

## **JRP SECOND WAVE PREPARATION FOR THE CONTENT ANALYSIS**

In preparation for the content analysis, several sequential stages need to be implemented: (a) all national teams must select their respective news media outlets to be included in the study; (b) the sampling units within each outlet need to be determined; (c) as the analyses in all countries will be conducted at the same time – during the same two constructed weeks during 2020 – the dates for the analysis must also be determined; and (d) deciding on the unit of analysis, that is, which individual items will be analyzed.

This document is designed to deal primarily with the choice of media outlets and the sample units, but brief references to the other issues are also presented and will be developed at later stages.

### **Selection of News Media Outlets**

Each national team will select two-to-four news media outlets per platform – television, radio, newspapers and online news websites – taking into account the following general criteria:

- The media outlets must be independent of each other, that is, no two outlets can belong to the same organization. However, some countries may have public and commercial radio and television broadcasts that are part of the same organization; in this case, both radio and television should be represented. Also, in order to avoid potential overlap in contents, the online news websites will *not* be those of the selected newspapers or television and radio stations. Although preference should be given to “net-native” media, in some countries they are not well developed and online news is dominated by websites of established media organizations; if that is the case, the sample should include both types.
- Preference should be given to outlets that are national in scope (in contrast with “local” or “regional” scope. Indeed, in most countries the dominant media (mainly broadcast) are national in scope. Nevertheless, in some countries there are considerable regional and local outlets that are as important as national media outlets. If relevant, the sample should aim to include both types.
- The outlets must be among the most popular (read, seen, or listened to) in their class, based on ratings or similar parameters.

Since the structure and format of media systems differ in many respects across the countries, such as with regards to size, media audience orientation (popular vs. elite), media ownership, media political leaning, and multiple languages, the selected outlets should represent as best as possible the diversity of the country’s media system. This leads to three relevant considerations:

- The number of media outlets that a country must include in order to represent the variation of its media system may vary from country to country.
- If most of the critical characteristics of the media outlets are constant in your media systems, you will need fewer outlets per platform. In other words, greater heterogeneity in your media system will lead you to include more outlets, and vice versa.
- Exceptions to the general criteria can be made when needed.

Some examples may be helpful to clarify potential variations among countries leading to decisions that need to be taken:

- While popular-oriented news outlets are not present in some countries, they are highly prevalent in other countries, so if they exist they should be included in the study.
- In some countries all media are private/commercial; in other countries there is a mix between private and public (mostly broadcast) media; while in still other countries all media may be public (state-owned). The sample of outlets should include all relevant types.
- In multilingual countries, in which language is an important feature of the media system, an attempt should be made to include media outlets that reflect all dominant languages.
- In countries in which media outlets are organs of political organizations, care should be taken to include outlets with different political orientations.

### **Selection of Sampling Units**

After selecting the media outlets, we need to determine which specific **sampling units** will be dealt with in each outlet. Specific criteria will be used in making the decision for each medium/platform.

For **television**, the sampling units will be the *most heavily viewed* newscasts within each selected outlet. Usually, the most viewed TV news program is the evening newscast but this may not be the case for all countries. It may also vary by day of the week (see below). It is important to determine whether the newscasts can be retrieved at a later date from the relevant online repository; if not, they will be recorded off-air.

For **radio**, the sampling units will be the *most listened-to news programs* within each selected outlet. Usually, the most listened-to radio news program is broadcast early in the morning, but this may not be the case for all countries. In some countries the relevant news broadcast (or bulletin) is brief (only several minutes) while in other countries it may last a full hour or even several hours. In the latter case, the goal should be to identify and suffice with the most popular one hour-long segment. Here, too, it is important to determine whether the broadcasts can be retrieved from the relevant online repository; if not, they will be recorded off-air.

For **Newspapers**, the sampling units will be the specific full issues of the selected newspapers published on the relevant dates. If it is possible to download the news from websites, this should be done. Alternatively, copies of the issues may be purchased and kept on file.

For **Online** news media, the sampling unit will be the entire homepage of the selected news websites (including its respective links) that will have appeared on the relevant dates. Whereas our selected television and radio news programs as well as newspapers are “static” in the sense that they are unique and appear at fixed times, news on the web is dynamic and changes constantly. We will therefore “capture” the homepages of the websites at two fixed points during the relevant days: once at 12:00 noon and once at midnight. The rationale for this is that the noon capture would be several hours after the appearance of the newspapers and morning radio news broadcasts whereas the midnight capture will tap the news several hours after the main evening TV newscast and about the time that the newspapers of the following day go to print. This 12-hour gap between both captures will most likely provide the greatest amount of content variability. The home pages and their respective links will be opened in real time (around noon and midnight) and saved.

### **Selection of Sample Days**

In designing the sample of the days that we will include in the study, we will use the same effective strategy as in the previous wave of the JRP Project.

Because daily and monthly variations are important factors to consider when conducting a news content analysis, we will divide the year 2020 into two 6 month periods: January-June, and July-December. For each six-month period we will create a constructed week, that is, stratified-systematic samples.

Accordingly, for each of the two 6-month periods, we will randomly select starting dates on a Monday in January and a Monday in July. Then, using 3-4 week skip intervals we will select each of the subsequent six days day: a Tuesday, a Wednesday, a Thursday, a Friday, a Saturday, and a Sunday. We will thus have 7 days in each 6-month period, for a total sample of 14 days during the year.

In some countries there is no news provided by some outlets on weekends or the format and/or time slots may differ from weekdays. In fact, the main day off (part of the weekend) varies between Sundays, Saturdays and Fridays. Thus in some countries there may be no newspaper published or no news program on television and/or radio on weekends or part thereof. These situations will be considered as “missing data.” If, however, only the time of a television and/or radio news program is shifted (e.g. in the case of sports events) then the newscast should be coded in the actual time slot.

Since all countries will be analyzing the same days, we will send you the list of the specific dates at some point in the near future.

### **News to be analyzed**

Once we've recorded the news of the first day (sometime in January 2020) from all outlets, we can begin coding. This will be done using our specially created and validated Codebook that will be devised on the basis of the previous JRP wave (that dealt only with print news), and that is now adapted for all four platforms. Details of the codebook will be discussed in the coming months.

The **unit of analysis** to be coded for all outlets and sampling units will be the **news item**. A news item will be conceptualized as a set of contiguous verbal and, if relevant, also audio and/or visual elements that refer to the same event/issue/person. The same definition for the unit of analysis will be used for all news outlets.

For each of the selected issues/editions/news programs from newspapers, television, radio and online news websites, the following news **topics** will be coded: politics, economics, business, police and crime (law and order), courts, defense, health, education, public works, housing, transport, energy, environment, accidents, natural disasters, religion and faith, human rights, marches and protests, science and technology, sports, culture, and entertainment.

Of course, an event, issue, etc., may be reported by the media in more than one news item on a given day. In such cases, both (or more) items would be considered as separate items and will be coded separately.

For all media outlets the following will **not** be coded: editorials, opinions, weather forecasts, horoscopes, movie (or other cultural) reviews, and puzzles. Furthermore, we will not code supplements/magazines/special features programs, published or broadcast by any media outlets. Headlines on newspaper front pages and at the beginning of TV newscasts and radio news broadcasts will also not be coded *per se*, but in the Codebook there will be a reference as to whether or not there was a headline for the item.

Stories authored by news agencies and wires or news stories that are copied and pasted from other media outlets (where the media you are analyzing explicitly states that the story was not produced or written by them) should not be coded.

In the case of online media, we will only code items in the above topic categories that appear on the extended home page by clicking on, and thereby opening, each of the relevant items. Items that have video (or audio) clips embedded in the item will also be coded.

The coding of all the items will be carried out by coders trained by the PIs in their respective countries. In order to reduce bias, the corpus of the news items in each country should be divided randomly among coders (thus no coder should code an entire outlet by herself/himself). Ideally, each national team should have between two to four coders to establish good inter-coder reliability.

Before starting the coder training process, PIs in each country will be requested clarify all their uncertainties with the General Coordinator of the Project or a member of the Executive Committee. If necessary, the General Coordinator will travel to work with the countries that need help in the training process.

Before starting with the actual coding we will conduct pretests within and across countries in order to ensure comparable interpretations of the Codebook. This process will be explained in due course.