

Q-BEx Manual

**Cécile De Cat, Draško Kaščelan, Philippe Prévost, Ludovica Serratrice, Laurie Tuller,
Sharon Unsworth**

Manual	2
What information does the Q-BEx questionnaire collect?	2
How to register as a user?	3
How to create and customise a project?	3
Distributing your questionnaire	6
Accessing data files	6
Interpreting data files	9
Which of your data can the Q-BEx team access?	11

Manual

What information does the Q-BEx questionnaire collect?

The Q-BEx questionnaire contains seven modules, two of which are obligatory. The other five modules can be customised or excluded altogether to meet the individual needs of researchers or practitioners. There are two versions of the questionnaire: a caregiver version (filled in by a caregiver about their child) and a child version (filled in by a child about themselves). The child version is adequate for distribution to teenagers only. Note that teenagers might need some help with filling in the questionnaire on their own (e.g., remembering the date they started school, or knowing the education level of their caregivers). The questionnaire consists of the following modules and sub-modules:

1. **Background information** [obligatory].
2. **Risk factors** [obligatory]. This module is not distributed in the child version of the questionnaire.
3. **Language exposure and use** [optional]:
 - a. *Current estimates (detailed)* [required sub-module if the module is selected; alternative to *Current estimates (short)*]
 - b. *Current estimates (short)* [required sub-module if the module is selected; alternative to *Current estimates (detailed)*]
 - c. *Cumulative estimates* [optional]
 - d. *Weighted estimates* [optional]
 - e. *Age and place of first exposure* [required sub-module if the module is selected]
 - f. *Overheard speech at home* [optional]
4. **Language proficiency** [optional]:
 - a. *Speaking and listening proficiency in each language* [optional]
 - b. *Reading and writing proficiency in each language* [optional]
 - c. *Proficiency compared with other bilingual children* [optional]. This sub-module is not distributed in the child version of the questionnaire.
5. **Richness of linguistic experience** [optional]:
 - a. *Activities* [optional]
 - b. *Caregivers' education* [optional]
 - c. *Estimated diversity of speakers* [optional]
6. **Attitudes and satisfaction with child's language** [optional]
 - a. *Attitudes on language mixing* [optional]
 - b. *Preferred language* [optional]
 - c. *Satisfaction with child's speaking and understanding* [optional]
 - d. *Pressure (on caregivers) and (child's) willingness to speak* [optional]
7. **Language mixing** [optional]
 - a. *At home, between interlocutors* [optional]
 - b. *At home, within interlocutors* [optional]
 - c. *At home, types of mixing* [optional]


- d. *Outside the home, between interlocutors* [optional]
- e. *Outside the home, within interlocutors* [optional]
- f. *Outside the home, types of mixing* [optional]

See [this infographic](#) for a brief outline of the type of data that each module collects and the scores derived by the tool. In the next section we outline how you can register on our platform and use the questionnaire.

How to register as a user?

1. Go to <https://q-bex.org>
2. Click on the button *Register* found in the upper right corner of the Home page.
3. Provide your first name, last name, email address (academic or professional), and agree to terms and conditions.
4. Click *Submit*.
5. You will see the message *Thank you. You will receive an email when your registration has been approved*. If the automatic filter is unable to vet your email address, a member of the Q-BEx team will have to approve your registration manually (**which might take a couple of days**).
6. You will receive an email with your username and a link to reset your password. Click on that link.
7. You will be taken to a website with a suggested password. Delete this password and set up a new one. Click *Reset password*.
8. You will see the message *Your password has been reset. Log in*.
9. Click on *Log in*. Enter your username (which has been emailed to you) and your new password. Click *Remember me* (if you wish to) and then *Log in*. A double authentication will be required.
10. You will be taken to the platform where you can create projects. Each project will contain a link to the questionnaire which you can send to your respondents. In the next section, we explain how to create a project.
11. Once you register and log in for the first time, you can access the questionnaire platform by going to <https://q-bex.org> and clicking *Log in* in the upper right corner of the Home page.

How to create and customise a project?

1. Once you are registered and logged into the website, click on the small Q-BEx label in the upper left corner of the Home page → 
2. Click on *Projects* (from the left side menu on the screen).

3. Under *Projects*, click *Add New* and name your study in the box entitled *Add title*. You can also create a new project by clicking on *Add New* - just under *All Projects* in the upper left corner of the screen.
4. In the section *Project* (just under the study title), the following parameters have to be set by the researcher/practitioner:
 - *Respondent country of residence*: the country you choose will automatically appear in relevant questions throughout the questionnaire. You can also tick *Respondent can select country of residence*, in which case each respondent will be asked about where the child lives and their answer will overwrite whatever you specified as the respondent's country of residence (selecting this option is convenient when in the same survey you have respondents living in different countries).
 - *Respondent type*: select whether you will be distributing the questionnaire to caregivers (about their children) or to children (about themselves).
 - *Questionnaire language*: select the language in which you want the questionnaire to be distributed. If you are using a version other than English, we strongly recommend asking a speaker of that language to go through the questionnaire in that language before you use it with the target population. In case you identify any issues with translations, please let us know at qbex@leeds.ac.uk
 - *Languages of interest*: list all the languages that you expect your participants to speak/understand (these will appear as drop-down menu options in the relevant question). Note that in the actual questionnaire, the respondents will also be able to choose the option *other* and specify a language that you have not included.
 - *Common language of interest*: if you wish, you can specify a language here that the respondents will always be asked about first (this is convenient when you are aware of a language that all children in your sample are likely exposed to, e.g. English in the UK). In the output, this language will consistently show as Language 1 (which will be useful if you need to assign a special status to that language – e.g., “societal language”).
5. In case you want to enable your respondents to abandon the survey before finishing and save the answers they have given, select the box *Respondent can save and abandon the questionnaire* in the section *Allow respondents to save and abandon the questionnaire?* Note, however, once the respondents abandon the questionnaire, they will not be able to resume it later.
6. In the section *Purpose of your project*, you must provide the following details:
 - a. *What will the questionnaire data be used for?* Select one or more of the given options
 - b. *Are you targeting atypical development?* If yes, you will be asked to specify what type of atypical development you are interested in.This information will help us keep track of the use of the Q-BEx questionnaire and will inform its future development.
7. If you plan to include the *Language Mixing* module, particularly the sub-modules about types of mixing, in the section *Examples*, enter required examples of language mixing

(code-switching). These examples will then appear as illustrations in relevant questions of the *Language mixing* module. Make sure that the three examples are (a) a one-word switch (e.g., Puedes darme la *drink*), (b) a two-three-word switch (e.g., I really like your *nueva camisa*), and (c) a sentence switch (e.g., Podemos comenzar ahora. *Or we can wait a bit longer*). For each example, we recommend the use of single-clause sentences as examples rather than multi-clause sentences. For instance, in case of a two-three word switch, we encourage using a single-clause sentence such as “One of my best friends lives *en esa calle*” rather than a multi-clause sentence such as “One of my best friends lives *en esa calle* and she also works nearby”. This will make the switch and INTRA-clausal switch (i.e., within a clause) rather than an INTER-clausal switch (i.e., between clauses). Whatever your choice of examples, make sure to report them when disseminating your data. This will allow the comparability of studies using different mixing examples across languages. The examples you provide will be used for types of mixing at home and outside home, so make sure that you provide examples which could be heard in both contexts. We strongly encourage the use of examples from spontaneous speech corpora of relevant communities if available.

8. In the section *Modules*, select those modules **and** sub-modules that you want to distribute. *Background information* (18 questions) and *Risk factors* (3 questions) are always obligatory. If you select *Language exposure and use*, two of its submodules are automatically included: *Current estimates (detailed or short)* and *Age and place of first exposure*. If you select *Current estimates (short)*, its alternative sub-module *Current estimates (detailed)* will be unselected. Additionally, the sub-module *Weighted estimates* will not be possible to distribute either. Note that the *Risk factor* module and the sub-module about *Proficiency compared with other bilingual children* are not distributed in the child version of the questionnaire even if they are selected.
9. If you are a practitioner (e.g., a teacher, a Speech and Language Therapist) and you would like to distribute a shorter version recommended by the Q-BEx team, click the button **Select recommended modules** on the right side of the screen, which will automatically select relevant modules and the sub-modules for the short version of Q-BEx. You can then modify things as required.
10. If you would like to set all the above parameters as in a version of the questionnaire that you built previously as a separate project, you can use the **Clone** option on the right side of the screen and select a project from the drop down menu that you would like to clone.
11. When you have set all the parameters, click *Publish* (upper right part of the screen). Note that once you publish a project, you can't make any updates. In case you require a different questionnaire, you'll need to create a new project.
12. **Each project expires two years after its publication, and all its data is deleted from the server. Please make sure to download and safely store your data before the expiry date of the project. A month before the project expires, you will receive an automated email reminding you of the project expiry date.**
13. After publication, in the section *Questionnaire*, you can find the questionnaire URL. This link is to be sent to your respondents.

14. Once a respondent has filled in the questionnaire, their answers will appear in the sections *Respondents*, *Data*, *Tidy response data and calculations*, *Raw data and reports*, and *Time matrix data*. Please see the sections *Accessing data files* and *Interpreting data files* of the manual for instructions on how to download and interpret the data.
15. When you click on *Projects* or *All projects* (upper left corner), you will go back to the list of the projects which you created. Next to each project, you will see its creation date, its deletion date, a type of respondent, and the language of distribution. **Please make sure to download and safely store your data before the expiry date of the project. Otherwise, the data will be lost.**
16. In case you have any difficulties using the questionnaire, write to us at qbex@leeds.ac.uk. After the end of the funded part of the project (October 2023), it will be monitored a couple of times per week (except during holidays).

Distributing your questionnaire

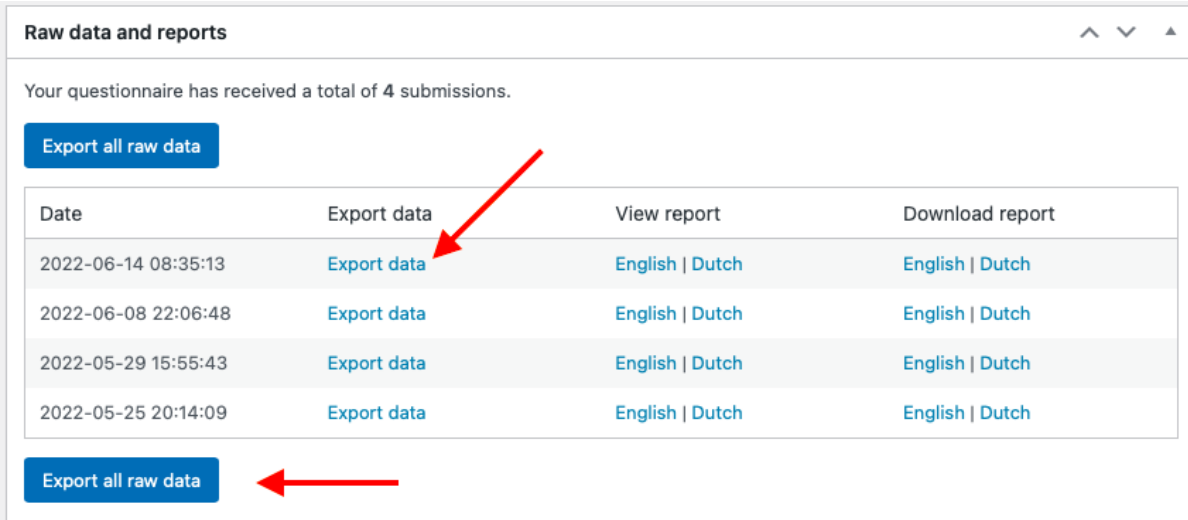
Once you publish a project, in the section *Questionnaire*, you can find the questionnaire URL. This link is to be sent to your respondents. You can also include this link into a larger, wrapper questionnaire by using a platform such as Google Form. This approach can be convenient for a variety of research projects, as the Google Form can include: the study information details, the consent form, any other questions which you need to ask, and finally the link to the Q-BEx questionnaire. The link will take the respondents to the Q-BEx platform page where they can complete the modules which the researcher/practitioner included. The questionnaire has been built primarily for use on computers and laptops. Tablets might also be convenient for distribution. Smartphones **are not** recommended as the size of their screen is not compatible with a series of questions and their visual outline. We also note that there is no paper version of the Q-BEx due to the questionnaire features which work only if distributed electronically.

Accessing data files

On the page of each project that you create, you will be able to access the following output files:

- Raw data (anonymised)
- Individual child reports (with names)
- List of respondents' IDs and names
- Tidy response data and calculations (anonymised)
- Time matrix data (if you distributed the sub-module *Weighted estimates*) (anonymised)
- All of the above in a single zipped folder

The output file containing the **raw data** can be found under the section *Raw data and reports*. You can either download the data (as a single .csv file) from all respondents by clicking *Export all raw data* or you can download the data for each participant individually by clicking *Export data* next to each submission (see Figure 1). Note that these files contain only raw data exactly as collected by the questionnaire without any data processing (i.e., no derived scores or tidying the data). All the data accessible in raw data files are anonymised as the names of children are automatically replaced by *respondent_id* (a variable present in the raw data file).



Raw data and reports

Your questionnaire has received a total of 4 submissions.

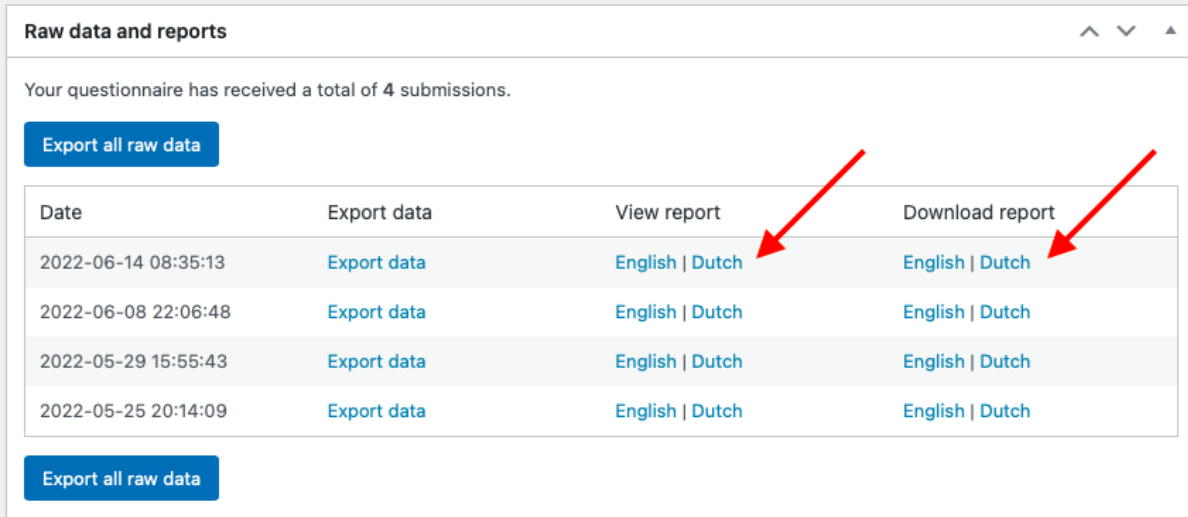
Export all raw data

Date	Export data	View report	Download report
2022-06-14 08:35:13	Export data	English Dutch	English Dutch
2022-06-08 22:06:48	Export data	English Dutch	English Dutch
2022-05-29 15:55:43	Export data	English Dutch	English Dutch
2022-05-25 20:14:09	Export data	English Dutch	English Dutch

Export all raw data

Figure 1. Exporting raw data

In the same section, you can view or download **reports** aimed at practitioners, which contain some of the data about participants' language history. Note that these files are not anonymised (i.e., they contain participants' names). You can access the reports in either English or Dutch (by clicking on the name of the language version which you prefer). See Figure 2 on how to access the report files.



Raw data and reports

Your questionnaire has received a total of 4 submissions.

Export all raw data

Date	Export data	View report	Download report
2022-06-14 08:35:13	Export data	English Dutch	English Dutch
2022-06-08 22:06:48	Export data	English Dutch	English Dutch
2022-05-29 15:55:43	Export data	English Dutch	English Dutch
2022-05-25 20:14:09	Export data	English Dutch	English Dutch

Export all raw data

Figure 2. Exporting report files

If you want to link individual data to each child, download the **list of respondents' IDs and names** from the section *Respondents* by clicking *Export respondent names* (see Figure 3).

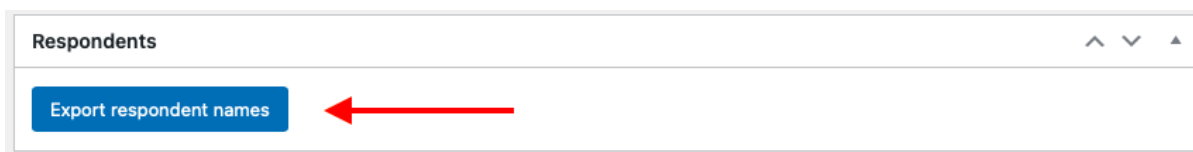


Figure 3. Exporting respondents' names

In the section *Tidy response data and calculations*, you can download **tidy raw and calculated data** for all participants by clicking *Export tidy response data and calculations* (see Figure 4).

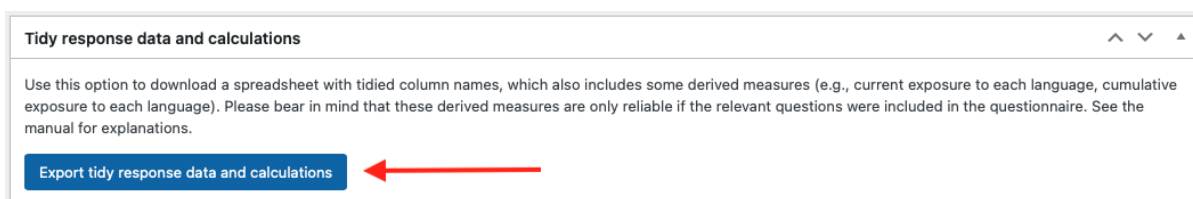


Figure 4. Exporting tidy response data and calculations

In case you distributed the sub-module *Weighted estimates*, you will be able to export the **time matrix data** by clicking *Export time matrix data* (see Figure 5). This file contains the data exactly as collected in the timetable questions (i.e., which hours of the day a child usually spends with each individual or in each context). Note that these data are not included in the raw data spreadsheet.

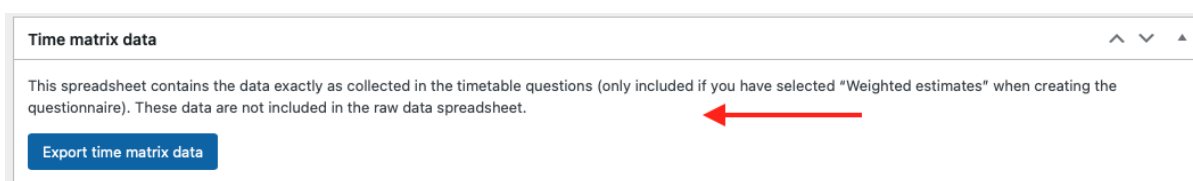


Figure 5. Exporting time matrix data

If you would like to export **all of the above data and reports** in English in a single folder, click *Export all data and reports* in the section *Data* (see Figure 6).

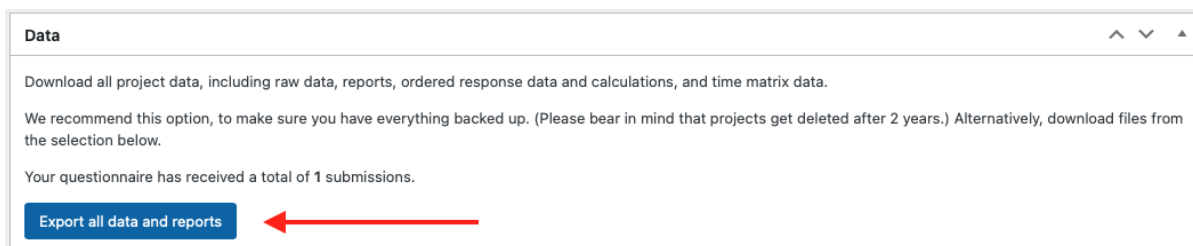


Figure 6. Exporting all data and reports in a single file

Interpreting data files

In order to interpret the data, you will require the *Output interpretation guide* excel file which can be downloaded from the [resources on our website](#). The *Output interpretation guide* file contains four sheets: raw data, tidy data and calculations, respondents, report, and time matrix data.

The sheet *raw data* contains variables necessary for the raw data interpretation (see Figure 7). For ease of interpretation, columns E and F contain the questions (the caregiver and the child version respectively), while column A contains the internal reference for each question (or NA if the data was generated automatically). Columns B and C indicate which module and sub-module each question belongs to. In column D, you can find the variable names assigned to each question in the output file. Column G contains possible responses as they will appear in the output file, while column H clarifies what each of those values stands for. Finally, column I contains some additional notes to help the interpretation. Note that throughout the *Output interpretation guide* file, we use square brackets around numbers when these stand for variables. For instance, the answer(s) to question Q.72 (“Please list all the languages that the child speaks and/or understands - no matter how well or how often.”) will be stored in the output file as *language 1* (if only one language is specified), and *language 2* (and possibly *language 3*) if more languages are specified by the respondent. Column D of the *Output interpretation guide* file only mentions these variables as *language [1]*, or *adult [1]* or *child [1]*, but they can be instantiated by several columns in the actual output file (depending on the number of languages, adults or children mentioned in the responses).

	A	B	C	D	E	F	G	H	I
1	question reference	module	sub-module	column name(s) in data output	question (caregiver)	question (child)	scale wording (data output)	scale wording (as seen by respondents)	notes
	NA	NA	NA	country of residence	NA	NA	abbreviation of a country or name of a country	NA	This country is normally predetermined by the superuser as a country of residence for all participants. However, if the superuser includes CQ.18 in the questionnaire (in which participants can specify themselves where the child lives), the data in this column will be overwritten by whatever the participant responded to CQ.18.
2									
3	CQ.15	Background information	NA	respondent_id	Please enter the child's name and surname.	Please enter your name and surname.	name replaced by an ID (a number)	insert	
4	NA	NA	NA	start_time	NA	NA	yyy-mm-dd hh:mm:ss	NA	These variables are meta data collected for every respondent without asking them any questions
5	NA	NA	NA	end_time	NA	NA	yyy-mm-dd hh:mm:ss	NA	
6	NA	NA	NA	version	NA	NA	version of the questionnaire [currently 8.1.0]	NA	
7	Q.72	Background information	NA	language [1]	Please list all the languages that the child speaks and/or understands - no matter how well or how often.	Please list all the languages that you speak and/or understand - no matter how well or how often.	names of a languages	select from a list of languages	
8	Q.73	Background information	NA	date of birth	What is the child's date of birth?	What is your date of birth?	yyy-mm-dd	insert date: dd/mm/yyyy	
9	Q.75	Background information	NA	gender	What is the child's gender?	What is your gender?	gender-male gender-female gender-non-binary gender-other [it will show what the respondent typed in]	male female non-binary other: [please specify]	
10									
11									
12									
13							gender-prefer-not-to-specify	I prefer not to specify	

Figure 7. Output interpretation, raw data

The sheet *tidy data & calculations* lists the variables that appear in the output file *Tidy response data and calculations* (see Figure 8). While column A lists the names of these variables as in the output file, column B specifies whether the variable is a raw or a calculated measure. Columns C specifies the unit measure of calculated variables as well as an explanation

or a reference to how the calculation has been derived. **Column D clarifies when you can rely on a certain calculated variable, which primarily depends on which modules and sub-modules were included in the survey.** Finally, in columns F-L, you can find additional notes to help you interpret the *Tidy response data and calculations* output.

Variable name	Raw or calculated measure	Calculation explanation or reference	When can you rely on this calculation?
Country of residence	raw	na	na
Respondent ID	raw	na	na
Start time	raw	na	na
End time	raw	na	na
Version	na	na	na
Language 1	raw	na	na
Language 2	raw	na	na
Language 3	raw	na	na
Date of birth	raw	na	na
Current age (months)	calculated	number of months (from the child's date of birth until the questionnaire completion date]	always
Gender	raw	na	na
Number of siblings	raw	na	na
Child birth order	raw	na	na
Country of birth	raw	na	na
Child attended daycare	raw	na	na

NOTES

This sheet contains a list of variables which can be found in the 'raw data and reports' output downloadable from each project (raw and calculated) for all participants in an ordered number of people, events, languages that a child may be irrelevant for a child (e.g., 'language 3' because the output will be empty. The only raw data not include of hours that the child spends with each person, gross a typical week day, during odd week days, during a typical day on holidays. This data can be found in raw data.

In the column 'Raw or calculated measure', we indicate whether the variable is raw or calculated.

In the column 'Calculation explanation or reference', we provide information about what the calculated measure represents or provide information about how the calculation was derived. For clarifications about raw data, see the 'raw data and reports' sheet.

In the column 'When can you rely on this calculation and sub-modules need to be distributed to the participant', we indicate when you can rely on this calculation.

Figure 8. Output interpretation, tidy data and calculations

The sheet *reports* outlines the data summarised in reports aimed at teachers and speech and language therapists (see Figure 9), while the sheet *respondents* clarifies where to find participants' names linked to their IDs (see Figure 10). The sheet *time matrix data* explains what is included in this datafile (see Figure 11).

NOTES

In the section 'Raw data and reports' on each project's page, we provide downloadable summary reports of each child's language profile. These reports are aimed primarily at practitioners (e.g., teachers and speech and language therapists), yet they can be useful to researchers as well. They do not include a summary of all the data. Specifically the reports include the following:

- Child's name
- Basic demographic data about the child's age, languages, country of birth, number of siblings, date the child started daycare/school, languages spoken by their caregivers, age of first exposure to each language
- If relevant (sub-)modules are distributed: cumulative estimates for exposure and use in each language
- If relevant (sub-)modules are distributed: current unweighted estimates for exposure and use in 4 contexts (home, local community, school, during holidays)
- If relevant (sub-)modules are distributed: current weighted estimates for exposure and use in each language
- If relevant (sub-)modules are distributed: richness of linguistic experience in each language
- If relevant (sub-)modules are distributed: proficiency estimates for each language (listening, speaking, reading, writing)
- concerns about the child's language development and periods when the child missed school (if relevant)
- early language milestones (age of first words and age of first word combinations)
- red flag score and a decision tree about a potential referral to a specialist for language assessment
- further information on each of the above measures

Figure 9. Output interpretation, reports sheet

- Device and browser
- Child's age
- Number of languages used by the family and what these languages are
- Which questions elicited "I don't know" responses

Q-BEx is fully [GDPR](#) compliant. Before using the questionnaire, please read the following sections of the project website carefully:

- [Disclaimer](#)
- [Privacy Policy](#)
- [Privacy and Cookies Policy](#)
- [Terms and Conditions](#)