the data collected by the EMA app (Fig. 5b).

4. Discussion

Compared to German test participants, Singaporeans reported skipping fewer questionnaires because it would be inappropriate or impolite to answer in that situation (Fig. 1). Rather the main reason for missed questionnaires was not hearing the notification sound. Similarly, they did not configure DND-times for social situations and in less than 20% of cases they considered answering questionnaires in social situations annoying or did not initiate them despite having something to report (Fig. 2). These results are based on the self-reports of subjects. Given that participants often have a desire to give socially desirable responses [14], responses of the subjects could be even more strongly influenced by social norms than their actual behaviour. However, the objective data finds no significant underrepresentation of speech in noise situations, but an overrepresentation of speech in quiet situations (Fig. 3), confirming the learnings from the subjective data.

Also in Singapore, there are differences between the distribution of situations in the entire wearing time and shortly before the random trigger questionnaires, such as an underrepresentation of car classed situations and an overrepresentation of SiQ. However, especially speech situations are crucial for research in audiology, as often it is the main goal to improve speech understanding of hearing-impaired individuals enabling them to fully take part in social life. Our results indicate that contrary to German test participants, Singaporeans tend to have no problems to answer questionnaires in speech situations, hence making it more feasible to use EMA to do research on these situations in real life.

Singapore test participants frequently reported technical difficulties as reasons for being annoyed by questionnaires and not initiating any themselves. While German test participants also encountered technical difficulties and loss of Bluetooth connection, they never mentioned this as a reason for not filling out a questionnaire but rather apologized for not being able to fix them themselves when calling the experimenter to get help. Thus, the way to deal with app malfunction and technostress also influence results differently for the two cultures.

While Singaporean subjects rarely (3.3%) reported social situations to be the reason for not carrying along the study phone, a significant underrepresentation of SiN situations is visible in the data collected by the EMA app compared to the data of the entire wearing time. However, with an effect size of 0.36 this effect is small.

In both studies, test participants answered with more delay in SiN situations (Fig. 4). One reason to use EMA is to avoid or reduce the memory bias that is often present in retrospective questionnaires. However, it is desirable to interrupt the test participants' life as little as possible. Our results seem to indicate, that it is advisable to give participants at least a few minutes time to react to the trigger in order to not miss out on social situations. This seems to be the case regardless of whether social norms permit the use of smartphone in that situation, or the answering of the questionnaire is simply interfering with the main task the test participant is currently pursuing. Also, as test participants in EMA studies soon know the questions by heart, they likely pay more attention to the topic of investigation and can give a more accurate answer slightly retrospectively, than if they did not know the question in advance.

One would expect more long questionnaires in difficult listening situations as there might be more to report. However, for the German subject population this was not the case for conversations with several, though these are often challenging. Possibly the reason for this was that Germans considered it impolite to use the study phone in these situations and thus did not want to spend too much time on it. While not significant, a trend towards more long questionnaires for conversations with several was observable for Singapore subjects (Fig. 4). This indicates that while it is advisable to shorten questionnaires for German participants to allow them a quick answer that does not violate social norms too much, it seems easier to get more detailed information in those situations from Singapore subject.

4.1. Limitations of the study

Payment schemes differed between the studies. While the German study participants were paid by questionnaire (and received more money for long than for short questionnaires) the Singapore participants only received an extra payment when filling out at least 8 current situation questionnaires per week. This could have influenced the percentage of long questionnaires. However, as the payment was independent of the situation where the questionnaire was filled out, it is unlikely that this changed the distribution of acoustic environments and listening intents in those situations where subjects answered a questionnaire.
Also, the study duration was shorter in the Singapore study than in the German counterpart. One could imagine that subjects get tired of answering questionnaires over time and start to skip them first in the situations where they are most inconvenient. However, the underrepresentation of SiN situations was only slightly lower in the first two weeks of the German study than over the total study duration (Hedges' g = 0.47 instead of 0.53). This cannot explain the differences between German and Singaporean participants.

While both studies were conducted around Christmas, the study in Singapore was performed during the Covid-19 pandemic and the German study two years earlier. At the time of the study the phase three of reopening was active in Singapore, meaning that gatherings of up to eight people, congregational worship services of up to 250 people and indoor and outdoor performances up to 250 and 100 people respectively (in zones of 50 people) were allowed. Karaoke and nightlife establishments were closed. Also, safe distances were strongly enforced. Despite these restrictions, there were many possibilities to participate in social life as for example, going to the office, to cafes and restaurants, shopping malls and even theatres.

While it is possible that the pandemic had an influence on the kind and absolute number of social situations subjects participated in, we would not expect that it has a major influence on the difference between the distribution of situations participants experienced and those they answered questionnaires in. Hence, an underrepresentation of social situations should be observable if it existed. Also, since the fitting of hearing aids was performed in person, people avoiding any nonessential social contact probably did not participate in the study.

Humans like to give socially acceptable answers [14] which can lead to different reporting biases for different cultures, i.e., if Germans perceive it as more socially acceptable to skip a questionnaire in social situations than Singaporeans, they are also more likely to report this reason even if several reasons for not responding may have been present. Thus, responses in the EoD-questionnaire and in the questionnaire enquiring about the reasons for missing a questionnaire, could overemphasize cultural differences. However, objective data are less prone to such biases and are consistent with the results of the subjective data suggesting that the tendency to skip questionnaires in speech in noise situations or social situations indeed is stronger in Germans than in Singaporeans.

This study and to a lesser degree also the German study [3] comprised a relatively small sample size. Hence, they could be influenced by selection bias and do not allow for a very precise quantification of the observed effects. Also, subject behaviour can be influenced by exact instructions of the study lead. While we tried to closely mimic the German study, both studies were performed by study leads of different cultural background who could have subconsciously influenced subjects in different ways. Hence, our initial results should be verified with a larger sample size after the end of the Covid-19 pandemic, preferably by one study lead performing the study for test participants in different cultural contexts.

5. Conclusions

Our results suggest that it may be easier to get subjective feedback in social situations when performing the study in a cultural context where smartphone usage is widely accepted also in social situations. This may be an advantage if the main goal of a study is for example to evaluate technical aspects of a hearing technology, such as objectively perceivable artifacts.

However, different cultures differ in many aspects, such as lifestyle, preference, and social norm. Thus, depending on the research question, results may differ greatly across cultures and are possibly not generalizable to a different target group as soon as the distribution of situations experienced or the preference of the individual subject matters. Hence it should be carefully considered, if the advantages of more subjective feedback outweigh any potential problems generalizing the study results to the desired target population, when deciding to choose a study population based on the ease of obtaining subjective feedback in EMA. In any case, it seems to facilitate reporting about social situations, if test participants are allowed to answer the prompts with some delay.

Acknowledgments

We thank Bastian Buder (Develappers GmbH) for development and maintenance of the EMA app.

Additional information

The authors declare no competing financial interests.

References

1. S. Shiffman, A.A. Stone, and M.R. Hufford; Ecological momentary assessment; Annu Rev Clin Psychol, 2008, 4, 1-32.
2. A.A. Stone, S. Shiffman; Ecological Momentary Assessment (EMA) in behavioral medicine; Annals of Behavioral Medicine, 1994, 16(3), 199-202.
3. N. Schinkel-Bielefeld, et al.; Evaluation of Hearing Aids in Everyday Life Using Ecological Momentary Assessment: What Situations Are We Missing?; Am J Audiol, 2020, 29(3S), 591-609.
4. R. Clapp; Singapore tops APAC for smartphone penetration; 2019; Available from: https://www.warc.com/content/paywall/article/warc-datapoints/singapore_tops_apac_for_smartphone_penetration/125723
5. L. Silver; Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally; Pew Research Center, Report, Washington, 2019; Available from: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally
6. Annual survey on infocomm usage in households and by individuals for 2019; Infocomm Media Development Authority, Singapore, 2019. Available from: https://www.imda.gov.sg/infocomm-media-landscape/research-and-statistics/infocomm-usage-households-and-individuals.
7. M. Kwon, et al.; The smartphone addiction scale: development and validation of a short version for adolescents; PloS one, 2013, 8(12), e83558.
8. Z. Yang, K. Asbury, M.D. Griffiths; Do Chinese and British university students use smartphones differently? A cross-cultural mixed methods study; International Journal of Mental Health and Addiction, 2019, 17(3), 644-657.
9. B. Lachmann, et al.; The role of empathy and life satisfaction in internet and smartphone use disorder. Frontiers in Psychology, 2018, 9, 398.
10. V. Chotpitayasunondh, K.M. Douglas; The effects of "phubbing" on social interaction; Journal of Applied Social Psychology, 2018, 48(6), 304-316.
11. R. Leuppert, S. Geber; Commonly done but not socially accepted? Phubbing and social norms in dyadic and small group settings; Communication Research Reports, 2020, 37(3), 55-64.
12. B. Efron; The jackknife, the bootstrap and other resampling plans; CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, 1982.
13. H. Hentschke, M.C. Stüttgen; Computation of measures of effect size for neuroscience data sets; European Journal of Neuroscience, 2011, 34(12), 1887-1894.
14. T.F. Van de Mortel; Faking it: social desirability response bias in self-report research; Australian Journal of Advanced Nursing, 2008, 25(4), 40-48.