The L' ART Research Assistant

Breit, Florian; Tamburelli, Marco; Gruffydd, Ianto; Brasca, Lissander

Cyswllt i'r cyhoeddiad / Link to publication

The L'ART Research Assistant: A digital toolkit for bilingualism and language attitude research[*]

Florian Breit, Marco Tamburelli, Ianto Gruffydd, Lissander Brasca

Bangor University

1. Introduction

The L'ART Research Assistant is a freely available open-source app to aid researchers in the collection, storage and transfer of data for research in bilingualism and language attitudes, with a particular focus on bilingual populations who speak a majority language and a regional / minority / minoritized language. The app aims to make research in bilingualism easier, more comparable and reproducible. The current version of the app [version 0.4.1] allows gathering of data on linguistic background, as well as other commonly used tasks, particularly in relation to investigating language attitudes in bilingual communities. A full list of the tools provided by the app is given below.

2. L'ART Research Assistant: Tools

Version 0.4.1 of the L'ART Research Assistant implements the following four tools:

1. A digital **informed consent** tool (see section 5 for details).
2. A digital adaptation of the **Language and Social Background Questionnaire**, or LSBQ (Anderson, Mak, Keyvani Chahi & Bialystok, 2018), which we term the **LSBQ$_e$** ("e" for electronic). See section 4.1 for details.
3. A digital implementation of the **Attitudes towards Languages Questionnaire** or AToL (Schoel, Roessel et al., 2012), which we call **AToL-C** ("C" for continuous, as our implementation uses continuous rather than ordinal scales. See section 6 for detail).
4. A digital tool for measuring language attitudes via the speaker evaluation paradigm. This tool enables users to run several evaluations of audio guises such as the **Matched Guise Technique** (Lambert, Hodgson, Gardner and Fillenbaum, 1960) and the **Verbal Guise Test** (e.g., Markel et al, 1967). Due to its flexibility as either MGT or VGT, we named this tool 'Audio Guise Task', or **AGT** for short.
5. A simple memory game to be employed as a distractor in a series of tasks (see section 8).

The main functionality of the L'ART Research Assistant resides in its format as a stand-alone app that can run on a large variety of desktop and laptop computers without the need for internet connectivity. This makes it highly usable both in lab environments and in the field, for example when collecting data in remote areas with inconsistent internet access. It also means that researchers using the app are not reliant on the availability of an online resource that they may not have full or direct control of at any given time.

3. The L'ART Research Assistant: Main Advantages

There are many advantages in using an app over a paper-and-pencil or a word-processor based questionnaire, especially where the use of a single tool finds broader adaptation across several studies and/or research groups.

---

[*] Version updated: 04 May 2023.

In short, adopting a toolkit like the L'ART Research Assistant means:

- **Less work for the researcher:** With research tasks pre-implemented, preparation for a new study only involves translation/localisation of the desired task's interface where a suitable one is not yet available for the target population. There is also no need to manage forms and manually enter data after collecting responses.
- **Enhanced consistency and comparability** within and across studies: The translation/localisation of tasks is the only thing that varies within tasks. The presentation, data types, data validation, coding, and output format stay constant across different use instances, whether as part of the same study or across different studies and research teams.
- **Improved transparency and reproducibility**: Because the entire source code for L'ART Research Assistant is publicly available and version-controlled, it's easy to reference the specific version and task or set of tasks that were used, allowing other researchers to easily view and reconstruct the tasks exactly as they were administered at the time the research was carried out.

Below are some concrete illustrations of some of these advantages, using the LSBQ$_e$ as an example.

3.1 Avoid mistakes before they happen

3.1.1 Data Collection

Have you heard the one about the perfectly balanced bilingual who only spoke Spanish? Paper-based and/or word-processor based questionnaires are prone to human error. A participant may omit answering one or more questions, forget to enter some of the information requested, or provide data that is inconsistent in some manner. A researcher can accidentally omit a question (or part of it) or alter the layout and formatting in an unintended manner which influences the participant's response in some way or makes it more likely that a question is not answered as intended. Additionally, the same questionnaire opened in a different word processor might not only look different but also inadvertently affect some of its functionality, such as a form field suddenly accepting text instead of requiring dates.

While an app cannot entirely eliminate such sources of error, it can considerably reduce their occurrence (e.g., Vergnaud et al., 2011). In the L'ART Research Assistant, questions that necessarily apply to all participants require that each participant provides an answer before they are allowed to continue, thus ensuring that required data will not be missed out due to oversights:

Fig. 1: The LSBQ$_e$ ensuring that bilingual participants provide information about at least two languages.
At the same time, participants can opt out of questions that may not apply to everyone, such as the question asking where, besides their place of birth, they have lived for a period of time over six months:

Fig. 2: Example of a question that can be left blank. Note the green text and tick marks, indicating that the empty fields are validated. Compare this to the empty field in Figure 1 above, which is flagged in red with an exclamation mark, indicating that it must be filled.

The LSBQ$_e$ can also be optimised for specific locations so that the two languages most relevant to a specific bilingual population appear at the top of the drop-down list, potentially followed by other languages of particular interest and then by a further list of languages most commonly associated with the target community. For example, when carrying out research with Welsh-English bilinguals in the UK, the options "Welsh" and "English" appear first:

Fig. 3: Language options presented according to location / population being studied.

At the same time, since participants might speak languages or dialects which could not be reasonably anticipated, the field also allows free text entry, as shown below:

Fig. 4: Example of free text entry for additional languages.

3.1.2 Enhanced Data Rigour

A common weakness of paper-based or word-processor based tasks is that they make it difficult, if not impossible, to control for consistent stimulus exposure across participants. This is particularly important in tasks that are meant to measure participants' attitudes, whether in questionnaire form such as the AToL (Schoel, Roessel et al., 2012), or in the form of auditory tasks such as the Matched Guise Technique (Lambert, Hodgson, Gardner and Fillenbaum, 1960). In the case of a paper questionnaire, some participants might go back on their answers, e.g., by changing their responses about one language after they were meant to have moved on to the next language. This interferes with any counterbalancing that researchers may have wished to build into the design and potentially introduces confounds into the data, as some participants will have provided more instinctive judgements than others. Similar, and possibly more intricately confounded issues can arise in relation to auditory tasks, for example if participants are able to listen to the same recording more than once.

To prevent the introduction of these confounds, the L'ART Research Assistant does not allow participants to go back to a previous section that they have already completed. For example, in the AToL-C (see [section 6](#)), a participant cannot go back to change their ratings for Welsh once the task has moved on to English. In the AGT ([section 7](#)), participants can only listen to a guise once, and they cannot go back to change ratings for a previous guise once they have moved on to the next guise. This allows for a more consistent data collection procedure and ultimately more rigorous datasets by minimising common confounds across participants. Researchers can, however, override this behaviour (e.g., in cases where an error occurred) by using the unlocking feature of the app [(User guide: unlocking the app).](#) Requiring the researcher's intervention in this way also ensures that the researcher will be aware of any issues as they occur and can keep appropriate records.

3.1.3 Data Entry

Another juncture at which human error can negatively affect data quality is when paper-based responses are digitalised or, in the case of word-processor based questionnaires, when responses are collated for use in a spreadsheet application or statistical package. The L'ART Research Assistant eliminates this step, as the data from each response is stored in an individual file employing a widely used data exchange format (JSON), which is easily exported, shared and inspected, as well as being

compatible with most tools in the modern "data-pipeline", and ready to be imported into any data-analysis software.
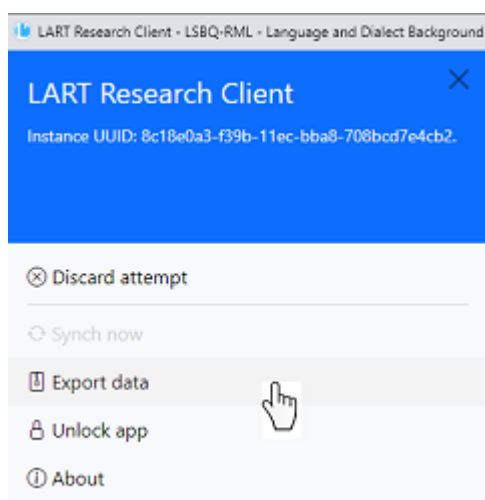

Fig. 5: Making data export simple, ready for backups or data analysis.

3.2 Data security
Data files that store participant responses in the L'ART Research Assistant are automatically assigned a universally unique identifier (UUID), which reduces the chance of ending up with duplicate IDs for separate participants and/or data instances, even when collating very large data sets from many different studies. With this system, the chance of duplication is practically *nil* – we would have to generate one data file per second for approximately a billion years in order to have 0.1% chance of repeating the same UUID.[2] Even the most eager team of researchers would likely be unable to accomplish such a feat.

Data back-up is also made easy and secure. Because each response is uniquely identified and stored in its own individual file in the standard roaming profile locations of the user's operating system, participant responses will be easily captured by many standard solutions for the synchronisation and backup of user profiles deployed by system administrators, as well as being easily targetable for cloud synching and backing up with user-level tools such as *rsync* [3]. This configuration also allows for simple merging of datasets by copy and paste without the risk of data loss due to accidental overwriting of a newer data file with an older one.

4   Research Tools: Adaptations and developments
Beside the many advantages that come with the modern digital implementation described above, the research tools available through the L'ART Research Assistant also involve several adaptations and developments in terms of their contents and – to some degree – their structure. Details for each tool are provided below.

4.1 From LSBQ to LSBQe
While retaining most items from the original LSBQ (Anderson, Mak, Keyvani Chahi & Bialystok, 2018), the *LSBQ$_e$* includes adaptations that broaden its use case, especially with regard to the study of bilingual populations who speak a majority language and a regional / minority / heritage or

---

[2] https://towardsdatascience.com/are-uuids-really-unique-57eb80fc2a87 – accessed on 24th February 2023.
[3] https://rsync.samba.org/

otherwise minoritized language. It additionally includes adaptations for greater flexibility (e.g., for use outside of Canada, where the original LSBQ was based) as well as allowing more standardised comparison across different bilingual populations (particularly in relation to educational level, see [section 4.4](#)). Each adaptation is described in some detail below.

4.2 LSBQ$_e$: Flexibility and configuration for specific research locations

We've designed and implemented the LSBQ$_e$ so that questions which just reference "the other language" in the original LSBQ (especially those under "*Language Use Bahavior*") can specifically refer to the language of the community being studied. This is particularly useful in cases where specific bilingual communities are being studied, so that the names of their languages can appear in the relevant questions in place of a generic "the other language".
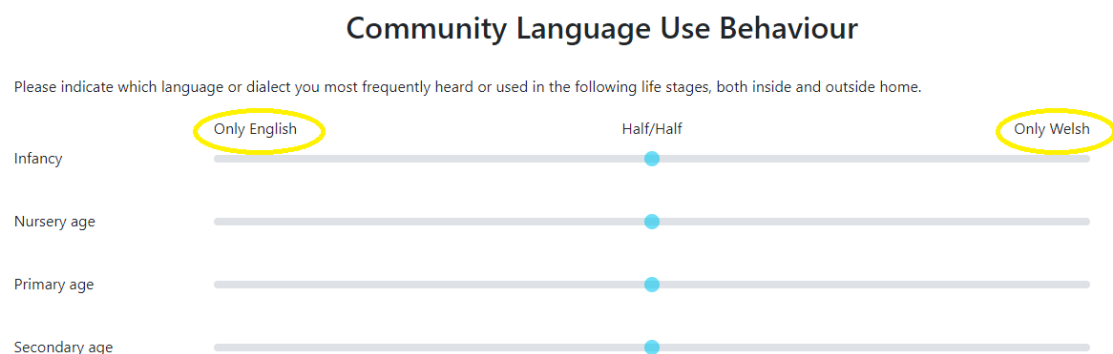

Fig 6: Explicitly referencing both target languages in the sections on linguistic behaviour.

However, seeing as research on bilinguals may also be carried out on more generic bilingual populations, for example on bilinguals who speak English and any other language, or one of a group of languages (e.g., English and any Slavic language), the LSBQ$_e$ retains the option of using the phrase "the other language" as a default. The current version offers implementations for several language pairs, each associated with a particular bilingual community: Welsh-English (available both in English and Welsh, localised for use in the United Kingdom), Moselle Franconian-German (available in German and localised for use in Belgium), Lombard-Italian (available in Italian and localised for use in Italy), Greek-English (available in Greek and localised for use in Greece), and Alemannic-German (available in German and localised for use in Germany). Additionally, the current implementation includes a generic option for English plus 'any other language', available in English (localised for the United Kingdom) and selectable from the dropdown list.
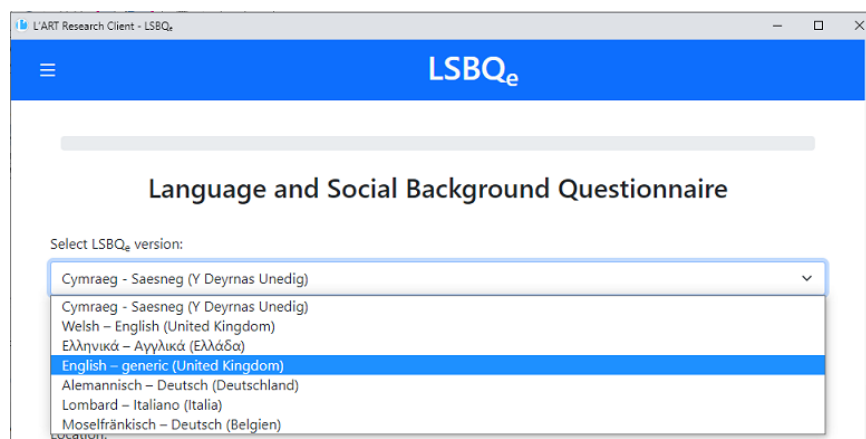

Fig. 7: Selecting the generic option when working with bilingual participants who speak British English and any other, unspecified, language.

Following the original LSBQ. a generic option simply refers to participants' "other language" which is to be rated in opposition to English.
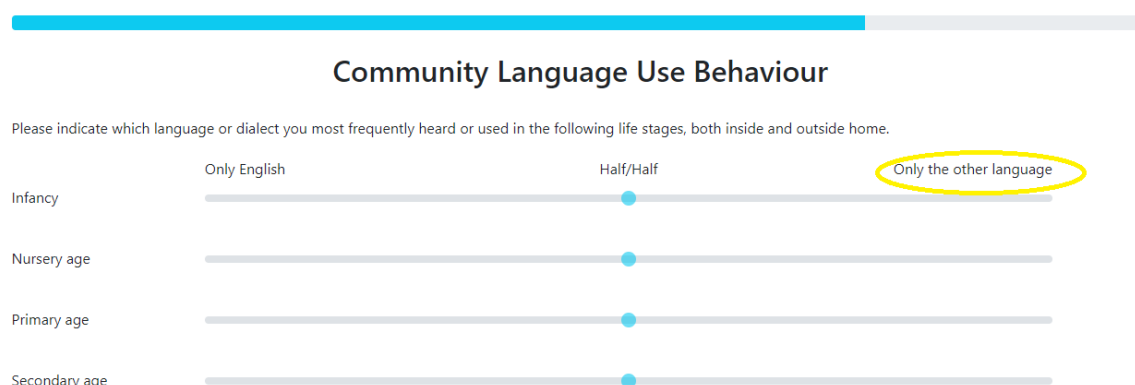

Fig 8: Reporting language behaviour for English and for "the other language" in the generic version for British English.

The app also stores further generic versions for German, Italian and Welsh, and while these are not available from the dropdown menu as default, they can be rapidly added, integrated and deployed where desired (see User Guide for details).

Importantly, users can relatively easily provide their own localised language version of the LSBQe (or any other task in the app) by following instructions in the localisation section of the User Guide.

4.3 LSBQe: Non-Linguistic habits and background

The original LSBQ begins with a host of questions on non-linguistic habits such as "*Do you play first-person shooting (FPS)/action video games?*" as well as background questions that may be relevant to a particular type of linguistic study but are not specifically about linguistic issues, such as "*Do you have vision problems?*" or "*Do you have any known neurological impairments?*". In the LSBQe, we treated some of these questions differently, as follows.

4.3.1 Video games

We excluded the question on FPS video games, as this was very specific to the research for which the original LSBQ was built. Our decision is in line with current GDPR (General Data Protection Regulation) which states that only data that are "*relevant and limited to what is necessary for the purposes for which they are processed*"[4] should be collected. We therefore omitted the question on FPS video games, as it pertains to data that are not relevant and/or necessary for most research projects on bilingualism and/or language attitudes.

4.3.2 Hearing and Vision

The LSBQe retains the questions on hearing and vision, for two reasons. First, the visual nature of the task is such that severely visually impaired participants might not be able to provide reliable data, making the question on visual impairment both relevant and necessary. Second, the question on hearing impairment is necessary as well as highly relevant to the majority of linguistic studies, specifically in relation to sampling (e.g., Horwitz et al., 2003, on the importance of excluding hearing-impaired participants from study samples). It is also directly relevant to at least one other task

---

[4] https://www.privacy-regulation.eu/en/recital-39-GDPR.htm – accessed on 24th February 2023.

implemented in the app itself, namely the AGT (see section 7). However, we amended the wording of the questions slightly, using "impairment" in place of "problem", in keeping with more modern phrasing.

Fig. 9: Background questions about hearing and vision. The questions are adaptive, so that the "if yes, …" questions only appear if the participant has already answered "yes" to the broader question.

3.2.3 Task-specific questions

The remaining questions on non-linguistic background may or may not be relevant and necessary depending on the task(s) that participants may need to complete after filling in the LSBQ$_e$. For example, the question on colour-blindness will likely be relevant and necessary if participants are to take part in a Stroop task, where colour blind participants are routinely excluded (e.g., Naber et al., 2016). In the interest of flexibility, we therefore excluded these questions from the LSBQ$_e$ proper, but catered for their inclusion in the digital consent form, where researchers can set eligibility criteria that are specific to their cohort (see the section on obtaining consent in the User Guide for details). In the "generic consent form", users will find all the remaining questions from the original LSBQ framed as eligibility criteria. In cases where meeting these criteria is particularly important for the applied research paradigm, we recommend that users introduce an additional, explicit step to confirm the criteria with the participants beforehand (e.g., during pre-screening).

Fig. 10: Eligibility criteria from the digital consent from (see obtaining consent in the User Guide for details). Note that the questions on hearing and visual impairments also appear in the LSBQ$_e$, as in Figure 9.

4.4 LSBQ$_e$: Educational level

To facilitate crosslinguistic and cross-cultural comparisons, the information on educational background – which usually forms the basis for the measurement of Socio-Economic Status (e.g., Calvo & Bialystok, 2014; Bialystok & Shorbagi, 2021) – is linked to the European Qualification Framework (EQF),[5] a framework that has been specifically developed to enable more accurate comparison between qualifications across different countries and education systems. Besides allowing alignment of educational achievements across 38 countries,[6] the EQF has also been recognised outside Europe as a vital reference framework when comparing qualifications across countries worldwide (Chakroun, 2010). Having fed directly into the International Standard Classification of Education (ISCED),[7] which is the UN's national qualifications framework (see e.g., Manuel Galvin Arribas, 2016), the 8 EQF levels are largely commensurate with levels 1-8 of the ISCED 2011 (UNESCO Institute for Statistics, 2012). Being based on specified learning outcomes that were developed with the direct involvement of stakeholders such as educators and potential employers, the EQF is highly comparable to the educational and vocational qualification frameworks in use in countries outside the EU, thus maximising the comparability of data collected via the LSBQ$_e$.

Therefore, while using a 5-point scale as in the original LSBQ, we have re-referenced the points of the scale to the following EQF levels:

1. EQF Level 1
2. EQF Levels 2-3
3. EQF Level 4
4. EQF Levels 5-6
5. EQF Level 7-8

In individual translations/localisations of the LSBQ$_e$, rather than giving labels such as "EQF Level 4" (which would be unintelligible to most participants), each EQF level is given a set of equivalent qualification(s) at the targeted EQF level for the respective locality. For example, in the Welsh-English localisation for the United Kingdom, the five EQF levels correspond to those illustrated in Figure 11.

Please indicate the highest level of education you have achieved:

○ No formal qualifications.
○ GCSE(s), NVQ Levels 1-2.
○ A-Level(s), AS-Level(s), NVQ Diploma, HNC, Apprenticeship.
○ Undergraduate degree, Foundation degree, HND.
○ Postgraduate degree, doctorate.

Fig. 11: Prompting the user with actual examples of qualifications at each level, rather than the abstract EQF levels these have been indexed to.

A useful tool to compare EQF levels across EU countries can be found at:
https://europa.eu/europass/en/compare-qualifications

---

[5] https://europa.eu/europass/en/european-qualifications-framework-eqf – accessed on 24th February 2023.
[6] At the time of writing. This includes 27 EU member states plus 11 countries outside the EU who have adhered to the scheme.
[7] International Standard Classification of Education (ISCED) | UNESCO UIS – accessed on 24th February 2023.

4.5 LSBQ_e: Location and dominance

The original LSBQ was developed within the national context of Canada, which is reflected in at least one of its questions, namely question 13, which asks: "*Were you born in Canada? If no, where were you born? When did you move to Canada?*". This question has been replaced by a set of questions requesting information about a participant's place of birth as well as potential periods spent in different geographic areas / language communities:

Fig. 12: Eliciting information about participants' place of birth and periods spent in other locations in a more neutral, non-presumptive manner.

This change allows researchers to gather information about the time participants may have spent in different locations, without specifically tying the LSBQ_e to any particular national experience. This also avoids introducing a potentially biasing presumption of nationality, which may lead to unwanted results if "Canada" in question 13 were simply replaced with the name of some other country where the research is carried out, especially in contexts involving either highly international or regional, minority and/or minoritized language communities.

The wording for question 14 ("*Have you ever lived in a place where English is not the dominant communicating language?*") was also changed, asking instead whether a participant has lived in a place other than their birthplace for any significant period of time (i.e., more than 6 months). This elicits the same data but avoids both the naming of a specific language (whether "English" or some other language) and the use of the phrase "not the dominant communicating language", for two reasons. First, it was felt that this is a rather technical term that many participants may not be familiar with, as suggested by the fact that some organisations report the phrase in their glossaries of specialised vocabulary[8]. Further, among participants who are familiar with the phrase, there may be radically different interpretations, seeing as even researchers disagree on what 'language dominance' is and how it should be measured (e.g., Treffers-Daller, 2011). It therefore seems wise to place the burden of identifying dominance on the researchers, who have the option of specifying their precise assumptions for analysis, rather than on the participants. For this reason, the LSBQ_e allows researchers to collect raw data about where a participant has lived for periods longer than six months, leaving the researchers to draw conclusions about how that location relates to the concept of language dominance when considered together with the other measures obtained.

4.6 LSBQ_e: Continuous data

The original LSBQ uses a combination of sliders (e.g., question 16.1) and ordinal, "Likert-type" choices (e.g., question 16.2) for many of its items. This creates a degree of inconsistency across responses, as the sliders are given a range of 0-10 and offer potentially infinitely fine-grained response options, while the Likert-type responses only include five options.

---

[8] See for example https://www.ldonline.org/glossary and https://www.colorincolorado.org/ell-basics/ell-glossary - accessed on 24th June 2022.

In the LSBQ$_e$ we implemented sliders in all questions that allowed them as a possible option, each accompanied by labels that inform the participant about the type of response required. Likert-type items that appeared in the original LSBQ were therefore replaced.
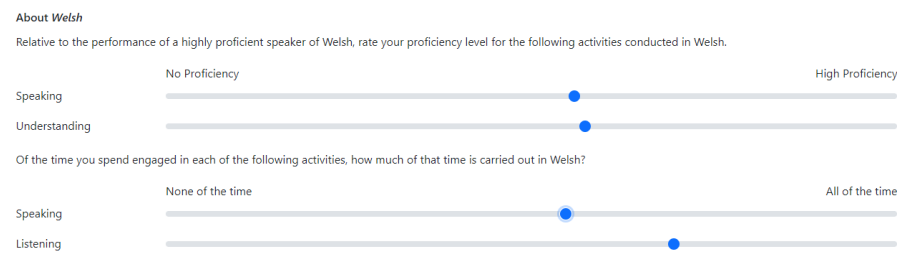

Fig. 13: All items requesting responses along a continuum have been implemented with continuous sliders.

Being a digital implementation, the LSBQ$_e$ has the advantage of allowing sliders to be measured as double-precision floating point numbers between 0 and 100, enabling the collection of potentially much finer-grained data without burdening the participant with such large number of options. The numerical value of the participant's choice is calculated behind the scenes, and all a participant needs to do is move the sliders across the polar continua. The result is data that is to some extent insulated from the "coarseness" bias (Aguinis, Pierce, & Culpepper, 2008; Russell & Bobko, 1992) that afflicts ordinal scales, and particularly 5-point Likert scales.

For ease of handling and inspection, and due to limitations when data is recorded with a keyboard rather than a pointing device,[9] the data is recorded as a real number between 0 and 100 (e.g., 34.6174789000817), but can be easily rescaled to whatever may be convenient in data analysis (e.g., to a 0-1 range).

4.7 LSBQ$_e$: Life stages

Questions relating to language use at different life stages of a participant (question 18-18.4; headed "*Please indicate which language(s) you most frequently heard or used in the following life stages, both inside and outside home*") is another example where the North American reference frame of the original LSBQ comes to bear. For instance, the labels "Infancy", "Preschool age", "Primary School age" and "High School age" do not transfer well to participants in countries outside of North America (e.g., in the UK, Italy, Belgium, or Germany). This is resolved by choosing appropriate labels for corresponding stages depending on the translation/localisation (for instance, our UK English localisation uses "Infancy", "Preschool age", "Primary age", "Secondary age"). Importantly, we have assumed here that cultural correspondence to early life stages is to be preserved, rather than exact age ranges, as for example the age at which children typically begin schooling differs across countries and cultures.

Do note, however, that a North American version (or any other version) can be easily integrated through an added localisation, which can then be selected depending on location of the study (see the User Guide for details on how to do this).

---

[9] To the best of our knowledge, all current web engines implement a 1/100th step as the default when a continuous range control is adjusted with the arrow keys of a keyboard. This would of course result in much coarser data, but a range of 0-100 here still ensures a reasonably continuous measure.

4.8 LSBQ$_e$: Interactions and interlocutors

The section on interlocutors (i.e., section 19 in the original LSBQ, headed by "*Please indicate which language(s) you generally use when speaking to the following people*") has an additional item "Children", as this is linguistically relevant in cases where the participants are adults. This, as well as other items, can then be skipped by the participant by ticking N/A in cases where it does not apply (e.g., if the participant has no children).

Fig. 14: Partial example of the Community Language Use Behaviour section of the LSBQ$_e$ showing how to indicate that the item "Children" is not applicable, which is then ignored in validation. Compare this to the accidentally omitted item "Grandparents", which is flagged up with a prompt (in red).

4.9 LSBQ$_e$: Open question

At the end of the LSBQ$_e$ we've added an open question asking participants to provide any additional information that they believe may be relevant. While this question can be left blank, it may be useful in some cases, for example if a researcher is piloting the questionnaire with a cohort that might not fit the standard LSBQ or LSBQ$_e$, but they are not sure what additional questions or criteria to include.

Fig. 15: The LSBQ$_e$ ends with an open question.

5.   Digital Informed Consent

The L'ART Research Assistant provides a digital consent tool for those who wish to forego pen-and-paper methods altogether. A consent form structure is built into the L'ART Research Assistant, while also allowing individual researchers to adapt it to their needs.

5.1   The Generic consent form

When selecting "Informed consent" from the side menu, an option for a "generic consent form" will appear in the drop-down field.
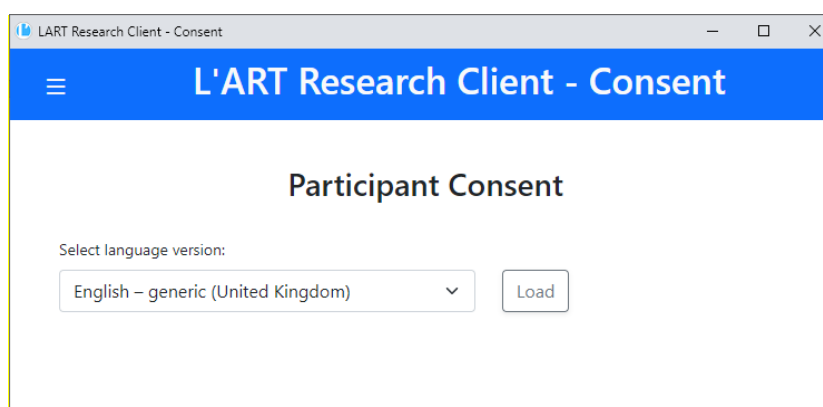

Fig. 16: Selecting the generic consent option for the British-English implementation

This will open a generic consent form that users can adapt and use in their own research. However, being a generic form, it does not have any specific sections that may be required for specific research purposes, particularly in cases where consent is required for more than just a background questionnaire (e.g., where a battery of tasks may be involved). If this is the case, users can produce and add their own form. Guidance on how to do this is provided in the User Guide.

6.   Explicit Attitudes: the AToL-C

Having been originally developed for the investigation of language attitudes in bilingual communities, the L'ART Research Assistant offers a digital implementation of the Attitudes towards Languages Questionnaire (AToL; Schoel, Roessel et al., 2012). Our implementation contains the full 15 bipolar pairs of adjectives of the AToL, but uses continuous digital scales as opposed to the 5-point Likert scale used in the original (see section 4.6 for an overview of why continuous scales were adopted in the L'ART Research Assistant) – we thus chose to refer to this specific implementation as the AToL-C ("C" for continuous). The 15 bipolar pairs represent three factors of language perception: *Sound* (e.g., harsh-soft), *Structure* (e.g., precise-vague) and *Value* (e.g., beautiful-ugly), thus fully replicating the original AToL. Version 0.4.1 of the app offers AToL-C implementations in four languages: English, German, Italian and Welsh[10]. Table 1 below shows the adjective pairs for each of these languages.

---

[10] The Welsh translation was provided by Bangor University translation services
https://www.bangor.ac.uk/canolfanbedwyr/cyfieithu.php.en

German	English	Italian	Welsh
logisch – unlogisch	logical - illogical	logico – illogico	rhesymegol - afresymegol
stillos – stilvoll	inelegant – elegant	non elegante – elegante	anghywrain - cywrain
stockend – fließend	choppy – fluent	frammentato – scorrevole	toredig - rhugl
eindeutig – missverständlich	unambiguous – ambiguous	chiaro – ambiguo	diamwys - amwys
anziehend - abstoßend	appealing – abhorrent	attraente – ripugnante	apelgar - atgas
stukturlos – sturkturiert	unstructured – structured	non strutturato – strutturato	distrwythur - strwythuredig
genau – ungenau	precise – vague	preciso – vago	manwl - amhendant
hart – weich	harsh – soft	duro – morbido	llym - meddal
flüssig – abgehackt	flowing – abrupt	fluido – brusco	llifo - swta
schön – hässlich	beautiful – ugly	bello – brutto	prydferth - hyll
systematisch – unsystematisch	systematic – unsystematic	sistematico - non sistematico	trefnus - di-drefn
angenehm – unangenehm	pleasant – unpleasant	piacevole – spiacevole	dymunol - annymunol
geschmeidig – rau	smooth - raspy	liscio – ruvido	llyfn - cryg
plump – anmutig	clumsy - graceful	goffo - aggraziato	trwsgl - gosgeiddig
eckig – rund	angular – round	spigoloso - arrotondato	onglog - crwn

The task begins with a screen that reads, "{The X language} is…" (where "{X}" is replaced with English, German, Italian, etc., depending on selection at the start) as shown in the excerpt in Fig. 17 below.

Fig. 17 – Section of the AToL-C screen for English, showing three of the fifteen bipolar adjective pairs.

Like all tasks in the L'ART Research Assistant, users can easily add their own language versions by following instructions in the section on [Localisation and Adding Translations](#).

7.   Speaker evaluation paradigm: from MGT and VGT to AGT

The speaker evaluation paradigm involves exposing participants to different audio-recorded guises representing different language varieties that the participant must rate (Dragojevic & Goatley-Soan 2022). The L'ART Research Assistant allows users to implement the speaker evaluation paradigm using the matched-guise technique (MGT; Lambert, Hodgson, Gardner & Fillenbaum 1960) as well as the verbal guise technique (VGT; Markel, Eisler & Reese 1967). We therefore label this implementation 'AGT', for 'Audio Guise Task', which users can run as either MGT (by adding recordings from six speakers, each performing in two different languages/varieties) or as VGT (by adding recordings from twelve different speakers, six per language/ variety). See the [AGT section](#) of the [User Guide](#) for details.

The current version offers fully functional MGT versions of the AGT in for three language pairs: Welsh-English, Moselle-Franconian-German, and Lombard-Italian. Templates for custom implementations in English, German and Italian are also included (ready to use upon addition of relevant recordings) and further language versions can be easily added by following instructions in the User Guide section on [Localisation and Adding Translations](#).

7.1 AGT: Experimental Design

Following standard procedure from both the MGT and VGT, the L'ART Research Assistant implementation (i.e., the AGT) asks participants to complete an evaluative questionnaire rating speakers on a range of traits. The AGT includes eighteen different traits that were selected based on the list used in the original MGT (Lambert et al. 1960) combined with more recent developments from attitudinal studies that focused on minority language contexts (Echeverria 2008; Loureiro-Rodriguez et al. 2013; Soukup 2013; Real Academia Galega 2014; Price & Tamburelli 2020). The trait list was constructed by the L'ART research team, which featured at least one member with linguistic expertise in each of the languages under consideration.

The trait selection was finalised via a two-step process. First, we produced a list of traits based on the list used in the original MGT (Lambert et al. 1960), grouping them into *solidarity* and *status* dimensions to provide consistency and maintain a balance of traits with the original MGT. Secondly, we reviewed further lists of traits that had been used in studies on regional/minority language contexts (Echeverria 2008; Loureiro-Rodriguez et al. 2013; Soukup 2013; Real Academia Galega 2014; Price & Tamburelli 2020) and brought in those that we thought were particularly relevant to the three European contexts that the L'ART Research Assistant was originally developed for. We filtered out any item that did not transfer well across different cultures and linguistic communities, as our final list had to cater for rather different sociolinguistic situations, where the regional/minority languages have sometimes radically different degrees of recognition and/or hierarchical relations with the majority language. For example, the trait "amusing" was included from Echevarria (2008) as it is likely to provide information on the type of attitude when describing a regional/minority language speaker (e.g., Ferrer & Sankoff, 2003). Similarly, "international" was included as it may elicit a potentially typical attitude when describing a majority language speaker (e.g., Byrne, Bertran, & Tudela Isanta, 2022). Through this process we collated an initial longlist of traits, as shown in Table 2. Traits that we deemed to be similar were included in the same cell to be

later finalised for one trait, and *status* vs *solidarity* and *valence* categories of the traits were recorded in order to feed into the decision-making process.

Table 2 - Initial long list of traits for AGT

Trait	Status / Solidarity	Valence
confident	*status*	*positive*
amusing / funny	*solidarity*	*negative*
open-minded	*status?*	*positive*
improper	*status*	*negative*
caring	*solidarity*	*positive*
attractive	*status*	*positive*
refined	*status*	*positive*
trustworthy	*solidarity*	*positive*
efficient	*status*	*positive*
has a sense of humour	*solidarity*	*positive*
ignorant	*status*	*negative*
boring	*solidarity*	*negative*
conservative	*status*	*negative?*
polite	*status*	*positive*
ambitious	*status*	*positive*
cosmopolitan / international	*status*	*positive*
rustic	*status?*	*negative*
arrogant / smug	*solidarity*	*negative*
cool	*status*	*positive*
cultivated	*status*	*positive*
lazy	*status?*	*negative*
artificial	*solidarity*	*negative*
intelligent	*status*	*positive*
kind	*solidarity*	*positive*
influential (leader)	*status*	*positive*
open	*solidarity*	*positive*
proud	*solidarity*	*positive?*
progressive	*?*	*positive?*
hard-working	*status*	*positive*
likeable	*solidarity*	*positive*
educated	*status*	*positive*
faithful	*?*	*positive?*
practical	*?*	*positive*
friendly	*solidarity*	*positive*
honest	*solidarity?*	*positive*
competent	*status*	*positive*
natural	*solidarity?*	*positive*
traditional	*status?*	*positive?*
pretentious	*solidarity*	*negative*
uptight	*?*	*negative*

Next, we considered the translatability of all items on the longlist. It was imperative that each item on our final list of traits would be translatable across all languages for which the L'ART Research Assistant was originally developed. For this reason, certain traits were adapted to a word with a clearer definition in all languages. For example, "cosmopolitan" was avoided due to a possible ambiguous overall meaning, as the connotations attached to being cosmopolitan may not be inherently positive, for instance, in the sense of relating to high society, rather than having a worldwide scope or outlook. Similarly, being "cosmopolitan" could relate to a potentially politically sensitive issue that could polarise participants, for example if interpreted in terms of favouring globalization and suggesting a "global vs local" opposition. The word "cosmopolitan" was therefore changed to "international" in order to better represent a trait that was closer to a sense of worldwide scope and outlook.

Other traits from previous research were excluded due to discrepancies they would create with the overall list, e.g., "a leader may appear idiosyncratic as the only phrasal noun in a list of adjectives. Furthermore, the trait list was made relevant to the research sample that the AGT was originally developed for, therefore we excluded traits from previous studies such as Price & Tamburelli (2020) which were aimed towards speakers outside our sample age group, e.g., 'goody-two-shoes' or 'likes a laugh' were not used, as these were specifically aimed at adolescents while our implementation of the AGT was developed for research on adults.

Several iterations of the above process led to a final list of eighteen traits, reported in Table 3 below with the respective classification for *status* vs *solidarity* and *valence.*

Table 3 - Final list of 18 traits for AGT

Trait	Status / Solidarity	Valence
amusing	*solidarity*	*negative*
open-minded	*status*	*positive*
attractive	*status*	*positive*
trustworthy	*solidarity*	*positive*
ignorant	*status*	*negative*
polite	*status*	*positive*
ambitious	*status*	*positive*
international	*status*	*positive*
cool	*status*	*positive*
intelligent	*status*	*positive*
influential	*status*	*positive*
likeable	*solidarity*	*positive*
educated	*status*	*positive*
friendly	*solidarity*	*positive*
honest	*solidarity*	*positive*
competent	*status*	*positive*
natural	*solidarity*	*positive*
pretentious	*solidarity*	*negative*

Note that it remains in principle possible for the user to adapt the list of adjectives via the general translation mechanism made available through the task localisation options, though an important current limitation in this is that when e.g., the adjective "amusing" is changed to "funny" in a localisation file, the *data file* with the response will still encode this as a score for *amusing*. An improved mechanism for arbitrary sets of adjectives may be implemented in future versions of the app.

7.2 The AGT: implementation features

Similar to what was discussed for the LSBQ_e implementation (see section 4.1), implementing the AGT digitally on the L'ART Research Assistant has several benefits concerning the effectiveness of executing various elements of the methodology.

7.2.1 Guise Timing

Central to the AGT are its guises. Guises may be as long or as short as the researcher requires, but it may be logical to align the length of the guise with the typical amount of time it takes a participant to work through the traits for a given guise. As the AGT set-up does not allow participants to relisten to a guise (see section 3.1.1), aligning the length of the guise with the time it takes to provide a rating for all eighteen traits ensures that participants rate speakers as they hear the guise, rather than forcing them to rate on the basis of what they may remember. For an 18-trait list, we found the optimal length of a guise to be between 1 minute and 1 minute and twenty seconds.

In order to avoid confound effects (e.g., participants listening to the same guise more than once, and thus potentially altering their attitude responses), the AGT for L'ART Research Assistant only allows participants to listen to each guise once. This is linked to the generic L'ART Research Assistant feature that does not permit participants to go back to previous screens on any of the tasks (see section 3.1.1). However, should the researcher wish to restart a guise, for instance because of a technical issue, such as trouble with headphones while a guise is playing, the researcher can unlock the app, reload the screen and re-initiate playback by clicking on the playback button (see User Guide: Locking & unlocking the app for details).

7.2.2 Progress

The AGT for L'ART Research Assistant implementation includes a progress bar indicating the elapsed time of the playing guise. The progress bar is located above the trait list and may aid participants with timekeeping whilst rating the guises (see Figure 1717 & Figure 1819).

Some participants may finish rating the guises before the guise has finished; however, the L'ART Research Assistant will not allow the participant to press "*Next*" to continue until the guise has been playing for at least 30 seconds or half of its total length (whichever is shorter), as illustrated in Figure 18 & Figure 20. This goes towards preventing participants from potentially rating guises too quickly, e.g., before they have had a reasonable chance to develop an opinion towards the guise.
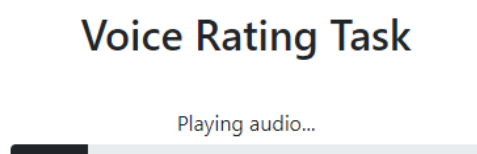

Figure 17 - AGT progress bar less than halfway through a guise

Figure 18 – Corresponding "next" button of a guise less than halfway through: the button is deactivated (i.e. "greyed out")
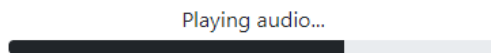

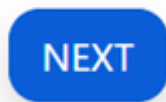*Figure 19 - AGT progress bar more than halfway through a guise*

Figure 20 - Corresponding "next" button of a guise more than halfway through: the button is now active (dark blue)

7.2.3 Practice Guise

The L'ART Research Assistant also implements a practice guise in order to familiarise participants with how the AGT works.

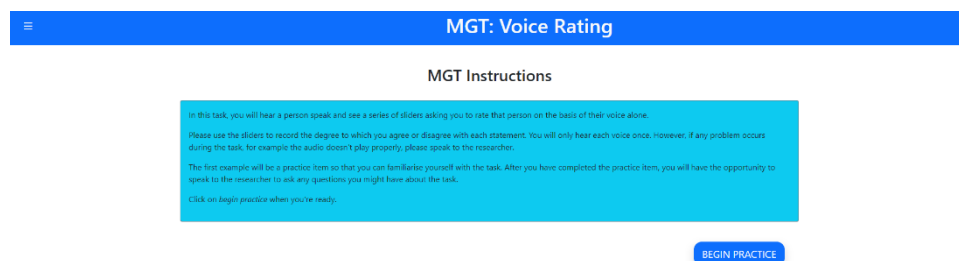*Figure 21 - Instruction screen notifying participants about the practice guise*

7.3 Stimulus Orders

As mentioned above, one distinguishing feature of the AGT is that it can produce either an MGT or a VGT. The AGT requires eight experimental audio files, and four filler files. If the researcher requires an MGT, they must prepare audio files of four speakers, with two recordings per speaker (one for each of the two language varieties). If the researcher requires a VGT set-up, they must prepare audio files of eight speakers, four speaking in one language variety, and another four speaking in another language variety. For information on loading files for the AGT please see the AGT section of the User Guide.

Guises appear in a pseudo-randomised order which ensures that recordings from the same speaker are always spaced maximally apart, something that is necessary in order for the AGT to allow an MGT presentation. The order of fillers (regardless of filler language), and which language variety is presented first (keeping alternation constant) is also randomised. Below are two potential pseudo-random orders generated by the L'ART Research Assistant's implementation of the AGT set up for an MGT presentation.

Table 4: Example of a pseudo-random order using Welsh and English as working languages

Speaker	Language Variety	Example
F1	Either	Filler: English
S1	Var1	Speaker 1, Welsh
S2	Var2	Speaker 2, English
F2	Either	Filler: Welsh
S3	Var2	Speaker 3, English
S4	Var1	Speaker 4, Welsh
F3	Either	Filler: Welsh
S1	Var2	Speaker 1, English
S2	Var1	Speaker 2, Welsh
F4	Either	Filler: English
S3	Var1	Speaker 3, Welsh
S4	Var2	Speaker 4, English

Table 5: Example of a pseudo-random order using Moselle-Franconian and German as working languages

Speaker	Language Variety	Example
F3	Either	Filler: German
S3	Var2	Speaker 3, Moselle-Franconian
S4	Var1	Speaker 4, German
F2	Either	Filler: Moselle-Franconian
S1	Var1	Speaker 1, German
S2	Var2	Speaker 2, Moselle-Franconian
F4	Either	Filler: German
S3	Var2	Speaker 3, Moselle-Franconian
S4	Var1	Speaker 4, German
F1	Either	Filler: Moselle-Franconian
S1	Var2	Speaker 1, Moselle-Franconian
S2	Var1	Speaker 2, German

8.   Memory Game

Finally, a simple memory game is also made available, which can be employed as a general distractor when running tasks in a series. This is particularly relevant for researchers investigating language attitudes, where it is It is standard practice not to fully disclose the aim of the study to participants in order to render attitudinal measures less direct (e.g., Pharao & Kristiansen 2019) and thus minimize acquiescence bias, where participants tend to give the response they believe the researcher is

looking for (Jackson & Messick 1965), and reduce the social desirability effect, where participants respond with the attitude they think is perceived as most desirable (Diekmann 2007).

For this purpose, the L'ART Research Assistant includes an implementation of a simple and widely available memory game (see Fig. 22) consisting of (n=20) cards presented on the screen.
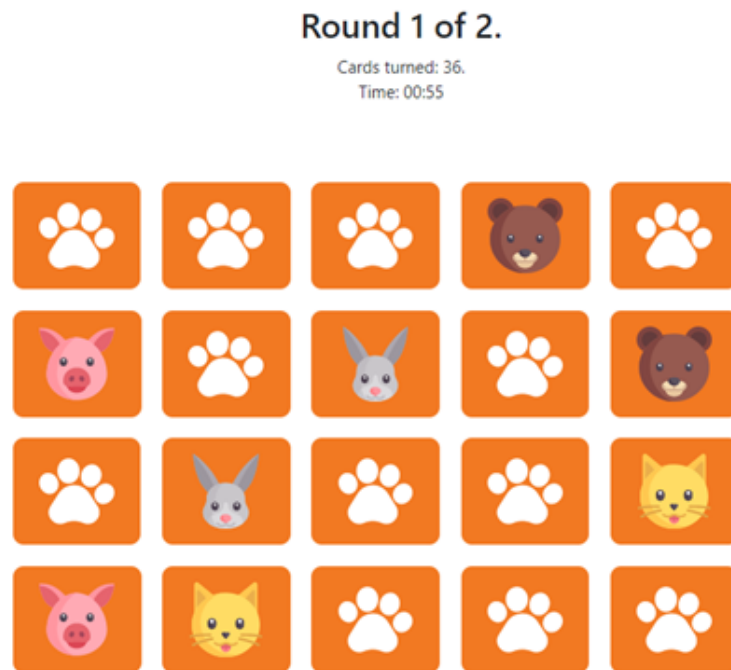

Figure 22 – Example set up of the memory game

The game requires participants to match all the identical card pairs in the shortest amount of time possible. The cards are faced down and only two cards can be revealed at a time, thus the participant must attempt to memorise card locations after revealing a pair.

Acknowledgements

References

Aguinis, H., Pierce, C., & Culpepper, S. 2008. Scale coarseness as methodological artifact: correcting correlation coefficients attenuated from using coarse scale. *Organizational Research Methods* *12*(4): 623-652

Anderson, J.A., Mak, L., Keyvani Chahi, A. and Bialystok, E., 2018. The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior research methods*, *50*(1), pp.250-263.

Bialystok, E. and Shorbagi, S.H., 2021. Subtle increments in socioeconomic status and bilingualism jointly affect children's verbal and nonverbal performance. *Journal of Cognition and Development*, *22*(3), pp.467-490.

Byrne, S., Bertran, A. and Tudela Isanta, A., 2022. Continuity and change: language ideologies of Catalan university students. *Journal of Multilingual and Multicultural Development*, pp.1-16.

Calvo, A. and Bialystok, E., 2014. Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, *130*(3), pp.278-288.

Dragojevic, M. and Goatley-Soan, S., 2022. Americans' attitudes toward foreign accents: Evaluative hierarchies and underlying processes. *Journal of multilingual and multicultural development*, *43*(2), pp.167-181.

Chakroun, B., 2010. National qualification frameworks: From policy borrowing to policy learning. *European Journal of Education*, *45*(2), pp.199-216.

Diekmann, A., 2007. Empirische sozialforschung. Grundlagen, *Methoden, Anwendungen*, 18.

Echeverria, B. 2005. Language attitudes in San Sebastian: The Basque vernacular as challenge to Spanish language hegemony. *Journal of Multilingual and Multicultural Development*, 26(3), 249–264.

European Commission. 2008. *Explaining the European Qualifications Framework for Lifelong Learning*. Luxembourg: Office for Official Publications of the European Communities.

Ferrer, R.C. and Sankoff, D., 2003. Identity as the primary determinant of language choice in Valencia. *Journal of Sociolinguistics*, *7*(1), pp.50-64.

Horwitz, S.M., Irwin, J.R., Briggs-Gowan, M.J., Heenan, J.M.B., Mendoza, J. and Carter, A.S., 2003. Language delay in a community cohort of young children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*(8), pp.932-940.

Jackson, P.W. and Messick, S., 1965. The person, the product, and the response: conceptual problems in the assessment of creativity. *Journal of personality*, 33(3), p.309-329.

Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology*, 60(4), 44.

Loureiro-Rodriguez, V., Boggess, M. M., & Goldsmith, A. (2013). Language attitudes in Galicia: Using the matched-guise test among high school students. *Journal of Multilingual and Multicultural Development*, 34(2), 136–153.

Manuel Galvin Arribas, J., 2016. Governance Dynamics and the Application of the Multilevel Governance Approach in Vocational Education and Training (VET) in the European Neighbourhood Countries: the case of the ENPI South region. *European Journal of Education*, *51*(4), pp.495-512.

Markel, N. N., Eisler, R. M., and Reese, H. W., 1967. Judging personality from dialect. *Journal of Verbal Learning and Verbal Behavior 6*(1): 33–35.

Naber, M., Vedder, A., Brown, S.B. and Nieuwenhuis, S., 2016. Speed and lateral inhibition of stimulus processing contribute to individual differences in Stroop-task performance. *Frontiers in psychology*, *7*, p.822.

Pharao, N. and Kristiansen, T., 2019. Reflections on the relation between direct/indirect methods and explicit/implicit attitudes. *Linguistics Vanguard*, 5(s1).

Price, A.R. and Tamburelli, M., 2020. Welsh-language prestige in adolescents: attitudes in the heartlands. *International Journal of Applied Linguistics*, 30(2), pp.195-213.

Russell, C. J. & Bobko, P. 1992. Moderated regression analyses and Likert scales: Too coarse for comfort. *Journal of Applied Psychology 77*(3): 336-342.

Schoel, C., Roessel, J., Eck, J., Janssen, J., Petrovic, B., Rothe, A., Rudert, S.C. and Stahlberg, D., 2013. "A ttitudes To wards L anguages"(AToL) Scale: A Global Instrument. *Journal of Language and Social Psychology, 32*(1), pp.21-45.

Soukup, B. 2012. Current issues in the social psychological study of "language attitudes": constructionism, context, and the attitude–behavior link. *Lang & Ling Compass*, 6(4), 212–224.

Treffers-Daller, J., 2011. Operationalizing and measuring language dominance. *International Journal of Bilingualism*, *15*(2), pp.147-163.

UNESCO Institute for Statistics. 2012. *International Standard Classification of Education: ISCED 2011.* Montreal: UNESCO Institute for Statistics.

Vergnaud, A.C., Touvier, M., Méjean, C., Kesse-Guyot, E., Pollet, C., Malon, A., Castetbon, K. and Hercberg, S., 2011. Agreement between web-based and paper versions of a socio-demographic questionnaire in the NutriNet-Santé study. *International journal of public health*, *56*(4), pp.407-417.