

Protocol

The Soundscape Indices (SSID) Protocol: A Method for Urban Soundscape Surveys—Questionnaires with Acoustical and Contextual Information

Andrew Mitchell ^{*} , Tin Oberman , Francesco Aletta ^{*} , Mercede Erfanian ,
Magdalena Kachlicka , Matteo Lionello  and Jian Kang ^{*} 

UCL Institute for Environmental Design and Engineering, The Bartlett, University College London,
Central House, 14 Upper Woburn Place, London WC1H 0NN, UK; t.oberman@ucl.ac.uk (T.O.);
mercede.erfanianghasab.18@ucl.ac.uk (M.E.); magdalena.kachlicka.18@ucl.ac.uk (M.K.);
matteo.lionello.18@ucl.ac.uk (M.L.)

^{*} Correspondence: andrew.mitchell.18@ucl.ac.uk (A.M.); f.aletta@ucl.ac.uk (F.A.); j.kang@ucl.ac.uk (J.K.)

Received: 26 February 2020; Accepted: 24 March 2020; Published: 1 April 2020



Abstract: A protocol for characterizing urban soundscapes for use in the design of Soundscape Indices (SSID) and general urban research as implemented under the European Research Council (ERC)-funded SSID project is described in detail. The protocol consists of two stages: (1) a Recording Stage to collect audio-visual recordings for further analysis and for use in laboratory experiments, and (2) a Questionnaire Stage to collect in situ soundscape assessments via a questionnaire method paired with acoustic data collection. Key adjustments and improvements to previous methodologies for soundscape characterization have been made to enable the collation of data gathered from research groups around the world. The data collected under this protocol will form a large-scale, international soundscape database.

Keywords: soundscape; SSID; binaural recordings; ambisonic recording; questionnaire; visual factors; ISO12913

1. Introduction

Soundscape studies strive to understand the perception of a sound environment, in context, including acoustic, (non-acoustic) environmental, and contextual, and personal factors. These factors combine together to form a person's soundscape in complex interacting ways [1]. In order to predict how people would perceive an acoustic environment, it is essential to identify the underlying acoustic and non-acoustic properties of soundscape.

The soundscape community is undergoing a period of increased methodological standardization in order to better coordinate and communicate the findings of the field. This process has resulted in many operational tools designed to assess and understand how sound environments are perceived and apply this to shape modern noise control engineering approaches. Important topics which have been identified throughout this process are soundscape 'descriptors', 'indicators', and 'indices'. Aletta et al. [2] defined soundscape descriptors as "measures of how people perceive the acoustic environment"; soundscape indicators as "measures used to predict the value of a soundscape descriptor"; and soundscape indices can then be defined as "single value scales derived from either descriptors or indicators that allow for comparison across soundscapes" [3]. (Please find the link to the Soundscape Indices (SSID) ERC-funded website in the Supplementary Materials.)

This conception has recently been formalized and expanded upon with the adoption of the recent ISO 12913 standard series [4–6]. ISO 12913 Part 1 sets out the definition and conception of Soundscape, defining it as the "acoustic environment as perceived or experienced and/or understood by a person

or people, in context". Here, the soundscape is separated from the idea of an acoustic environment, which encompasses all of the sound which is experienced by the receiver, including any acoustically modifying effects of the environment. In contrast, the soundscape considers the acoustic environment, but also considers the impact of non-acoustic elements, such as the listener's context and the visual setting, and how these interact with the acoustic environment to influence the listener's perception.

The ISO/TS 12913-2:2018 is the current reference document addressing data collection and reporting requirements in soundscape studies. In terms of methods, the ISO document covers two main approaches, namely: soundwalks combined with questionnaires (Methods A and B) and narrative interviews (Method C) [5], which relate to on-site and off-site data collection, accordingly. Part 3 of the ISO 12913 series builds on Part 2 and provides guidelines for analyzing data gathered using only those methods [6]. However, the range of possible methodological approaches to soundscape data collection is much broader and it includes, for instance, laboratory experiments [2,7,8], pseudo-randomized experience sampling [9], and even non-participatory studies [10]. The protocol described in this paper was designed having in mind the need for a relatively large soundscape dataset that could be used for design and modeling purposes, thus trying to expand the scope of soundwalks that typically deal with much smaller samples of participants [11]. For the sake of comparability and standardization with these methods, we chose to refer to the soundscape attributes reported in the ISO Part 2 (Method A).

Several studies prior to the formalization of the ISO standards on soundscape demonstrated the general, but inadequate, relationship between traditional acoustic metrics, such as L_{Aeq} , with the subjective evaluation of the soundscape [1,12–15]. These have typically aimed to address the existing gap between traditional environmental acoustics metrics and the experience of the sound environment. Yang and Kang (2005) showed that, when the sound level is 'lower than a certain value, say 70 dBA', there is no longer a significant change in the evaluation of acoustic comfort as the sound level changes. However, the perceived sound level does continue to change along with the measured sound level, showing that (1) measured sound level is not enough to predict soundscape descriptors such as 'acoustic comfort', and (2) there is a complex relationship between perceived sound level and soundscape descriptors which is mediated by other factors.

Subsequent studies have shown that, even with large data sets and several possible acoustic indicators examined, models that are based on objective/measurable metrics under-perform in predicting soundscape assessment when compared to models based on perceptual responses. Ricciardi et al. [16], with a methodology based on smart phone recordings, achieved $R^2 = 0.21$ with acoustic input factors L_{50} and $L_{10} - L_{90}$, whereas the same dataset and model building method achieved $R^2 = 0.52$ with perceptual input factors overall loudness (OL), visual amenity (VA), traffic (T), voice (V), and birds (B). This indicates that merely examining the acoustic level is not sufficient for predicting the assessed soundscape quality, and that additional objective factors and a more holistic and involved method of characterizing the environment is required. This protocol is trying to extend the scope of objective measurements that are being collected in conjunction with perceptual responses by including other environmental and visual data. These previous studies have generally been limited by one or many of the following factors: limited number or types of locations, limited responses sample size, and no non-acoustic factors, generally limiting the generalizability of their results beyond the investigated locations.

The ability to predict the likely soundscape assessment of a space is crucial to implementing the soundscape concept in practical design. Current methods of assessing soundscapes are generally limited to a post-hoc assessment of the existing environment, where users of the space in question are surveyed regarding their experience of the acoustic environment [11,17]. While this approach has proved useful in identifying the impacts of an existing environment, designers require the ability to predict how a change or proposed design will impact the soundscape of the space. To this end, a model that is built upon measurable or estimate-able quantities of the environment would represent a leap forward in the ability to design soundscapes.

Developing soundscape indices is a process that requires consideration of how people perceive, experience, and understand the surrounding sound environment. For the purpose of modeling and comparisons, it is important that such indices are numerical entities and that these quantities are collected consistently across all investigated spaces and soundscapes. Although the soundscape approach taken in this protocol represents a step-change away from existing methods of noise exposure measurements, strong cues particularly in the realm of acoustic measurement methods should be taken from existing standards both to make use of the significant knowledge and experience that has gone into the creation of these standards and to facilitate compatibility between soundscape and traditional measurements. In general, the measurement methods and best practice given in environmental noise standards such as ISO 1996-1:2016 and ISO 1996-2:2017 [18,19] should be followed wherever possible, including the use of standardized acoustic equipment such as standard sound level meters.

An European Research Council (ERC) Advanced Grant project is ongoing to develop the proposed “Soundscape Indices” (SSID), which adequately reflect levels of human comfort and preference while integrating measurable and observable quantities. The framework proposed for the SSID project is laid out in detail by Kang et al. [3], the first step of which is generating a large-scale and coherent database of the required soundscape characterization data. Given the already recognized differences in soundscape assessment across various countries and cultures [20,21] and the success of existing international soundscape efforts such as the Soundscapes of the World project [22], the collection of soundscapes from many different countries and in many different contexts is an important component of the SSID project.

Therefore, the following protocol has been conceived and implemented within the SSID framework to collect data about urban soundscapes for use in general soundscape research and toward the design of Soundscape Indices. Thus far, the collected database includes nearly 4000 participants’ responses from 59 locations in 10 cities and provinces across the UK, China, Spain, and Italy. This protocol has been refined and adjusted as needed during this extensive data collection process to arrive at this final version. This work was conducted by nine associated research groups and coordinated by the SSID group based at University College London and has already produced several pieces of published work towards the creation of Soundscape Indices [23–30]. Additional collaborations and data collection efforts are currently underway in France, the Netherlands, and Croatia.

Purpose

This protocol was designed to achieve two primary goals: (1) gather in situ soundscape assessments from the public, which can be further analyzed and utilized in designing a soundscape index; (2) conduct recordings needed to reproduce the audio-visual environment of a location in a laboratory setting for conducting controlled experiments on soundscape. These two goals represent two levels of data required for developing a general soundscape model. The first enables large scale data collection, resulting in a database with thousands of perceptual responses and their corresponding quantitative data which can be statistically analyzed on a large scale, or used for training in machine learning modeling. In situ assessments also represent the most holistic assessment, ensuring all factors that influence the soundscape are present, including those which cannot be reproduced elsewhere.

However, there are questions that cannot be practically addressed in situ, such as soundscape assessment of less- or un-populated areas, the influence of mismatched acoustic and visual cues, physiological and neural responses to various soundscapes, and so on [31]. Laboratory experiments with controlled environments are required to address these aspects. Toward the development of a coherent SSID, however, it is important that these two forms of data are collected simultaneously and with compatible methods, such that the results of the two approaches can be confidently combined and compared. In addition, since this protocol is intended to be used for the creation of a large-scale international database with additions carried out by several different and remote teams, it has been designed for efficiency, scalability, and information redundancy.

2. Protocol Design and Equipment

The first goal is achieved by conducting in situ questionnaires using a slightly altered version of Method A (questionnaire) from Annex C of the ISO/TS 12913-2:2018 technical specification [5] collected either via handheld tablets or paper copies of the questionnaire. Typically, a minimum of 100 responses are collected at each location during multiple 2–5-h sessions over several days. During the survey sessions, acoustic data are collected via a stationary class 1 or class 2 Sound Level Meter (SLM) (as defined in IEC 61672-1:2013 [32]) running throughout the survey period and through binaural recordings taken next to each respondent. These acoustic and response data are linked through an indexing system so that features of the acoustic environment can be correlated with individual responses or with the overall assessment of the soundscape, as required by researchers.

The second goal is achieved by making First-Order (or higher) Ambisonic recordings simultaneously with 360° video which can be reproduced in a virtual reality environment. It has been shown that head-tracked binaural and multi-speaker ambisonic reproduction of recorded acoustic environments recorded in this way have high ecological validity [33], particularly when paired with simultaneous head-tracked virtual reality video [22,34,35].

The on-site procedure to collect these data are separated into two stages, which will be outlined in detail in Section 4. The stage during which the spatial audio-visual recordings are made for lab experiments is called the **Recording Stage**, while the stage during which questionnaires and environmental data are captured is called the **Questionnaire Stage**.

The procedure has been designed to include multiple levels of data and metadata redundancy, making it robust to on-site issues and human error. The most crucial aspect of the redundancy is ensuring perceptual responses can be matched with the appropriate corresponding environmental and acoustic data even when some information is lost or forgotten.

2.1. Labeling and Data Organization

In order to be able to identify all of the many data components of the Recording and Questionnaire Stages and to associate these with their various corresponding data, the following labeling system is suggested. This system is focused on (1) relating all of the separate recordings and factors to specific questionnaire responses and (2) efficiency and consistency on site. A recent paper by Aumond et al. [14] demonstrated the importance of addressing multiple levels of factors which influence perception, from individual-, to session-, to location-level. The successful pleasantness models building incorporating these information levels showed a marked improvement over the equivalent individual-level or location-level only models. The data organisation system proposed here was designed in order to maintain this important information, and the levels of information for the data collected on site are shown in Table 1.

At the top level is the **Location** information. This includes information about the location which does not change day-to-day, and generally characterizes the architectural character of the space, or typical climate conditions for the area. As described in Section 2.2, each 'environmental unit' should be considered a new location. Therefore, if researchers want to investigate the differences in soundscape assessment in the middle of a small urban park and along the road next to the same park, these would be considered different locations since they would (typically) have different environmental factors, and should be given different names. The name chosen should be concise, but it should be obvious what location is referred to.

The next level is information which is specific to each session, labeled with a **SessionID**. This SessionID should contain the name of the location and a numerical index which will increase with each repeated session at that location. The SessionID is associated with the data collected during the Recording Stage, and with the data which are continuous throughout the Questionnaire Stage, SLM, and ENV data. For easy automatic processing, correct spelling and consistency with the format is crucial so that data can be filtered according to the SessionID or the location, as is often necessary.

In addition, for ease of automatic processing, it is recommended not to include spaces in the SessionID to avoid string splitting issues in analysis code.

Underneath each SessionID will be a set of **GroupIDs**. One GroupID is assigned for *each group of participants*. This should correspond to a single binaural recording and a single 360° photo. This will be used to (1) relate multiple surveys taken simultaneously and (2) link the recording and photo with the surveys. The GroupID is particularly crucial as it allows commonly missing data to be shared across multiple collection methods. For instance, occasionally paper questionnaires will be missing start and end time information. In this case, this information can be pulled directly from other questionnaires with the same GroupID. Where no questionnaires have the times, it is possible to extract an approximate start time from the binaural recording or 360° photo and then estimate an average end time.

The GroupID should have the following format: [a set of letters representing the location name][the SessionID index number][an incrementing index for each group]. For example, for the second session at Regent’s Park Japanese Garden, the location name is RegentsParkJapan, the GroupID letters might be ‘RPJ’; the SessionID would be ‘RegentsParkJapan2’ so the GroupIDs for that session would start at ‘201’. Therefore, for example, the tenth group of participants for that session would be labeled ‘RPJ210’. This format ensures that, if the location or SessionID are not recorded for a questionnaire, it is still obvious which session it belongs to.

Table 1. Labeling system for on site data collection. Regent’s Park Japanese Garden is used as an example location. SLM: Sound Level Meter (acoustical factors); ENV: Environmental factors; BIN: Binaural; QUE: Questionnaires; PIC: Site pictures.

| Level of Information | Example Label | | | | | Factors Measured at This Level |
|----------------------|-------------------|-------------------|-----|--------|-----|--|
| Location | RegentsParkJapan | | | | | GPS, Architectural typology, visual openness, etc. |
| SessionID | RegentsParkJapan1 | RegentsParkJapan2 | | | | SLM, session notes, ENV |
| GroupID | RPJ101 | RPJ102 | ... | RPJ201 | ... | BIN, PIC |
| Questionnaire | 1, 2, 3 | 4, 5 | ... | 25, 26 | ... | QUE, Start & End time |

2.2. Location and Measurement Point Selection

To select the appropriate measurement point, it should be ensured that the following contextual factors representative of the site are present in the spatial recording: openness, greenness, presence of landmarks, dominant use (walking, staying), and social presence (related to the dominant use). These are identified as objective metrics often used in urban and landscape research [36–40], possibly contributing to soundscape assessment [23,41]. This relies on researcher’s opinion-driven assessment—it is advised to observe the location for a moment and then choose the point representative of the context and the first-person user experience. For instance, in a park, it would probably be near a bench in the central area near the fountain; in a busy square, it would be a place where most people gather and have the best view on the landmark. While doing so, the placement too near the prominent vertical objects such as a statue, a wall, or a mast should be avoided as it might cause issues in later handling the visual data (3 m is considered a safe distance from these features). Similar concerns are also true for the audio data and careful attention should be paid to avoid placing the recording equipment near extraneous noisy equipment or in acoustic shadows. Further guidance on this is given in Point 4 of Section 4.1. It is important to avoid placing the recording equipment at a position where no users are expected (i.e., don’t put the equipment in the middle of a flower bed or a grass area that nobody uses).

For the purposes of this protocol, a single location was considered to be an ‘environmental unit’ wherein the environmental factors are consistent and is typically perceived to constitute a single distinct area. The exact dimensions and delineation of the environmental unit will vary depending on the characteristics of the space, so it is ultimately up to the judgment of the researchers on site to select an appropriate measurement point to best capture the character of the environmental unit.

2.3. Equipment

The equipment listed in Table 2 is designed to facilitate both the audio-visual recording of the location and the collection of objective environmental factors, as given in Table 3. What equipment is brought on site should be adjusted depending on availability, needs of the researchers, and whether only one of the protocol stages will be carried out, or both. The equipment selected should be neutral and not noticeable. In general, this means dark or neutral colors as opposed to high-visibility colors and selecting compact equipment.

The use of class 1 or 2 sound level meters has been stipulated to maintain verifiable consistency and quality of data across all soundscape studies which make use of this protocol, as well as with data collected under various other environmental acoustics purposes. As the accuracy of acoustic information gathered at the site is the most vital in the discussion of soundscape indices, specific requirements have only been set out for the acoustic equipment. Class 1 is highly preferred, but consideration is made for cost and availability of equipment. It should be noted what standard of SLM was used in the data collection and appropriate consideration of the precision and tolerances of the equipment should be taken during the data analysis.

Table 2. Recommended equipment for implementing the SSID protocol. SLM: Sound Level Meter; AMB: Ambisonics; BIN: Binaural; QUE: Questionnaires.

| Equipment | Requirements |
|---|---|
| Tripod stand | With add-on hooks/holders for AMB microphone, SLM, environmental meter(s) and 360° camera with suitable suspension for microphones |
| 360° camera | 4 K, 5.1 K or better resolution video, with suitable battery life and optional remote control |
| Spatial audio/Ambisonics (AMB) microphone system | Min. quality should be First-order Ambisonics (FOA) capability, however systems which achieve higher-order ambisonics would be preferred where available. |
| Multi-channel field recorder | Min. inputs to accommodate output from AMB microphone |
| Windshield(s) for AMB and SLM microphones | This can be a single large windshield which can accommodate both microphones or separate windscreens for each microphone |
| Sound Level Meter (SLM) | class 1 (preferred) or class 2 with omnidirectional pattern measurement microphone |
| Binaural recording system | Portable, worn by the researcher or with a mounted binaural head |
| Sound calibrator for SLM, AMB microphones and binaural system | According to IEC 60942: 2017 Electroacoustics—Sound calibrators [42] |
| Environmental meter(s) | See Table 3 for the recommended metrics |
| Tablets and/or printed questionnaires | Internet connectivity or offline app to submit the questionnaires on site |

Table 3. Table of recommended context and acoustic measurement factors.

| Factor Category | Category Code | Factors Collected | Protocol Stage | Measurement Duration |
|---|---------------|--|---------------------|--|
| Spatial Audio | AMB | Ambisonics A format 44.1 kHz, 24 bit resolution Min. first-order ambisonics (FOA) | Recording Stage | 15 min |
| 360° Video | VID | 4K, 5.1K or better resolution video | Recording Stage | 15 minutes |
| 360° Photos | PIC | 4K, 5.1K or better resolution still photos | Questionnaire Stage | Captured with each GroupID |
| Binaural Audio | BIN | Binaural audio recording Note down the corresponding GroupID in recording metadata | Questionnaire Stage | 30 s of clean audio captured with each GroupID |
| Sound Level Meter Acoustic Data * and Audio | SLM | Acoustic data: (a) 1-second logging period (b) L_{Aeq} , L_{AFmax} , 1/3 rd Octave Band L_{Aeq} , Octave Band L_{Aeq} , Full statistics, and Full Spectral Statistics Recording: (a) .wav audio recordings (b) 44.1 kHz, 24 bit resolution | Both | Span of survey (approx. 3–4 h) |
| Environmental Data ** | ENV | 10-second logging period: (a) Temperature (°C) (b) Lighting Intensity, Lux (LI) (c) Air quality (CO ₂) (d) Relative Humidity (RH) (e) Dew Point (°C) | Both | Span of survey (approx. 3–4 h) |
| Questionnaires | QUE | SSID Questionnaire given in Appendix C Additional data: (a) GroupID for each group of participants (b) SessionID (c) Start and End time for each participant (if electronic) or each group (if paper) (d) GPS Location (if electronic) | Questionnaire Stage | On average, questionnaires last 5–10 min per GroupID |

* The recommended acoustic data settings are given here in order of importance. In cases where researchers do not have access to a meter capable spectral logging, L_{Aeq} logging should be prioritized over spectral analysis. During both stages, spectral data can typically be extracted from the audio recordings, but accurately tracking the sound level is crucial. ** The recommended environmental factors are given here in order of importance. More flexibility is allowed in selecting which factors to record and investigate (compared to the acoustic data) as it is still unclear how and to what extent environmental factors influence soundscape assessment. However, previous studies have indicated visual (i.e., lighting levels) and temperature are significant factors [43].

3. Techniques for Field Data Collection

There are several methods available for characterizing the physical environment and collecting soundscape assessments. Here, we will address the techniques employed in this protocol and general best practice for each of them.

3.1. Questionnaire Surveys

As stated above, the questionnaire is primarily based on Method A of ISO/TS 12913-2:2018. This method begins with a set of questions relating to the sound environment which are assessed on a 5-point Likert scale, coded from 1 to 5. A sample codebook to demonstrate the recommended variable naming and response coding is included in Appendix D.

The first section includes four questions relating to sound source identification, where the sound sources are divided into four categories: Traffic noise, Other noise, Sounds from human beings, and Natural sounds (labeled SSI01 through SSI04, respectively). These taxonomic categories of environmental sounds are based on the work done by Guastavino [44] and Brown, Kang, and Gjestland [45].

Next are the 8 scales which make up the circumplex model of the Swedish Soundscape Protocol [46], describing the Perceived Affective Quality (PAQ). These are assessed on a 5-point Likert scale from 'Strongly Disagree (1)' to 'Strongly Agree (5)'. These are included as follows: Pleasant, Chaotic, Vibrant, Uneventful, Calm, Annoying, Eventful, and Monotonous (labeled PAQ01 through PAQ08, respectively).

Following this are five questions addressing the participant's overall assessment of the surrounding sound environment, addressing overall acoustic quality, the appropriateness of the sound environment to the location, perceived loudness, and how often the participant visits the place and how often they would like to visit again (labeled SSS01 through SSS05, respectively).

The fourth section comprises the WHO-5 well-being index, asking how the participants have been feeling over the last two weeks, such as 'I have felt calm and relaxed'. The WHO-5 index is constructed to constitute an integrated scale in which the items add up related information about the level of the individual's general psychological well-being [47,48]. This information can provide additional insight into how exposure to pleasant or annoying soundscapes may impact psychological well-being as was investigated by Aletta et al. [27] or, alternatively, how a person's current psychological status may influence their perception of the sound environment as recently investigated by Erfanian, Mitchell, Aletta, and Kang [49]. Each of the five WHO questions (labeled WHO01 to WHO05) are assessed on a 6-point scale coded from 0 to 5.

The final section of the participant-facing questionnaire comprises five questions on the participant's demographic information (age [AGE00], gender [GEN00], occupational status [OCC00], education level [EDU00], ethnicity [ETH00], and local vs. tourist [MISC03]) and a free response for the participant to provide any additional comments they would like to make on the sound environment [MISC01]. It is important to note that the section on ethnicity, and to a lesser extent education level, will need to be adjusted to ensure the available responses are appropriate for the location where the survey is being conducted.

At the end of the questionnaire are a set of spaces available for the researcher conducting the survey to fill out, adding additional information about the observed behavior of the participants, indexing and labeling metadata, and space for any additional notes. More information and guidance on this information is included below.

This questionnaire is intended to collect a consistent core set of perceptual responses and information about the participant, with space to add additional questions as required by specific research goals. Some examples of this which have been implemented by the various research groups are specific questions calling attention to water sounds and features, the perception of visual features, and an open response for identifying the dominant sound source. Given the proper labeling and coding, these additional questions can be fully integrated into the overall dataset, allowing the researchers the freedom to pursue their own research interests while maintaining consistency and compatibility with the overall database.

General notes for conducting the questionnaires:

- The core questionnaire is reported in Appendix C. The labels and corresponding scales are also reported. Ideally, the form should be submitted and filled on a tablet via a survey app (e.g., REDCap, Qualtrics, KoBoToolbox, or similar) so that data can then be easily downloaded in an .xlsx or .csv file. Using paper forms is also acceptable; however, researchers on site will need to take more careful note of information such as the time of response and the information will need to be manually input after the session is completed. If using an electronic version, the system should be set up to record the start and end times and GPS coordinates for each survey.

- If using an electronic version, be sure to have enough tablets with internet connectivity (if required by the survey system) and sufficient battery life; if using the paper version, be sure to print enough copies. Even if using the electronic version, it is recommended to also print a number of paper versions as a backup or if a large group agrees to participate at once.
- Regardless of the translation of the items, it is important that the label (e.g., SSI01) is kept, as well as the size and direction of the scales (1–5, etc.) to maintain data consistency.

3.2. Contextual and Environmental Factor Data Collection

During each survey, the equipment listed in Section 2.3 is set up to capture the contextual and environmental data for the location. Table 3 lists the factors to be collected and at what stage they should be collected.

3.2.1. Spatial Audio-Visual Recordings

In order to capture the acoustic and visual information in the space for replication in a laboratory setting, 360° video and AMB audio are recorded to be used in Virtual Reality (VR) playback. The goal of this is two-fold: first, to enable researchers to document and replicate the in situ environment of the space as it was during a questionnaire survey session for lab experiments and, second, to capture environments in which performing a questionnaire survey is not feasible.

Typically, questionnaire surveys are carried out over a period of several days at the same location. The goal of these multiple sessions is to capture as many questionnaire responses as needed (100 for a particular soundscape is typically recommended [11]), which, in the experience of the authors is prohibitively difficult to achieve in a single session in most locations. It is recommended that the repeated sessions are conducted under similar circumstances and environmental conditions. As such, it is not entirely necessary to repeat the spatial recordings each time a questionnaire survey is conducted. Instead, it is useful to use the spatial recording as a chance to gain a different perspective on the space under investigation. For instance, if the questionnaires are conducted in the middle of a large urban park, the first session could collect a spatial recording within the environmental unit of the questionnaire site, but the subsequent returns to the site could collect spatial recordings in a different environmental unit, say, along a road bounding the park, or in a space in the park which does not typically have many people. This enables the simultaneous expansion of the questionnaire database and the gathering of additional environments to investigate in a laboratory setting.

General notes for spatial recordings:

- The audio-video recordings can be done before or after the questionnaire survey.
- The purpose of the audio-video recordings is to capture representative recordings which can be reproduced in a laboratory setting. During the first time at a location, the focus should be on capturing the environment as experienced by the respondents to the questionnaires at that location. Therefore, the recordings should be performed in nearly the same spot, with similar lighting and environmental conditions. For further survey sessions, provided the conditions are similar, other recordings could be taken which provide additional perspectives around the space for reproducing in the lab.
- These recordings can be performed entirely separately from the questionnaire survey, if desired. Reasons for doing this may be (but are not limited to): location is not populated, making questionnaires impossible; specific locations or conditions are required for a lab experiment; time limitations require many sites in an area to be captured and in situ questionnaires could not be completed in time.
- The 360° video will take a significant amount of storage space. Researchers should ensure that there is ample free space on the camera SD cards prior to going out on site. If conducting multiple surveys away from their home institution (i.e., in another city), teams are recommended to bring a large external hard drive so that videos can be offloaded after each session.

3.2.2. Reference Recordings

A soundscape index, or any investigation of the impact of the physical environment on the soundscape, requires consistent and accurate measurement of the environment, most importantly calibrated measurement and recording of the acoustic environment. For this protocol, this has been achieved through the use of separate calibrated binaural recordings and measurements made with a calibrated sound level meter (SLM).

4. Procedure

Figure 1 shows the whole process of the on site soundscape protocol. The relevant equipment in each row should be operating when the row is colored in, such that when multiple rows are shaded this means that multiple pieces of equipment should be running during that time period. The following section prepares step-by-step instructions for conducting the in situ surveys, including the Recording Stage and Questionnaire Stage. Figure 2 shows an example of the recommended equipment setup.

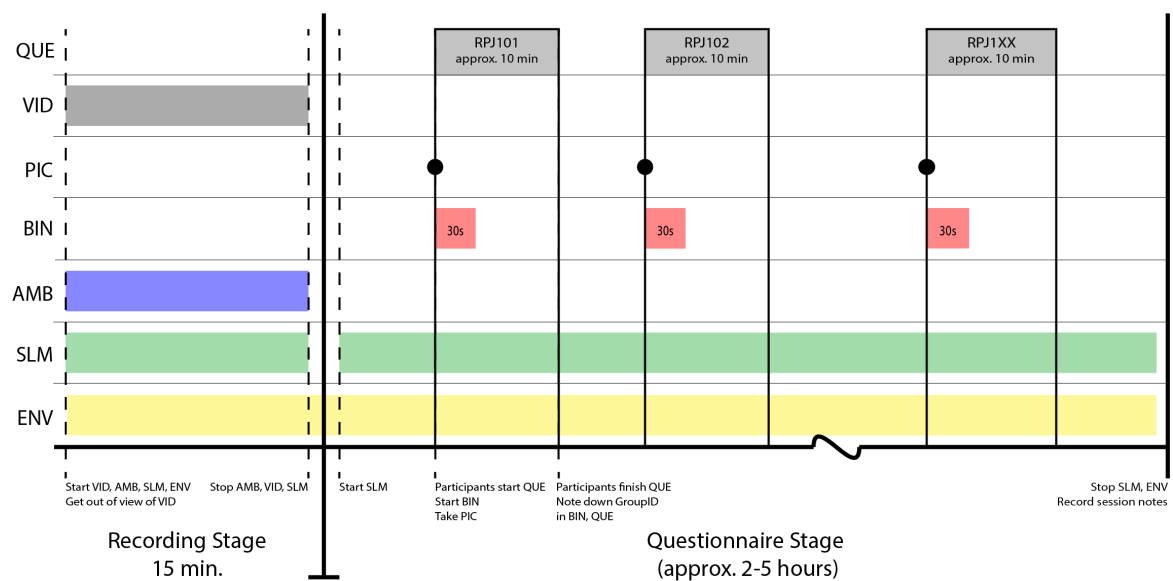


Figure 1. Timeline of the on site soundscape protocol. RegentsParkJapan (RPJ) is used as an example. Abbreviations as defined in Table 3—QUE: Questionnaires; VID: 360° video; PIC: Site pictures; BIN: Binaural Recording; AMB: Ambisonic recording; SLM: Sound Level Meter (acoustical factors); ENV: Environmental factors.

The equipment should be assembled, checked, and calibrated prior to arriving at the measurement location. Calibrate the equipment according to the manufacturer’s instructions. All sound level meters should have built-in methods to calibrate using a standard 94 dB 1 kHz tone calibrator. If a similar method is available for the ambisonic microphone, this should be used. If a built-in method is not available, but a calibrator can be fitted to the microphone capsules, then the ambisonic microphone should be calibrated by recording the 1 kHz signal through the system for each microphone capsule after the gain settings have been finalized on site (see below). If it is not possible to calibrate the ambisonic microphone, then the levels recorded will need to be compared to the levels taken simultaneously with the SLM. This is why it is crucial to have an appropriate quality, calibrated SLM included within the same setup as the AMB recordings.

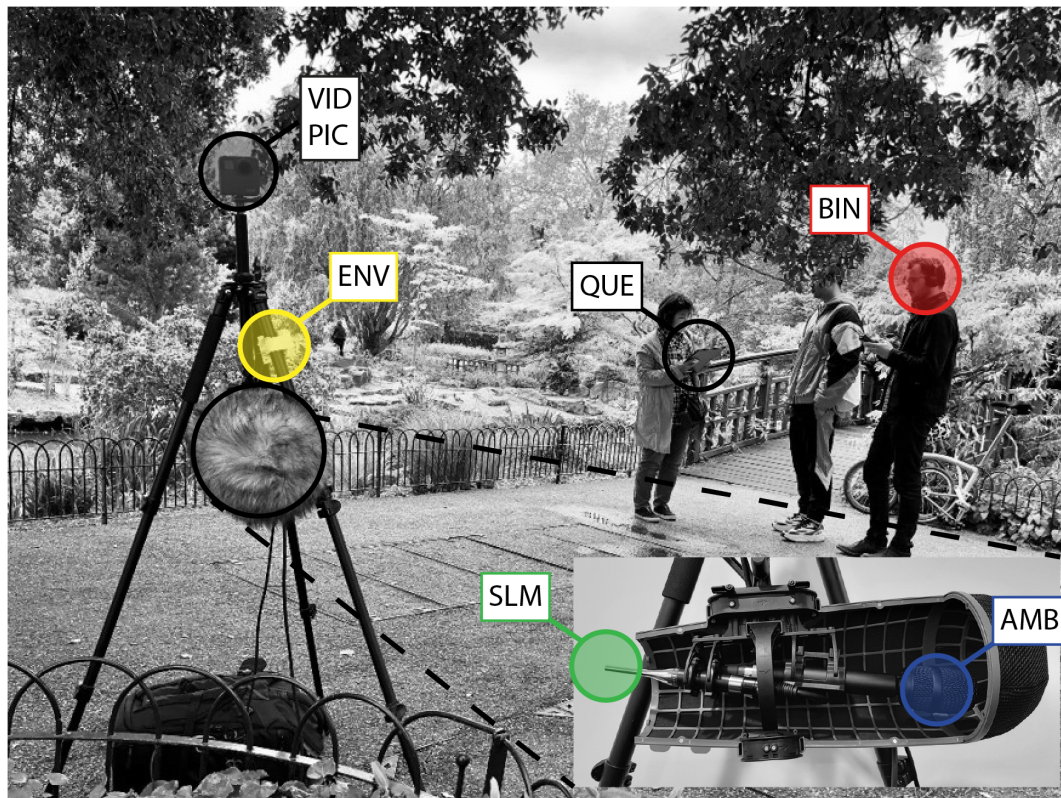


Figure 2. Photo of a full survey carried out in a park in London during the Questionnaire Stage. To the left is the equipment (color-coded to match Figure 1), with the ambisonic microphone and SLM microphone in the windscreen, with the 360° camera on top of the tripod and to the right are one researcher interacting with the participant while the second researcher conducts the binaural recording. The body of the SLM and the multi-channel recorder are stored in a bag under the tripod which can contain all of the pieces of equipment for easy transport.

4.1. Assembling the Equipment

1. Set up the equipment by prioritizing the position of the 360° camera and position the lens at the average eye level 160–180 cm, as shown in Figure 2.

It is advisable to test the setup for video stitching issues and reconfigure if needed (e.g., the equipment will be partially visible in the raw video recording, so you need to test if the chosen setup allows for efficient erasing/hiding/patching of the exposed parts in the post-processing). Companies selling 360° cameras usually offer free software for basic editing and previewing. It is advisable to position the camera as the highest item in the set to avoid the need for editing both the sky and the ground.

2. Carefully position the AMB microphone so its axes are aligned with the axes of the 360° camera; the microphone's front (usually marked by the logo) and the camera's front should be looking in the same direction. Many AMB microphones allow them to be oriented vertically or horizontally (end-fire), this should be noted and adjusted in the relevant software settings.

This is essential for informed post-processing. It is advisable to position the capsules of the AMB microphone and the capsule of the SLM as near to each other as possible, without introducing scattering effects. It can usually be done within the same windshield unit, but it is not essential to do so and depends on the available clamps and stands.

3. The gain settings for the four ambisonic audio channels should be set to the same level. In some devices (such as the MixPre 10), this can be set by locking the channel gain settings to a

single channel. Many devices also offer ambisonic plugins which simplify these settings and automatically link the gain settings—these should be used where available.

4. Set the SLM to log sound levels and simultaneously record .wav audio. The recommended logging settings are given in Table 3. The SLM should be mounted and positioned according to standard guidance for environmental noise measurements, like that given in Section 9 of ISO 1996-2:2017 [19] or Section 5 of ANSI/ASA S12.9-2013/Part 1 [50]. Generally, the microphone should be a minimum of 1.2 m above the ground and a minimum of 1 m from any vertical reflecting surfaces.
5. Attach the environmental meter(s) to the tripod. Care should be taken when positioning the environmental monitor. Most units will include guidance on their use from the manufacturer—these should be followed where available. Some general items to keep in mind include not accidentally covering air quality sensor holes, not positioning light sensors in the shade of the other equipment, and not positioning temperature sensors in direct sunlight unless this is how they are intended to be positioned.

4.2. Recording Stage

The following section prepares step-by-step instructions for conducting the Recording Stage of the on site protocol, as shown in Figure 1.

1. Double check all settings and file save locations on the recording equipment.
2. Adjust gain settings to ensure there is no clipping. Good practice is to listen for what is expected to be the loudest sound event during the recording period (e.g., sirens) and set the gain such that the level is comfortably under clipping during this event.
3. Start recording on all devices, including the ambisonic microphone, 360° camera, SLM, and environmental meter.
4. Stand at the front of the camera/ambisonic microphone and clap. The clap can help synchronize the audio with the video, if necessary, and ensuring you are standing in line with the front of the 360° video can help with lining up the directionality of the two, if necessary.
5. Retreat out of view of the camera, blending into the surrounding crowd, or otherwise make sure not to be obvious to someone watching the video.
6. Record at least 5 min of consistent and representative audio and video. It is recommended to record for 15 min to give the best chance of being able to extract a solid 5 min of useful video and audio.
7. Stop recording on all devices and ensure all files are saved properly.

4.3. Questionnaire Stage

The following section prepares step-by-step instructions for conducting the in situ questionnaires and their accompanying reference recordings as part of the Questionnaire Stage. Typically these are performed during the same working session as the Recording Stage, using the same set of equipment. The selection of an appropriate location and setup of the equipment should follow the guidance given in Section 2.2, while making sure the location selected is representative of where the respondents will be stopped. Wherever possible, the equipment should be assembled and located so as not to draw the attention of the respondents and particularly to avoid influencing their perception of the space.

1. Double check all settings and file save locations on the recording equipment. If starting this stage immediately after the Recording Stage, make sure to rename or advance the index on the filenames for the SLM and environmental meters.
2. Start recording on the SLM and environmental meter (or leave running from preceding Recording Stage). These will continue running until the end of the Questionnaire Stage.
3. Gather the tablets and/or paper questionnaires and prepare to approach potential participants.

4. Approach participants and ask if they would be willing to take part in a research study. If the participants are in a group, they can participate at the same time, but should each fill out a separate questionnaire. When approaching participants, you should identify yourself as a researcher or student researching urban sound. We advise avoiding phrases such as “noise”, “noise pollution”, “noise disturbance” or other terms which carry a negative connotation. In general, explanations and answers to questions should strive to be as neutral as possible regarding the nature of the soundscape.
5. Once the participant has consented to participate, hand them the questionnaire or tablet and provide them with basic instructions for answering the questionnaire. Emphasize that they should be responding and assessing the current sound environment, in the current place. Note that this is a common misunderstanding—many participants assume the questionnaire is focused on the sound environment at their home, or in the city in general. Where a mix of tablets and paper questionnaires are being used, each group should have at least one participant using a tablet such that start and end times and precise GPS coordinates can be pulled from the accompanying electronic questionnaire. While one researcher is interacting with the participants, the second should arrange the equipment for taking the binaural recordings and 360° photo.
6. Once the participant has started answering the questionnaire, start recording the binaural audio. If the participants are in a group and all are taking the survey at the same time, only one binaural recording is needed for the whole group. The researcher conducting the recording should strive to keep their head as stationary as possible and to avoid making any extraneous noise.

Make sure that at least 30 s of consistent audio is recorded while the participant is filling in the questionnaire. This should not include talking either from the researcher or the participant. If talking or other intrusive (non-representative) sound occurs, extend the recording period to end up with a solid 30 s of good audio. The goal is to capture the sound environment which the participant was exposed to while filling out their questionnaire, but to exclude sounds which the participant is not likely considering as part of their assessment. Most commonly, this would be the researcher talking, or the participant themselves talking. Any other sounds which the participant was “naturally” exposed to should be included.

When taking the binaural recording, attempt to orient the head (artificial or researcher wearing a headset) in the same direction as the participants. This is not crucial as it is often impossible to achieve, but it is preferable. Be careful not to move the head during the recording.

7. Note the GroupID in the metadata for the binaural recording, or make a manual note of the binaural recording file name and the GroupID separately.
8. Take one 360° photo with the camera to capture the general setting. This can also be done at regular intervals during the survey session.
9. When the participant has finished filling in the questionnaire, thank them for their participation and fill in the additional researcher questions at the end of the questionnaire. These help to both track the data collected and to document the conditions on site. The most important of these are:
 - (For paper versions) Start and End time. If a Start time was not noted, at minimum, the End Time must be recorded and an average survey duration can be subtracted to estimate the Start Time
 - GroupID
 - SessionID
10. Repeat steps 4–9 for the remainder of the session, incrementing the GroupID by one with each new group of participants. If there are more than two researchers on site, the additional researchers can stop new groups of participants simultaneously. The researcher operating the binaural equipment can then shift between the groups once they have finished the 30 s recording. This researcher should also have the responsibility of keeping track of the GroupID numbers for each group.

Experience has shown this is possible up to about three groups at a time, with four researchers on site.

11. Once the session is finished, stop the equipment and ensure all files are saved properly.
12. After each session, make note of the character of the site and the environmental conditions during the survey. This might include, but is not limited to:
 - Site typology and intended use (e.g., urban park, transit station, urban square, etc.)
 - Weather
 - Crowdedness (i.e., how many people are present in the space)
 - Dominant sound sources and any key soundmarks
 - Visual character (e.g., amount of greenness, enclosed vs. open, etc.)

5. Lessons from International Data Collection

As this protocol has already been implemented by several research groups across four countries, it has undergone a rigorous testing and development process. Throughout this process, adjustments have been made which resulted in the final protocol presented here. However, no process is perfect or applicable in all situations. As such, after consultation with the research groups involved, we have compiled the most common feedback and guidance to keep in mind when implementing this protocol.

5.1. Sampling

The research groups were instructed to try keeping the structure of respondents well-balanced. This often led to longer times and larger sample sizes required as most comments from five research groups addressed age and type of location as the most influential factors for participant sampling. However, while some reported higher response rates from younger (students) members of public, the others reported higher response rates in case of older high educated people. A common observation was that public parks are the locations with the highest response rates, most probably due to a high number of people taking part in activities that allow enough time to take part in a survey. The type of space was also reflected in the sense of privacy. In locations that were more public, people in groups were more likely to take part in the survey, while in the more private locations it was the opposite. Amongst other comments, whether a participant was a tourist or a local also had an influence on the response rate. Tourists seemed more likely to participate in the survey.

Several groups reported excessive heat and cold to be negatively affecting the response rates. One research group, which conducted the survey also in a residential area, distinguished privacy/ownership of the survey site as a major factor.

5.2. Data Collection

A group of three researchers seems to be the minimum number needed to conduct the survey, as observed by the partner research groups. The group of nine researchers on-site proved to be the most effective number. The time needed to complete the survey varied greatly depending on the location.

Although the questions are written in a manner that emphasizes the focus on the actual acoustic environment perceived at the moment, additional care should be made to ensure the proper understanding of that concept while approaching the participants. Researcher's comments are invaluable here to keep track of the outliers if a researcher feels similar issues or other factors (i.e., wearing headphones) lead to collecting invalid/misleading data.

5.3. Equipment

Some partners had previous experience in soundscape research, but for all this was the first study that featured surveying large number of public participants around a single measurement point. All the

research groups found it very important to delegate one researcher/technician to care exclusively about the equipment and the quality of the recordings.

The intention of the recording stage is to record a first-person experience most representative of the location. Therefore, the researchers are instructed to ‘make themselves invisible’ in the recording. However, at some locations, various research groups decided to put out a sign asking members of public not to touch or come near the measurement point as they experienced passers-by touching the windshield out of curiosity.

The equipment setup has been designed to be as compact and unobtrusive as possible so as to limit any intrusion on the participant’s experience of the space. From our experience, most participants do not end up with the equipment within their field of view during the questionnaire and often do not notice the presence of the stationary equipment. In some locations, this is not possible and participants may comment on its presence; however, over the thousands of surveys collected, only a small number of respondents have commented on the equipment as noticeably impacting their experience.

5.4. Translation

Regarding the on-site soundscape survey, the translation of the questionnaires (and in particular the perceptual adjectives used for the soundscape appraisal) is a key point to consider when using the protocol in regions where English is not the local language. Indeed, while the ISO/TS 12913-2:2018 document from which the soundscape-related questions of this protocol are derived aims at providing standardized scales, it does not provide official translations in languages other than English. Some perceptual constructs are difficult to render in different languages and people might assign different meanings to them (e.g. [51–54]). For this reason, in the soundscape research community, there is a growing interest in testing and validating reliable translation of the ISO soundscape adjectives [24], which will hopefully lead to a wide-spread use of this soundscape tool. It is expected that these validated translations could simply be substituted for their English counterparts in this protocol, when they become available.

Supplementary Materials: Please find the link to the Soundscape Indices (SSID) ERC-funded website at: <https://www.ucl.ac.uk/bartlett/environmental-design/soundscape-indices-ssid>.

Author Contributions: Conceptualization, A.M., T.O., F.A., and J.K.; methodology, A.M., T.O., F.A., M.E., M.L., and J.K.; validation, A.M., T.O., M.E., M.K., and M.L.; investigation, A.M., T.O., F.A., M.E., M.K., and M.L.; resources, T.O., F.A., and J.K.; data curation, A.M. and T.O.; writing—original draft preparation, A.M.; writing—review and editing, A.M., T.O., F.A., and J.K.; visualization, A.M.; supervision, J.K.; project administration, F.A. and J.K.; funding acquisition, J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant Agreement No. 740696, project title: Soundscape Indices—SSID). More information and related publications can be found at the CORDIS webpage of the project: <https://cordis.europa.eu/project/rcn/211802/factsheet/en>.

Acknowledgments: The authors would like to thank our partner research groups, particularly those in Universidad de Granada, Harbin Institute of Technology, Harbin Institute of Technology (Shenzhen), Shenyang Jianzhu University, Tianjin University, Northwest A&F University, Huazhong University of Science and Technology, and Chongqing University for their valuable feedback and for their data collection efforts. We would also like to thank the Masters students, visiting researchers, and other PhD students at UCL who have helped contribute to the refinement of this protocol and collection of the current database. Study data were collected and managed using REDCap electronic data capture tools hosted at University College London (UCL) [55,56]. REDCap (Research Electronic Data Capture) is a web-based software platform designed to support data capture for research studies.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SSID Soundscape Indices
FOA First Order Ambisonics

| | |
|-----|-----------------------------|
| PAQ | Perceived Affective Quality |
| ERC | European Research Council |
| VR | Virtual Reality |
| QUE | Questionnaires |
| VID | 360° Video |
| PIC | Site Pictures |
| BIN | Binaural Recording |
| AMB | Ambisonic Video |
| SLM | Sound Level Meter |
| ENV | Environmental Data |

Appendix A. Sample Information Sheet

PARTICIPANT INFORMATION SHEET: PERCEIVED SOUNDSCAPE

You are being invited to take part in a research project. Before you decide it is important for you to understand why the research is being done and what participation will involve. Please take your time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part. Thank you for reading this.

I am XXXXX YYYYY and I am a student at ZZZZZ University. I am working on a project about “soundscapes” and the aim of this questionnaire is collecting data on how people perceive urban acoustic environments, and what are the relationship between acoustic environment and well-being. Results from this survey will help us to gather further insights into these relationships and the knowledge we’ll gain from this research should inform urban sound planners. Other participants will be randomly approached on site and will be invited to take part, like you. You can participate to this survey only if you are over eighteen years old.

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep. Please observe that participating in a scientific experiment is voluntary and you are free to withdraw at any time without giving any reason. Should you choose to withdraw before completing the questionnaire, your partial responses will be deleted. However, once your response is submitted, all data are completely anonymised, and it will not be possible for us to identify or remove your response. The questionnaire is designed to take approximately 10 minutes to complete, and it will allow you to evaluate the surrounding acoustic environment in the public area where you are now, related with your well-being. Whilst there are no immediate benefits for those people participating in the project, it is hoped that this work will raise environmental awareness for the soundscapes of our cities.

This study has been approved via the ethics review procedure of the Bartlett School of Environment, Energy and Resources. If you wish to raise a complaint you should contact the principal researcher of this project, Prof Jian Kang at the e-mail kang@ucl.ac.uk or telephone number 020 3108 7338. If you feel your complaint has not been handled to your satisfaction, then you can contact the Chair of the UCL Research Ethics Committee (ethics@ucl.ac.uk).

All the information that we collect about you during the course of the research will be kept strictly confidential. You will not be able to be identified in any ensuing reports or publications. All the data collected will be stored in on UCL secure servers with password protection and at the end of the project all data collected will be securely archived.

This project is funded through a European Research Council (ERC) Advanced Grant (no. 740696) on “Soundscape Indices” (SSID) (Principal Investigator: Prof Jian Kang).

If you experience any problems or need further information, contact me on email: uclssid@gmail.com

Thank you in advance for your time and participation!

Appendix B. Sample Consent Form

2019 London Questionnaire

Page 1 of 7

Please complete the survey below.

Thank you!

PARTICIPANT CONSENT FORM

Title of Research Project: Soundscape

Name of Researchers: Mercede Erfanian, Magdalena Kachlicka, Matteo Lionello & Andrew Mitchell

By clicking this, I confirm that:

1. I have read and understood the information sheet explaining the experiment and I have had the opportunity to ask questions about the project.
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without there being any negative consequences. In addition, should I not wish to answer any particular question, I am free to decline.
3. I understand that my responses will be kept strictly confidential.
4. I give permission for members of the research team to have access to my anonymised responses and I understand that I will not be identifiable in reports that result from the research.
5. I agree to take part in the above-mentioned project.
6. I am over 18 years old and under 80 years old.

- Yes, I agree
 No, I disagree
-

Thank you in advance for your time and participation!
This questionnaire should take approximately 3 minutes.

Appendix C. SSID Questionnaire

Page 2 of 7

| To what extent do you presently hear the following four types of sounds? | | | | | |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | Not at all | A little | Moderately | A lot | Dominates completely |
| Traffic noise (e.g. cars, buses, trains, airplanes) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Other noise (e.g. sirens, construction, industry, loading of goods) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Sounds from human beings (e.g. conversation, laughter, children at play, footsteps) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Natural sounds (e.g. singing birds, flowing water, wind in vegetation) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

| For each of the 8 scales below, to what extent do you agree or disagree that the present surrounding sound environment is... | | | | | |
|---|-----------------------|-----------------------|----------------------------|-----------------------|-----------------------|
| | Strongly agree | Somewhat agree | Neither agree nor disagree | Somewhat disagree | Strongly disagree |
| Pleasant | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Chaotic | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Vibrant | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Uneventful | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Calm | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Annoying | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Eventful | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Monotonous | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Overall, how would you describe the present surrounding sound environment?

- Very good
- Good
- Neither bad nor good
- Bad
- Very bad

Overall, to what extent is the present surrounding sound environment appropriate to the present place?

- Not at all
- Slightly
- Moderately
- Very
- Perfectly

How loud would you say the sound environment is?

- Not at all
- Slightly
- Moderately
- Very
- Extremely

How often do you visit this place?

- Never / This is my first time here
- Rarely
- Sometimes
- Often
- Very often

How often would you like to visit this place again?

- Never
- Rarely
- Sometimes
- Often
- Very often

Please indicate for each of the five statements below which is closest to how you have been feeling over the last two weeks.

| | All of the time | Most of the time | More than half of the time | Less than half of the time | Some of the time | At no time |
|--|-----------------------|-----------------------|----------------------------|----------------------------|-----------------------|-----------------------|
| I have felt cheerful and in good spirits | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| I have felt calm and relaxed | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| I have felt active and vigorous | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| I woke up feeling fresh and rested | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| My daily life has been filled with things that interest me | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

How old are you?

What is your gender? (Optional)

- Male
- Female
- Non-conforming
- Rather not say

What is your occupational status?

- Employed
- Unemployed
- Retired
- Student
- Other
- Rather not say

Please specify "Other":

What is the highest level of education you have completed?

- Some high school
- High school graduate
- Some college
- Trade/technical/vocational training
- University graduate
- Some postgraduate work
- Postgraduate degree

Please specify your ethnicity.

- White
- Mixed/Multiple ethnic groups
- Asian/Asian British
- Black/African/Caribbean/Black British
- Middle Eastern
- Rather not say
- Other ethnic group

Please specify "Other ethnic group":

Is there anything else you want to let us know about the sound environment? (Optional)

Would you consider yourself...

- A local (~15 km)
- A tourist
- Other

Please specify "Other":

Thank you for your participation, please hand the tablet back to the researcher.

Filled by the Researcher:

LocationID

Group ID

Was the participant...

- Staying
- Arriving
- Leaving
- Passing through

Was the participant...

- Alone
- In couple
- In a group of 3 or more

Recordings taken?

- Continuous sound level
- Binaural recording
- 360 photo
- Spatial audio
- 360 Video

Any other notes?

Was this a test?

- Yes
- No

Is this a paper input?

- Yes
- No

Appendix D. Questionnaire Codebook

| SSID Questionnaire Codebook | | | |
|-----------------------------|--|-------------------|------------------------------------|
| Variable / Field Name | Displayed Question | Response code | Response text |
| ssi01 | Traffic noise (e.g. cars, buses, trains, airplanes) | 1 | Not at all |
| ssi02 | Other noise (e.g. sirens, construction, industry, loading of goods) | 2 | A little |
| ssi03 | Sounds from human beings (e.g. conversation, laughter, children at play, footsteps) | 3 | Moderately |
| ssi04 | Natural sounds (e.g. singing birds, flowing water, wind in vegetation) | 4 | A lot |
| | | 5 | Dominates completely |
| paq01 | Pleasant | 5 | Strongly agree |
| paq02 | Chaotic | 4 | Somewhat agree |
| paq03 | Vibrant | 3 | Neither agree nor disagree |
| paq04 | Uneventful | 2 | Somewhat disagree |
| paq05 | Calm | 1 | Strongly disagree |
| paq06 | Annoying | | |
| paq07 | Eventful | | |
| paq08 | Monotonous | | |
| sss01 | Overall, how would you describe the present surrounding sound environment? | 5 | Very good |
| | | 4 | Good |
| | | 3 | Neither bad nor good |
| | | 2 | Bad |
| | | 1 | Very bad |
| sss02 | Overall, to what extent is the present surrounding sound environment appropriate to the present place? | 1 | Not at all |
| | | 2 | Slightly |
| | | 3 | Moderately |
| | | 4 | Very |
| | | 5 | Perfectly |
| sss03 | How loud would you say the sound environment is? | 1 | Not at all |
| | | 2 | Slightly |
| | | 3 | Moderately |
| | | 4 | Very |
| | | 5 | Extremely |
| sss04 | How often do you visit this place? | 1 | Never / This is my first time here |
| | | 2 | Rarely |
| | | 3 | Sometimes |
| | | 4 | Often |
| | | 5 | Very often |
| sss05 | How often would you like to visit this place again? | 1 | Never |
| | | 2 | Rarely |
| | | 3 | Sometimes |
| | | 4 | Often |
| | | 5 | Very often |
| who01 | I have felt cheerful and in good spirits | 5 | All of the time |
| who02 | I have felt calm and relaxed | 4 | Most of the time |
| who03 | I have felt active and vigorous | 3 | More than half of the time |
| who04 | I woke up feeling fresh and rested | 2 | Less than half of the time |
| who05 | My daily life has been filled with things that interest me | 1 | Some of the time |
| | | 0 | At no time |
| age00 | How old are you? | (integer, Min:18) | |
| gen00 | What is your gender? | 1 | Male |
| | | 2 | Female |
| | | 3 | Non-conforming |
| | | 4 | Rather not say |

| | | | |
|-----------|---|--------|---|
| occ00 | What is your occupational status? | 1 | Employed |
| | | 2 | Unemployed |
| | | 3 | Retired |
| | | 4 | Student |
| | | 5 | Other |
| | | 6 | Rather not say |
| edu00 | What is the highest level of education you have completed? | 1 | Some high school |
| | | 2 | High school graduate |
| | | 3 | Some college |
| | | 4 | Trade / Technical / vocational training |
| | | 5 | University graduate |
| | | 6 | Some postgraduate work |
| | | 7 | Postgraduate degree |
| eth00 | Please specify your ethnicity. | 1 | White |
| | | 2 | Mixed / multiple ethnic groups |
| | | 3 | Asian / Asian British |
| | | 4 | Black / African / Caribbean / Black British |
| | | 5 | Middle Eastern |
| | | 6 | Rather not say |
| | | 7 | Other ethnic group |
| misc00 | Would you consider yourself... | 1 | A local |
| | | 2 | A tourist |
| | | 3 | Other |
| misc01 | Is there anything else you want to let us know about the sound environment? | (text) | |
| sessionid | SessionID | (text) | |
| groupid | GroupID | (text) | |
| use00 | Was the participant... | 1 | Staying |
| | | 2 | Arriving |
| | | 3 | Leaving |
| | | 4 | Passing through |
| res00 | text | (text) | |

References

- Berglund, B.; Nilsson, M.E. On a tool for measuring soundscape quality in urban residential areas. *Acta Acust. United Acust.* **2006**, *92*, 938–944.
- Aletta, F.; Kang, J.; Axelsson, Ö. Soundscape descriptors and a conceptual framework for developing predictive soundscape models. *Landsc. Urban Plan.* **2016**, *149*, 65–74. [CrossRef]
- Kang, J.; Aletta, F.; Oberman, T.; Erfanian, M.; Kachlicka, M.; Lionello, M.; Mitchell, A. Towards soundscape indices. In Proceedings of the 23rd International Congress on Acoustics, Aachen, Germany, 9–13 September 2019; pp. 2488–2495.
- International Organization for Standardization. *ISO 12913-1:2014 Acoustics—Soundscape Part 1: Definition and Conceptual Framework*; Standard ISO 12913-1:2014; International Organization for Standardization: Geneva, Switzerland, 2014.
- International Organization for Standardization. *ISO/TS 12913-2:2018 Acoustics—Soundscape Part 2: Data collection and Reporting Requirements*; Standard ISO/TS 12913-2:2018; International Organization for Standardization: Geneva, Switzerland, 2018.
- International Organization for Standardization. *ISO/TS 12913-3:2019 Acoustics—Soundscape Part 3: Data Analysis*; Standard ISO/TS 12913-3:2019; International Organization for Standardization: Geneva, Switzerland, 2019.
- Sun, K.; De Coensel, B.; Filipan, K.; Aletta, F.; Van Renterghem, T.; De Pessemer, T.; Joseph, W.; Botteldooren, D. Classification of soundscapes of urban public open spaces. *Landsc. Urban Plan.* **2019**, *189*, 139–155. [CrossRef]
- Oberman, T.; Šćitaroci, B.B.O.; Jambrošić, K. Towards a virtual soundwalk. In *Handbook of Research on Perception-Driven Approaches to Urban Assessment and Design*; IGI Global: Hershey, PA, USA, 2018; pp. 317–343.
- Craig, A.; Moore, D.; Knox, D. Experience sampling: Assessing urban soundscapes using in-situ participatory methods. *Appl. Acoust.* **2017**, *117*, 227–235. [CrossRef]
- Lavia, L.; Witchel, H.J.; Aletta, F.; Steffens, J.; Fiebig, A.; Kang, J.; Howes, C.; Healey, P.G. Non-participant observation methods for soundscape design and urban planning. In *Handbook of Research on Perception-Driven Approaches to Urban Assessment and Design*; IGI Global: Hershey, PA, USA, 2018; pp. 73–99.

11. Engel, M.S.; Fiebig, A.; Pfaffenbach, C.; Fels, J. A Review of Socio-acoustic Surveys for Soundscape Studies. *Curr. Pollut. Rep.* **2018**, *4*, 220–239. [[CrossRef](#)]
12. Yang, W.; Kang, J. Acoustic comfort evaluation in urban open public spaces. *Appl. Acoust.* **2005**, *66*, 211–229. [[CrossRef](#)]
13. Rychtáriková, M.; Vermeir, G. Soundscape categorization on the basis of objective acoustical parameters. *Appl. Acoust.* **2013**, *74*, 240–247. [[CrossRef](#)]
14. Aumond, P.; Can, A.; De Coensel, B.; Botteldooren, D.; Ribeiro, C.; Lavandier, C. Modeling soundscape pleasantness using perceptual assessments and acoustic measurements along paths in urban context. *Acta Acust. United Acust.* **2017**, *103*, 430–443. [[CrossRef](#)]
15. Alsina-Pagès, R.M.; Freixes, M.; Orga, F.; Foraster, M.; Labairu-Trenchs, A. Perceptual evaluation of the citizen's acoustic environment from classic noise monitoring. *Cities Health* **2020**, 1–5. [[CrossRef](#)]
16. Ricciardi, P.; Delaitre, P.; Lavandier, C.; Torchia, F.; Aumond, P. Sound quality indicators for urban places in Paris cross-validated by Milan data. *J. Acoust. Soc. Am.* **2015**, *138*, 2337–2348. [[CrossRef](#)]
17. Zhang, X.; Ba, M.; Kang, J.; Meng, Q. Effect of soundscape dimensions on acoustic comfort in urban open public spaces. *Appl. Acoust.* **2018**, *133*, 73–81. [[CrossRef](#)]
18. International Organization for Standardization. *ISO 1996-1:2016 Acoustics—Description, Measurement and Assessment of Environmental Noise—Part 1: Basic Quantities and Assessment Procedures*; Standard ISO 1996-1:2016; International Organization for Standardization: Geneva, Switzerland, 2016.
19. International Organization for Standardization. *ISO 1996-2:2017 Acoustics—Description, Measurement and Assessment of Environmental Noise—Part 2: Determination of Sound Pressure Levels*; Standard ISO 1996-2:2017; International Organization for Standardization: Geneva, Switzerland, 2017.
20. Ren, X.; Kang, J.; Zhu, P.; Wang, S. Soundscape expectations of rural tourism: A comparison between Chinese and English potential tourists. *J. Acoust. Soc. Am.* **2018**, *143*, 373–377. [[CrossRef](#)] [[PubMed](#)]
21. Kang, J.; Aletta, F.; Gjestland, T.T.; Brown, A.L.; Botteldooren, D.; Schulte-Fortkamp, B.; Lercher, P.; van Kamp, I.; Genuit, K.; Fiebig, A.; et al. Ten questions on the soundscapes of the built environment. *Build. Environ.* **2016**, *108*, 284–294. [[CrossRef](#)]
22. De Coensel, B.; Sun, K.; Botteldooren, D. Urban Soundscapes of the World: Selection and reproduction of urban acoustic environments with soundscape in mind. In Proceedings of the 46th International Congress and Exposition on Noise Control Engineering (Inter-Noise) 2017, Hong Kong, China, 27–30 August 2017; p. 7.
23. Aletta, F.; Kang, J. Towards an urban vibrancy model: A soundscape approach. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1712. [[CrossRef](#)] [[PubMed](#)]
24. Aletta, F.; Guattari, C.; Evangelisti, L.; Asdrubali, F.; Oberman, T.; Kang, J. Exploring the compatibility of “Method A” and “Method B” data collection protocols reported in the ISO/TS 12913-2: 2018 for urban soundscape via a soundwalk. *Appl. Acoust.* **2019**, *155*, 190–203. [[CrossRef](#)]
25. Mitchell, A.; Kang, J. The spectral structure of acoustic time series can predict the perceptual assessment of urban soundscapes. *J. Acoust. Soc. Am.* **2019**, *146*, 2795. [[CrossRef](#)]
26. Mitchell, A.; Aletta, F.; Oberman, T.; Erfanian, M.; Kachlicka, M.; Lionello, M.; Kang, J. Making cities smarter with new soundscape indices. *J. Acoust. Soc. Am.* **2019**, *146*, 2873. [[CrossRef](#)]
27. Aletta, F.; Oberman, T.; Mitchell, A.; Erfanian, M.; Lionello, M.; Kachlicka, M.; Kang, J. Associations between soundscape experience and self-reported wellbeing in open public urban spaces: A field study. *Lancet* **2019**, *394*, S17. [[CrossRef](#)]
28. Lionello, M.; Aletta, F.; Kang, J. On the dimension and scaling analysis of soundscape assessment tools: A case study about the “Method A” of ISO/TS 12913-2: 2018. In Proceedings of the International Conference on Acoustics ICA, Aachen, Germany, 9–13 September 2019.
29. Aletta, F.; Molinero, L.; Astolfi, A.; Di Blasio, S.; Shtrepi, L.; Oberman, T.; Kang, J. Exploring associations between soundscape assessment, perceived safety and well-being: A pilot field study in Granary Square, London. In Proceedings of the International Conference on Acoustics, Aachen, Germany, 9–13 September 2019.
30. Oberman, T.; Jambrošić, K.; Aletta, F.; Kang, J. Towards a soundscape surround index. In Proceedings of the International Conference on Acoustics, Aachen, Germany, 9–13 September 2019.
31. Kogan, P.; Turra, B.; Arenas, J.P.; Hinalaf, M. A comprehensive methodology for the multidimensional and synchronic data collecting in soundscape. *Sci. Total Environ.* **2017**, *580*, 1068–1077. [[CrossRef](#)]

32. International Electrotechnical Commission IEC 61672-1: 2013 *Electroacoustics—Sound Level Meters—Part 1: Specifications*; Standard IEC 61672-1: 2013; International Electrotechnical Commission: Geneva, Switzerland, 2013.
33. Davies, W.J.; Bruce, N.S.; Murphy, J.E. Soundscape reproduction and synthesis. *Acta Acust. United Acust.* **2014**, *100*, 285–292. [[CrossRef](#)]
34. Hong, Y.J.; He, J.; Lam, B.; Gupta, R.; Gan, W.S.; Hong, J.Y.; He, J.; Lam, B.; Gupta, R.; Gan, W.S. Spatial Audio for Soundscape Design: Recording and Reproduction. *Appl. Sci.* **2017**, *7*, 627. [[CrossRef](#)]
35. Hong, J.Y.; Lam, B.; Ong, Z.T.; Ooi, K.; Gan, W.S.; Kang, J.; Jing, F.; Tan, S.T. Quality assessment of acoustic environment reproduction methods for cinematic virtual reality in soundscape applications. *Build. Environ.* **2018**, *149*, 1–14. [[CrossRef](#)]
36. Lynch, K. *The Image of the City*; MIT Press: Cambridge, MA, USA, 1960; Volume 11.
37. Kaplan, R.; Kaplan, S. *The Experience of Nature: A Psychological Perspective*; Cambridge University Press: Cambridge, UK, 1989.
38. Ewing, R.; Clemente, O. *Measuring Urban Design: Metrics for Livable Places*; Island Press: Washington, DC, USA, 2013.
39. Quercia, D.; O’Hare, N.K.; Cramer, H. Aesthetic capital: What makes London look beautiful, quiet, and happy? In Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing, Baltimore, MD, USA, 15–19 February 2014; pp. 945–955.
40. Joglekar, S.; Quercia, D.; Redi, M.; Aiello, L.M.; Kauer, T.; Sastry, N. FaceLift: A transparent deep learning framework to beautify urban scenes. *R. Soc. Open Sci.* **2020**, *7*, 190987. [[CrossRef](#)] [[PubMed](#)]
41. Pheasant, R.J.; Fisher, M.N.; Watts, G.R.; Whitaker, D.J.; Horoshenkov, K.V. The importance of auditory-visual interaction in the construction of ‘tranquil space’. *J. Environ. Psychol.* **2010**, *30*, 501–509. [[CrossRef](#)]
42. International Electrotechnical Commission. IEC 60942: 2017 *Electroacoustics—Sound Calibrators*; Standard IEC 60942: 2017; International Electrotechnical Commission: Geneva, Switzerland, 2017.
43. Jeon, J.Y.; Lee, P.J.; Hong, J.Y.; Cabrera, D. Non-auditory factors affecting urban soundscape evaluation. *J. Acoust. Soc. Am.* **2011**, *130*, 3761–3770. [[CrossRef](#)] [[PubMed](#)]
44. Guastavino, C. Categorization of environmental sounds. *Can. J. Exp. Psychol.* **2007**, *61*, 54–63. [[CrossRef](#)] [[PubMed](#)]
45. Brown, A.L.; Kang, J.; Gjestland, T.T. Towards standardization in soundscape preference assessment. *Appl. Acoust.* **2011**, *72*, 387–392. [[CrossRef](#)]
46. Axelsson, Ö.; Nilsson, M.E.; Berglund, B. A principal components model of soundscape perception. *J. Acoust. Soc. Am.* **2010**, *128*, 2836–2846. [[CrossRef](#)]
47. Topp, C.W.; Østergaard, S.D.; Søndergaard, S.; Bech, P. The WHO-5 Well-Being Index: A Systematic Review of the Literature. *Psychother. Psychosom.* **2015**, *84*, 167–176. [[CrossRef](#)]
48. Hall, T.; Krahn, G.L.; Horner-Johnson, W.; Lamb, G. Examining functional content in widely used Health-Related Quality of Life scales. *Rehabil. Psychol.* **2011**, *56*, 94. [[CrossRef](#)]
49. Erfanian, M.; Mitchell, A.; Aletta, F.; Kang, J. Psychological Well-being, Age and Gender can Mediate Soundscapes: A large sample study. **2020**, Under Review.
50. American National Standards Institute, Inc. *ANSI/ASA S12.9-2013/Part1 Quantities and Procedures for Description and Measurement of Environmental Sound—Part 1: Basic Quantities and Definitions*; Standard ANSI/ASA S12.9-2013/Part1; American National Standards Institute, Inc.: Washington, DC, USA, 2013.
51. Tarlao, C.; Fernandez, P.; Guastavino, C. Comparing soundscape evaluations in French and English across three studies in Montreal. In Proceedings of the INTER-NOISE 2016, Hamburg, Germany, 21–24 August 2016.
52. Jeon, J.Y.; Hong, J.Y.; Lavandier, C.; Lafon, J.; Axelsson, Ö.; Hurtig, M. A cross-national comparison in assessment of urban park soundscapes in France, Korea, and Sweden through laboratory experiments. *Appl. Acoust.* **2018**, *133*, 107–117. [[CrossRef](#)]
53. Nagahata, K. Examination of Soundscape-Quality Protocols in Japanese. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*; Institute of Noise Control Engineering: San Diego, CA, USA, 2019; Volume 259, pp. 437–446.

54. Almagro Pastor, J.A.; Vida Manzano, J.; García Quesada, R. Soundscape approach for noise management of conflict urban areas. The PTS case in Granada (Spain). In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*; Institute of Noise Control Engineering: San Diego, CA, USA, 2019; Volume 259, pp. 1802–1813.
55. Harris, P.A.; Taylor, R.; Thielke, R.; Payne, J.; Gonzalez, N.; Conde, J.G. Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* **2009**. [[CrossRef](#)] [[PubMed](#)]
56. Harris, P.A.; Taylor, R.; Minor, B.L.; Elliott, V.; Fernandez, M.; O'Neal, L.; McLeod, L.; Delacqua, G.; Delacqua, F.; Kirby, J.; et al. The REDCap consortium: Building an international community of software platform partners. *J. Biomed. Inform.* **2019**, *95*, 103208. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).