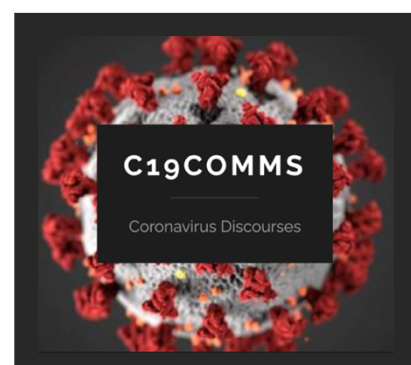# Participant Information Sheet, February 2022

Coronavirus Discourses Online: Participating in Studies on the Reception and Production of Public Health Messaging

Emma McClaughlin, Dawn Knight, Svenja Adolphs, Jeremie Clos, Elena Nichele, Pepita Barnard, Derek McAuley, Alexandra Lang
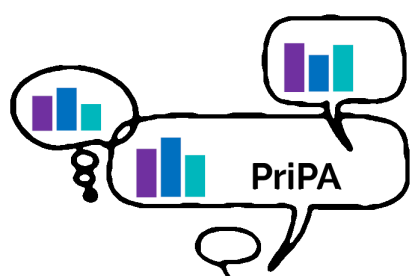
Background to the Study

Researchers on the Coronavirus Discourses project at the University of Nottingham and Cardiff University are working with project partners Public Health England, Public Health Wales, and NHS Education for Scotland to understand the **flow and impact** of **public health messages** in **public** and **private communications**.

In the UK, official public health communications are released by the Government and Public Health Agencies. These are shared through many other channels, including the news and on social media. In this document, we focus on the use of the software we have developed for the project and how the public can get help us gather information and data.

We are interested in how language about the coronavirus pandemic reaches the public, whether it is changed by those who reproduce it, and if so, how it is changed.

Our findings will help our project partners in the design of future public health campaigns

How Does the Study Work?

We have developed a 'web browser extension' called **PriPA** – short for **Pri**vacy **P**reserving **A**nalytics. PriPA will help us to gather insights from the public about Coronavirus to support our public health agency partners, without compromising the privacy of the people providing the data. Individuals should be able to **trust that their personal data is safe**, secure, and their identity remains private, as only the general insights are shared with researchers.

The following information explains **how you might get involved** in the research and provides answers to frequently asked questions about the study.

The University of Nottingham in collaboration with Cardiff University is working in partnership with Public Health England, Public Health Wales, and NHS Education for Scotland to investigate linguistic evidence for effective public health messaging.

We are interested in how official Coronavirus public health messages have been received and reproduced by members of the public. This means investigating:

- Individuals' browsing activities on specific news websites to understand what information people are accessing about COVID-19 and from where.

- Language that is shared in private (e.g., conversations happening on social media and other sites that require a login) concerning COVID-19.

How to Get Involved in the Study

You are invited to install the PriPA extension on your Google Chrome web browser to submit simple language data about COVID-19 from the news websites you browse and, if you wish, from the language you use when you talk about Coronavirus with others. Please read through this information sheet carefully for more information about how PriPA works to gather language data in a privacy preserving way.

What language data does the study need from the public?

We are interested in the information provided in the table below. Nothing else will be detected by the PriPA extension for the Coronavirus Discourses project. Importantly, you will be able to view everything the browser extension collects before you decide to submit any data to the study.

Sources

The mainstream news and information sites where you read information about COVID-19 (e.g., the BBC, Sky News). We are interested in the content you are shown.

With your additional permission, we are also interested in where you are contributing to conversations online (e.g., on social media platforms such as Twitter).

Word Frequencies

Lists of the words that you read and/or write online, ordered by frequency.

We are not interested in individual differences across word lists, only aggregated results. In other words, we will combine the individual word lists from several participants before analysis.

Collocates

Words that most frequently co-occur in language use (e.g., salt + pepper). We are interested in how combinations of words reveal patterns of meaning. For example, the word 'substance' appears to be neutral but in reality it often co-occurs with words like 'harmful' and 'toxic', so substance takes on a negative meaning from its context. Other words take on positive meaning from their context.

Concordance Lines

A line of text showing the linguistic context for a particular word. Concordance lines give us a bit more context to explore patterns of meaning. To preserve privacy, they will not be collected from personal messages on social media.

In the example below, we can see the language appearing before and after the words 'family' and 'ill' from BBC news.

#	Left	KWIC	Right	Count	Source	Exclude?
1	Christmas, reducing their mixing so they could spend time with	family	he says. "But... with schools back and people working again	1	www.bbc.co.uk	☐
2	care homes have not led to significant numbers becoming seriously	ill	The boosters seem to have had the desired effect. \n"There	1	www.bbc.co.uk	☐

Example concordance lines from bbc.co.uk for the words 'family' and 'ill' from PriPA version 0.0.0.6

The Parameters of the Study

There are two strands to the study: one explores **reception and production across news and social media websites** and the other explores **reception** of language about COVID-19.

News Websites

If you would like to join the study on COVID-19 language reception, you can submit words, collocates (co-occurring words) and concordance lines (longer phrases or sentences) generated by PriPA from any of the websites that you browse from the news outlets shown.

- The Daily Telegraph
- Huffington Post
- The Independent
- The Daily Mirror
- The Daily Express
- The Evening Standard
- The Daily Star
- Nottingham Post

- The Guardian
- The Observer
- Daily Mail
- Google News
- Buzzfeed UK
- Yahoo News
- MSN News
- Financial Times

- Metro
- i News
- The Sun
- BBC News
- Sky News
- ITV News
- LADbible

Social Media

If you agree to join the study on reception and production, you can submit counts of words and collocates (co-occurring words) generated by PriPA, but no concordance lines (longer phrases or sentences) from your direct messages on Twitter and/or Facebook.

For the collocation analysis, we are investigating the use of **modal verbs** as **stance markers** that tell us about attitudes towards Coronavirus information, guidance, and instruction.

These modal verbs are **may**, **might**, **must**, **should**, **would**, **could**, **can**, **will**, **shall**, **ought to**, **need**, and **dare**, along with their negatives e.g., **wouldn't**, **would not**, **would never**.

We are interested in how these words co-occur with other words relating to Coronavirus, which are provided below.

Coronavirus Language of Interest

For the language reception and production study, PriPA will count frequencies of the following words. It will also calculate how often they co-occur with the words above (collocation analysis). <u>For the language reception study only</u>, PriPA will gather extracts of text containing these words and ten words either side (concordance lines), in addition to the frequencies and collocation.

Public health measures

- mask
- face covering
- vaccine
- hands
- face
- space
- fresh air
- do your bit
- 2m
- two metres
- metre
- wash
- f.a.c.t.s
- social distancing
- stay home
- protect the NHS
- save lives
- rules
- protect
- don't guess
- get a test
- self-isolate
- mouth
- nose
- face
- essential
- safe
- safety
- protect
- sanitise
- stay alert
- control the virus
- virus
- avoid crowds
- clean hands
- book
- test
- stop the spread
- safe
- bubble
- home
- stay
- control the virus
- work from home

Social groups and populations

- black
- asian
- minority
- ethnic
- of colour
- bame
- white
- people
- vulnerable
- disability
- vulnerable
- disability
- socio-economic
- income
- public transport
- at risk
- family
- friends
- neighbours
- community
- social
- everyone
- bipoc

Geographic and social borders

- scotland
- wales
- england
- britain
- english
- scottish
- welsh
- british
- bubble
- national
- global
- local
- regional
- high alert
- medium alert
- medium alert
- low alert
- united kingdom
- UK
- tier

Effects and consequences

- long covid
- fatigue
- spread
- die
- risk
- danger
- symptoms
- poorly
- ill
- sick
- unwell
- virtual burnout
- cold
- flu

Answers to commonly asked questions about our study

More information including a jargon buster is available on our website **c19comms.wp.horizon.ac.uk**

What is Corpus Linguistics?

Corpus linguistics is the study of language using large, principled **digital collections of texts** called 'corpora' (singular 'corpus'). It uses specialist software to identify common patterns that occur in language, including the **frequency and usage of specific words**. It has been used for many purposes including language teaching and learning, dictionary writing, and to explore language about different topics. PriPA is a browser extension that **anyone can use** to submit simple corpus linguistic analysis from their own computers.

What is a 'browser extension'?

Browser extensions are **computer programs which run within Web browsers** (e.g., Google Chrome) on your computer. Extensions usually add extra functions to the browser. For example, you might have heard of extensions to remove advertising from Web pages (adblockers), or safely store your passwords.

How will this information be used for the Coronavirus Discourses study?

Information on browsing activities and language use generated by the PriPA extension will be used **anonymously** for research publications and presentations and shared with our public health partners to help shape their **future messaging campaigns**. This means that people who choose to submit their results cannot be identified by the researchers carrying out the study or in these outputs. We will not use this data for anything else and **the data will be deleted when the project is finished**.

Why use a privacy-preserving browser extension for Corpus Linguistics?

Having access to the text that people read whilst browsing and the conversations that take place between individuals online will help us investigate how people have understood what they have read. We want to make sure that **we only gather the general insights and remove all identifiable data**, so we do this within a browser-extension, only recording the results of the privacy preserving analysis. This is the **PriPA browser-extension**. PriPA is dormant until a user signs up to a specific project, which will clearly list the websites that will be processed, and the words that the tool will search for. The results of that analysis are only shared when the user has reviewed them and **explicitly agreed to submit them**.

Submit results
By clicking 'Yes, submit results', you agree to storing the results externally for research purposes only.
If you do not agree with that please press 'No, don't submit results and delete results'.

[No, don't submit results and delete results]　[Yes, submit results]

Participant option to submit or delete results

How can I be certain that I will not be identifiable if I take part?

> You do not share any of the PriPA results until you decide to submit them.

> You have the option to view and delete your results (i.e., word list and concordance lines).

> You can decide to exclude certain results from the data if you do not wish to share them.

> If, after sharing, any clearly identifiable personal data is detected, it will be removed from the data by the researchers.

> We cover these points in more detail on page (5).

Answers to commonly asked questions about taking part

More information including a jargon buster is available on our website **c19comms.wp.horizon.ac.uk**

What if I'm interested in viewing my results but I don't want to share them?

That's fine! The extension is **free to use**, and you don't have to make your results available to anyone else if you don't want to. Remember, all the analysis is carried out in your browser (Google Chrome) and no one else can see the results unless **you choose to share** them.

Can I look at my results and then decide whether I want to share them?

Yes of course. If you do decide to share your results when you've seen them, you can even choose to exclude any extracts (concordance lines) that you want to keep private before sharing. **The researchers will never see language data you have excluded from your shared results**.

Can I download my results to look at later?

Yes. Your results will be **available to download** in a comma-separated values (.csv) format for your own personal reference. This type of file is readable by any spreadsheet software such as Microsoft Excel.

Can I decide to withdraw my results later?

Yes, this is an important part of **ethical research**. Participants have the option to withdraw their data from the study. It's easy to do and instructions for how to withdraw are available to everyone who takes part. After data (e.g., word counts) have been combined with data from other participants and published, this cannot be withdrawn, but we will always remove your responses from the data that we hold.

Ok, I'm interested in taking part in the study. Where do I start?

Thank you! Please go to our website **c19comms.wp.horizon.ac.uk** for details on how to register your interest for the study. You will receive instructions by email if a suitable study is available at the time of registration. Alternatively, we may be able to put you on a reserve list for the next phase of our project.

I have another question!

Further information about the extension and the Coronavirus Discourses study is available on our website. Scan the QR code or visit c19comms.wp.horizon.ac.uk. Here you can find out more about the **project team**, read the **jargon buster**, access **resources**, and sign up to **take part in the research**.

If you have further questions, please contact our Principal Investigator Professor Svenja Adolphs by email: svenja.adolphs@nottingham.ac.uk.