Dear Dr. Maehler,

Dear Reviewers,

We sincerely thank you for the time and effort you have devoted to reviewing our manuscript and for your detailed, constructive, and thoughtful feedback. We are convinced that your suggestions have substantially improved the quality, clarity, and conceptual precision of the paper.

We carefully considered all comments and have revised the manuscript accordingly.

Below, we provide a point-by-point response to each comment. Reviewer comments appear in plain black text, our responses are provided in *italic* and introduced by "Response:".

We have uploaded a marked version of the revised manuscript via Editorial Manager, with all changes highlighted in yellow.

Once again, we are very grateful for the opportunity to improve our manuscript based on your valuable input.

Sincerely,

The Authors

I. Editor

1. Improve the transparency and completeness of documentation regarding measures, translations, questionnaire versions, and the distinction between the YOURGROWTH and YOURCARE components.

   Response:

   We have now added information on the original language versions of the instruments included in the study and on the translations required to administer them in the participants' native languages. We have also included a brief statement indicating that the questionnaires are available from the authors upon request, thereby avoiding potential copyright issues associated with making them publicly available. The distinction between the YOURGROWTH and YOURCARE sub-studies is now outlined in the section on the basic research questions at the beginning of the manuscript to draw readers' attention to this important distinction.

2. Provide a clearer description of the longitudinal design, participant flow, and attrition across waves.

   Response:

   Thank you for this suggestion. We have now added Figure 1 which provides an overview of the longitudinal design of the study (including participant flow, attrition rates, employed questionnaire version, and timely distance between the measurement points.

3. Clarify missing data handling procedures, including the imputation methods, affected variables, and whether imputed values are included in the deposited dataset.

   Response:

   If imputation procedures were used, they are documented in the respective papers. The data file itself does not contain any imputed data. This information was added to the manuscript.

4. Expand the description of quality control procedures, including data entry checks, tablet-based data collection procedures, and anonymization strategies.

   Response:

   We have added more information about quality control procedures for the case of manual and tablet-based data collection procedures. Moreover, we have provided more detail about data anonymization strategies to prevent re-identification of individual participants.

5. Improve documentation for data reuse by providing English versions of the README file and codebook.

   Response:

   English versions of the Readme-file and of the codebook were prepared and additionally uploaded to the data repository at ZPID. Thank you for this suggestion.

6. Strengthen methodological reporting by adding further information on data collection procedures, assessment duration, and reliability estimates.

Response:

We have expanded the sections reporting on data-collection and assessment procedures (s. Section 2.4 and Table 2). Regarding reliability estimates: Where reliability analyses had already been conducted, we reported the published reliability estimates. We did not conduct additional reliability analyses for the published dataset, as reliability should be assessed based on the specific data used in each analysis. For further details, please see our response to point 10 of Reviewer 2 below.

7. Section 3.5 "License and limits to sharing" requires specific attention. The current description of data access needs to be revised to accurately reflect the formal access procedure. Access upon author request is inappropriate, I assume that the dataset is officially deposited with the ZPID Research Data Center, where access is presumably granted through a standardized procedure (e.g., application and completion of a Data Use Agreement). Please verify the actual procedure and clearly describe the repository responsible for access, the steps required for obtaining the data, the conditions of use, and any restrictions. For guidance, please review data papers published in the Special Issue "Data for Psychological Research in the Educational Field" to ensure that the description of controlled data access and licensing follows established standards. If there is uncertainty regarding whether the current access model is compatible with publication as a data paper, a further discussion with the editor before resubmission may be helpful to clarify this point.

Response:

We have expanded the description of the data access procedure. We have also removed the reference to the embargo period ending on May 1, 2026, as this date is no longer relevant. The revised section now reads as follows:

Researchers can order the data set by concluding the associated data use agreement. To access the data use agreement the researcher has to click on the "Access dataset"-button presented on the corresponding dataset-metadata page (see Section 3.1)*, fill in a short form (asking for name, affiliation, and contact information as well as the intended use) and generate the contract. Afterwards researchers have to email the signed contract to RDCservice.info@leibniz-psychology.org. After a short formal review of the contract via the Research Data Center (RDC) team the contributors of the dataset are asked for their permission to provide the data for the intended use. If the RDC team has received the permission, the researcher gets an email including a download link as well as the contract, also signed by the data center, for his/her own records.*

II. Reviewer 1:

1. As a first note, I was not able to access the dataset. I followed the procedure described in the paper and sent the data use agreement to RDCservice.info@zpid.org on June 1, 2025. As of June 5, 2025, I had not received access. This may simply be due to a short delay, but I am therefore submitting this review without having been able to look at the data directly.

Response:

We had an E-mail exchange with the responsible person at the ZPID regarding this issue. We had forwarded the reviewer's comment and asked how long it typically takes to receive a response. The answer was as follows:

In general, depending on the number of requests we receive and the amount of review required, it takes us between one and four days to make the data available. In your case, the email to the reviewer was sent on June 5, 2026, together with the duly signed agreement and the download link. However, the email was inadvertently sent to our service address rather than to the reviewer's email address. In other words, we simply replied to the wrong email address on our end. This means that the workflow was correctly described on your side, and the reviewer did not do anything wrong either.

This means that something obviously went wrong in this particular case, but that requests are typically handled promptly.

2. The manuscript states that the instruments were provided in the participants' native languages and in German, and that translations and back-translations were conducted when needed. It would be useful to specify which questionnaires were already available and validated in the relevant languages, and which ones were newly translated for this study. If possible, the authors could also make the translated items available, for example as supplementary material or repository documentation, provided this is allowed by copyright and ethical restrictions. This would make the dataset more useful for researchers working with multilingual populations.

Response:

We have now added information on the language of the original questionnaire versions to Table 2 (mainly English or German). In some cases, versions were available in both languages (English and German), but versions in the participants' native languages (Arabic, Persian, or Kurmancî/Sorani Kurdish) were rarely available. Thus, forward and back-translations were necessary in almost all cases. You are right that making the translated items publicly available could potentially raise copyright issues. However, we have added a sentence indicating that the instruments used in this study are available from the authors upon request.

3. The manuscript could provide more information on attrition across waves. If available, the authors could report whether attrition differed by study site, age, gender, country of origin, language, questionnaire version, or baseline mental health indicators. This would help future users assess how suitable the dataset is for longitudinal analyses.

Response:

The new manuscript version provides a flowchart with information about the attrition rates across the three assessment waves (including information about the questionnaire version (child versus adolescent version versus short COVID version), assessment wave (wave 1 versus wave 2), and time lag between the assessments. The flowchart provides a clear overview on

the longitudinal design of the study and – in combination with Table 1 – also basic information about age, gender, country of origin, questionnaire language and additional information about the participants.

4. The paper briefly mentions that imputation was used, but does not specify the method or the variables/scales concerned: "Nevertheless, for a variety of reasons – further compounded by the COVID-19 pandemic – missing data occurred, which in some cases had to be addressed through imputation procedures." It would be useful to clarify whether imputed values are included in the shared dataset, which variables were imputed, and what imputation method was used.

   Response:
   If imputation procedures were used, they are documented in the respective papers. The data file itself does not contain any imputed data. This information was added to the manuscript.

5. The manuscript would benefit from a short clarification of the distinction between the YOURGROWTH and YOURCARE components, ideally early in the paper. Some explanation is already given in sections 2.5.6 and 4, but a brief summary of the aims of each subproject, and of which measures mainly belonged to each, would help readers navigate the dataset from the beginning.

   Response:

   We have now added a brief description of the differences between the YOURGROWTH and YOURCARE study components at the beginning of the paper (Section 1.1, Basic Research Questions). However, it is difficult to clearly assign measures to either subproject, as many were relevant to both.

6. The note to Table 1 states that -95 and -99 are used as general missing value codes. It would be helpful to specify directly what each code means, for example in parentheses.

   Response:

   Thank you for pointing this out. We originally used -95 and -99 to code missing values. However, it proved difficult to systematically distinguish between different reasons for missing data. We therefore decided to use -95 consistently throughout the file. It should be noted, however, that five COVID-related items also use the codes 6 (four items) and 11 (one item) to indicate missing data. In these cases, code 6 indicates "I do not know," while code 11 indicates "I do not have any friends" for an item concerning friendships. These codes and their meanings are clearly specified in the description of the item values.

7. Section 2.5.7 on sociodemographic variables appears after Table 2, although these variables are listed at the beginning of the table. It may be more logical and easier to read if this section appeared before the questionnaire sections, or before Table 2.

   Response:

Table 2 was completely reorganized to be in line with the sequence in which the instruments are described in the manuscript text. This should indeed be more logical and easier to follow for readers.

8. The manuscript lists so many publications that already used parts of the dataset. This is useful, but it could be even more informative to include a small table indicating, for each publication, the main questionnaires or constructs used and the wave(s) included. This is not essential, but it would help readers understand both the existing use of the dataset and its remaining potential for reuse.

Response:

We have added a Table which lists the measurement instruments used in our previous publications. Thus, readers can now easily see which instruments were repeatedly used in previous publications and which instruments were seldom or not at all used.

9. I would suggest mentioning directly in Section 2.4 that a trained psychologist was present during data collection. It would also be useful to report the approximate duration of the assessments in this section.

Response:

A short statement on the supervision of the data-collections by a trained psychologist was added to Section 2.4. Moreover, the approximate duration of the assessment (complete assessment battery and shortened COVID-19 version) is also added now.

10. As a minor point, I suggest making the reporting of Cronbach's alpha coefficients more consistent. The current approach, using coefficients from previous publications based on subsamples of the present dataset, is understandable. However, because these coefficients may come from different subsamples, they are not always directly comparable. If feasible, the authors could report a consistent set of reliability estimates for the deposited dataset, for example using the first available assessment of Wave 1. I understand that this may require additional analyses, so I leave it to the authors to decide whether they want to include this.

Response:

Where reliability analyses had already been conducted, we reported the published reliability estimates. However, we decided against conducting additional reliability analyses for the published dataset, as providing further reliability estimates might create confusion rather than add clarity. In future reanalyses of the data, reliability estimates should be calculated and reported based on the specific data used in each analysis. In addition, reliability analyses have not yet been conducted for all questionnaires, as some of these measures have not been used in any of the manuscripts published to date.

III. Reviewer 2:

1. Introduction & Background: The first couple of sentences mention more current immigration trends in Germany—how relevant are those statistics when the sample is related to earlier displacements? It might be useful to provide more context on the population and timeframe

of the data instead.

Response: *Thank you for this helpful comment. We have revised the Introduction and Background section accordingly by removing the more current immigration statistics and instead providing historical context that reflects the period relevant to our sample. Specifically, we now include information on the scale of refugee arrivals in Germany, the age distribution of asylum applicants, and the main countries of origin during 2015 and 2016, which correspond to the period in which most participants in our sample arrived in Germany.*

2. Methods: The study design section might benefit from a participant flow chart. It's a bit confusing who started when, and how much attrition happened for which group at which time. A CONSORT diagram might not quite work due to the multiple waves, but a general timeline indicating number of participants from each wave and reasons for attrition might work? This will help with understanding data availability for longitudinal studies (one of the possible reuse designs you mention later in the paper).

   Response:

   A flowchart was added to the manuscript (new Figure 1). It illustrates the longitudinal design, attrition rates, and the time intervals between assessments, while also distinguishing between Waves 1 and 2.

3. For Table 2: Is it possible to indicate if a COVID-SF was used differently? At the moment, it looks like if COVID-SF has a checkmark, that this form was used across all 3 waves? Is there a different way to indicate more precisely which versions of the test were used for whom at which timepoint?

   Response:

   We have added more information about the use of the COVID-19 short version. This version was used for a subsample of participants, mainly in 2020. The total number of participants who received this version, as well as the respective sample sizes at the initial and second assessments, are now specified both in the manuscript text and in the new Figure 1.

4. Quality control: I'm confused about the imputation process. Does this mean that some of the data in the scientifically available data has already been imputed? If yes, I think this needs more justification. How would a user know which values were imputed?

   Response:

   Imputation procedures were used in some published manuscripts and are documented there. The published data files do not contain imputed data. This information is explicitly added to the manuscript now.

5. For manual entry, was an error entry rate calculated? How were discrepancies resolved?

   Response:

For manual entries, random checks were conducted on an intended proportion of 10% of the entries. Moreover, out-of-range entries were identified by inspecting item frequency tables for incorrect values. Any discrepancies were resolved by checking the participants' original responses. This information is now provided for manual data entries.

6. For the tablet data collection, was there quality control on setting up data collection? For example, range restrictions for numerical items with finite options. While it may seem data is collected "correctly", this may not be the case depending on how collection systems were set up.

   Response:

   The range of possible entries was restricted to valid responses. This information is added to the new version of the manuscript. For both tablet and manual entries, additionally plausibility checks were conducted (e.g., by comparing participants' reported ages across assessment time points). This was also added to the manuscript now.

7. Data anonymization: This section is more about data privacy and deidentification from the front end. Can you specify how data were checked to prevent re-identification? I understand no other de-identification methods were used, such as crosstabulation or truncation?

   Response:

   Thank you for raising this important point. Before the data were made available, measures were taken to minimize the risk of re-identification of individual participants. Potentially identifying information was either removed or recoded into broader categories. For example, exact age information was replaced by age categories, and information on the type of school attended was reduced to a distinction between primary and secondary school. In addition, responses to open-ended questions were removed because they frequently contained personal information. This information was added to the manuscript in Section 2.7.

8. Reuse Potential: There's great potential for reuse in the scientific community and many great ideas for new papers are presented. I do think reuse is limited by making the data available to the scientific community only through a DUA. Does this limit to research papers only, or would you consider data use for teaching purposes, for example? I'm also wondering if there is potential for data aggregation, and if someone had identified possible datasets with similar variables? What is the rationale behind this decision?

   Response:

   As these are sensitive data, it is necessary to prevent any use beyond scientific contexts. However, the DUA explicitly permits the use of the data for teaching purposes. While it states that "forwarding the materials made available to third party is prohibited," the following sentence specifies that "an exception shall apply to use of the materials by researcher groups/project groups and in academic training." This makes clear that the use of the data for teaching purposes, as well as within broader scientific contexts, is explicitly permitted. The process of getting access to the DUA is now explicitly described in section 3.5.

9. Regarding the datasets themselves, the process of obtaining data was easy, even though the DUA might deter reuse. The datasets and accompanying codebook and readme file are .txt files, which is appropriate. The ZIPD site has a good short description of the project and potential constructs, as well as a public facing codebook. The dataset has a DOI attached, making citations easy to track—I may have missed it, but I did not find the license for the dataset. This would be good information to make clear to user.

   Response:

   As described above, we have now provided information about the data access procedure in section 3.5. This is a general procedure used for all data set accesses via the Research Data Center at the ZPID.

10. My only real concern is that the readme file and codebook are only in German. I was able to understand enough to get the gist, but I think this may limit the potential of reuse of the data. Providing a readme and codebook in English would be helpful to broaden the audience.

    Response:

    The readme file as well as the codebook are provided additionally in English language at the ZPID data repository now. Both files can be accessed there.