

The RAJAR Methodology

RAJAR modified its national listening survey from the Autumn of 2021 following a period of disruption caused by the COVID pandemic. The design changes were to make the survey more resilient through using less face-to-face recruitment resources and to minimise as much as possible the level of data volatility over time through the incorporation of continuous panels.

What is the survey format?

It is a hybrid method comprising of three data collection components combining to produce a unique sample of over 22,000 adults per quarter:

- a 15,000 quarterly sample completing a 7-day diary, with a choice of online, mobile or paper formats – *The Diary Sample*
- a continuous panel of 5,000 participants providing app-based passive listening data – *The Passive Panel*
- a targeted diary panel of around 1,000 to boost representation of young people and selected ethnic minority groups – *The Diary Boost Panel*

(The research contractor for all of the above is Ipsos UK. All sample sizes are for Adults aged 15+.)
The main survey component remains face-to-face recruited, self-complete diaries, but now supported by two distinct panels. The first takes the form of a panel of people registering their listening via an app-based passive measurement system called MediaCell, with a second and separate specialised panel keeping a diary once a month.

Each month, around 5,000 people provide listening data collected by the MediaCell method. This is supplemented by a 1,000-strong diary panel designed to boost demographic representation of young people and those belonging to selected ethnic groups or living in remote regions. Finally, these samples are merged with 15,000 RAJAR respondents, recruited face-to-face in the traditional way. All panel and survey respondents are counted equally in the sample across all reporting periods.

How does the MediaCell measurement work?

MediaCell, as used by RAJAR, uses a process called audio-matching to identify the stations being listened to. This takes the form of an app that is installed on the panellist's smartphone that grabs regular short samples of the ambient sound, which are subsequently compared to the programmes being aired at that precise moment in time across all of the stations on RAJAR. An algorithm determines if there are enough matches of sufficient quality to attribute a sound sample to a particular radio station. The science behind this process is similar to that used by music-recognition apps such as Shazam.

Is MediaCell accurate?

RAJAR has conducted an exhaustive analysis of passive data measurement systems over the past 20 years, and thanks to developments in battery life, storage capacity and component miniaturisation, these techniques have made significant advances. They are not perfect – people may not keep their smartphones with them at all times, or they may not be charged, or a noisy listening environment might contaminate the audio sample quality - but generally they give a good account of people's actual listening and smart edits can be used to fill in gaps in listening records.

Are the diary and the passive data compatible?

Not initially. Because MediaCell captures a continuous record of a person's radio exposure, minute-by-minute, 7 days a week, 52 weeks a year, it is necessary to convert this large volume of data into a RAJAR-compatible format. There are two stages to this process. The first is to apply standard RAJAR diary rules to the data (e.g. a listening session must be at least 5 minutes within a quarter hour to be reported as listening). The second stage is to turn a full month's listening into an average week for each panellist. To achieve this, RAJAR has developed a proprietary model that analyses all listening sessions across all weeks of the month and ranks their frequency by time of day, day of week and station. Then, starting with the most frequently occurring session, the model gradually populates a virtual diary, indistinguishable from a standard RAJAR weekly diary, but with the added benefit that it delivers the purest account of that person's typical listening behaviour across the whole time period.

The model has a calibration setting that matches the MediaCell panel's listening levels to those produced by the 15,000 diarists. As in any hybrid system, it is important to calibrate to one established currency so that all stations in all areas are treated equally, regardless of panel dispersion. This is usually the largest component, which in RAJAR's case, is the standard diary.

With three survey components now instead of one, are the data-sets different?

RAJAR does not produce data based on any single survey component, although we do conduct comprehensive quality control checks at component level. This is because the components are complementary to one another and are designed to work as a whole. For example, when sampling for the face-to-face element, the locations of panellists are taken into consideration so as to maximise sample efficiency and coverage. For a similar reason, RAJAR does not break out its paper v mobile v laptop/desktop diary samples since each of these appeals to different population groups, which, by definition, would demonstrate quite different listening patterns that would not be comparable.

What are the advantages of the current RAJAR methodology?

The MediaCell panel helps to reduce volatility because of its relatively stable sample profile, as well as presenting us with the best possible average of each person's listening. The diary panel guarantees us a minimum level of representation among population groups that are often difficult to engage with on a random basis. And should another pandemic (or a severe weather event or any major disruption) compromise face-to-face interviewing, the panels give us some insurance so that we can continue to monitor radio listening in the intervening period.