General Definitions

This document

Here we collect definitions regarding ManyBabies projects that are aimed to facilitate collaboration for all parties involved. This is VERSION 0.2 - mention this version number whenever referring to definitions. To suggest updates and changes, you can (anonymously) comment here:
https://docs.google.com/document/d/1UYSevbWnBQwd_eaBe1oKkOBX-8sMsBfiPz2kwNp7T1c/edit#

Last edited: 2020-03-10 by Christina Bergmann
Written by: The ManyBabies Governing Board
Reviewed by: Project Leads: Kelsey Lucca, Rebecca Lundwall, Tobias Schuwerk, Angeline Tsui, Ingmar Visser

What is ManyBabies?

- A 6-continent, 30+-nation, grassroots open science network for large-scale infancy research
- A “way station” between basic research and application of study results
- A healthy vision for the future of developmental science, providing a practical and coordinated approach to replication

The goals of ManyBabies include:

- Validity: Assess the replicability, generalizability, and robustness of key findings in infancy (effect + effect size + moderators) in consensus paradigms. Note that these paradigms are not necessarily exact replications of seminal studies.
- Theory: Understand infant development across a wide range of ages and experiences (e.g., language and cultural background)
- Methods: Quantify sources of variability across labs and establish best practices
- Diversity: Increase diversity of participants, researchers, & research questions

Types of ManyBabies Projects

Main projects: Main ManyBabies projects are numbered sequentially (e.g., ManyBabies 1, ManyBabies 2). Each project tests a question about infant development of key theoretical or methodological interest.

Spin-off projects: Spin-off projects are ManyBabies projects that build on the methods or data from the main ManyBabies projects, and are pre-registered before all the data necessary to the spin-off project has been collected. Some examples include extending the study to a different population of infants (e.g., bilinguals; additional geographically/culturally diverse samples), follow-up data collection (e.g., longitudinal correlation with CDIs; test-retest reliability), and additional analyses of the main dataset beyond the main project that were pre-registered prior to the completion data collection (e.g., methods moderators).

Secondary analyses: Secondary analyses are conducted on publicly-available ManyBabies data, from the main project and/or spin-off projects. In some cases, these analysis may include additional data not collected as part of ManyBabies, for example if an individual lab investigates relationships between their lab’s ManyBabies data and other data collected by their lab.
Key Principles of ManyBabies Projects

ManyBabies projects are collaborative experimentally focused projects in developmental research. They are built on the following key principles. The responsibilities for different parties in a ManyBabies project defined below are derived from these principles.

1. Transparency: All decisions in a ManyBabies project, from design to publication, are as transparent as possible.
   a. If possible, ManyBabies projects are published as registered reports. When this is not possible, analysis plans are preregistered and made public as soon as practicable, at the latest upon project completion.
   b. Preprints are strongly encouraged.
   c. Project websites should serve as the central access point to all relevant information and documentation about a specific ManyBabies project.
   d. Openness: ManyBabies projects incorporate openness in four ways
      i. Open to participation: Constructive participation is welcome. Everybody who wants to contribute should be included barring substantial reasons for exclusion (e.g. breaching the code of conduct).
      ii. Open materials: Stimuli, questionnaires, and protocols are freely shared with the public (at the latest upon publication, but ideally throughout the project).
      iii. Open code (using open source tools when possible): All analysis code is made available so that others can independently verify analysis decisions and results.
      iv. Open (anonymized) data: To maximize transparency, support the validation of all analyses, and enable exploratory analyses that could generate new hypotheses, all data should be shared openly if possible (see 4. Ethical research for limitations). This requires advance planning on the part of both project leaders and individual contributors to ensure that appropriate permissions and procedures are in place to allow for open data-sharing.
      v. Open publication: When possible, it is strongly encouraged to share preprints and/or publish open access.

2. Collective governance: When possible and feasible, decisions are made jointly and transparently
   a. Each project will have a leadership team whose mandate is to facilitate decision-making and ensure timely progress of the project. Wherever possible, consensus-based decision-making is encouraged, defaulting to democratic (vote-based) decisions where a reasonable consensus cannot be reached.
   b. In rare situations the leadership team will make executive decisions (e.g. when consensus is not possible or in the presence of external time pressure).

3. Inclusivity and diversity: ManyBabies explicitly encourages efforts to increase diversity. This includes active applications for funding to assist laboratories from under-represented communities to participate, outreach efforts, and planning workshops.
4. Ethical research: ManyBabies projects are committed to best practices and high ethical standards, they follow the Declaration of Helsinki (version of 64th WMA General Assembly, Fortaleza, Brazil, October 2013) without exception.

   a. Data will only be shared with informed consent or when fully anonymized.
   b. Participation by research participants must be both voluntary and informed.
   c. Participating laboratories are expected to follow the ethical and legal standards of their country of operation and granting agencies.

5. Respect: Everyone agrees to a general code of conduct which ensures respectful interactions. We have no reason to believe this necessary, but prefer to err on the side of caution.
Roles and Responsibilities in ManyBabies

The Governing Board

The governing board ensures that ManyBabies projects adhere to the above-mentioned principles, coordinates different projects, support project leads and contributors, enforces the code of conduct, and serves as central contact point.

**Responsibilities:** The governing board responds to concerns regarding compliance to policies laid out in this document and additional policies set by project leads (assuming they are not conflicting, in which case this document takes precedence) brought forward by anyone. Concerns and complaints are treated confidentially and can be made to any member of the governing board or to the whole board.

To support specific ManyBabies projects, the governing board will regularly check in unless a governing board member is already part of the leadership team. If desired, ManyBabies project leads can nominate specific governing board members as liaison.

The governing board also supports grant proposals related to ManyBabies projects. Interested parties can contact the governing board to request materials from previously submitted grants, letters of support, and advice.

Project Leads

Project leads either propose a study to the governing board or respond to a proposal by the community or the governing board.

Project leads can be a single person or a group. Based on experience, it is useful to appoint leads for sub-tasks (e.g. analysis, writing, stimuli, coordination and communication). There are no restrictions regarding the current type of funding / employment for project leads.

**Role:** Project leads are free to determine specifics of a given ManyBabies study, including the precise topic, design, research question, and venue for publication(s). The project leads can independently of the governing board pursue funding options for different aspects of the study, and can ask the governing board for support.

Spin-off projects and secondary analyses to a specific study can have different project leads (see below for details). For a spin-off project that is taking place in parallel or closely after a main study, it is recommended to coordinate with the project lead from the main study. Secondary analyses independent of ManyBabies are encouraged as soon as the data have been made public. ManyBabies project leads are expected to provide a reasonable level of support to such endeavours to the extent expected of standard Open Science practices outside of the ManyBabies context. Details on the role of spin-off project leads and secondary data analysts are provided below.

**Responsibilities:** Project leads are responsible for all aspects of a specific study and for adhering to the ManyBabies principles. This includes transparency, openness to participation, open materials, code, and data. For concrete decisions, project leads should facilitate collective decision making by the group whenever possible or make carefully considered executive decisions when needed (e.g. when consensus cannot be reached or under time pressure). Project leads also ensure that the information provided to data contributors is clear and permits data collection in the desired way, provide contact points for contributors, and ensure that authorship of any resulting manuscript is attributed correctly to all (including data and project contributors). Concretely, support for contributors entails that project leads provide a format for
data contributors to pre-register, support local set up of the study, and templates for data submission.
To facilitate fulfilling those responsibilities, it is advisable to create a project page early on and provide contact information for the project leadership team; if applicable with specific assigned topics (e.g. design, analysis, general questions).

Data Contributors

ManyBabies projects rely on laboratories to collect data. Each laboratory is typically represented by one or two researchers, the level of seniority is not limited and ManyBabies explicitly encourages the participation of junior researchers from the undergraduate level onwards.

**Role:** Contributors agree to collect data following the experimental protocol as provided by the project lead and will be listed as authors of a final manuscript.

**Responsibilities:** Contributors pre-register their contribution and thus pre-determine all degrees of freedom (most saliently planned sample size) within a ManyBabies project in the format provided by the project lead (e.g. via filling in a questionnaire) before starting data collection. All changes to the preregistration will be noted and communicated to the project lead (e.g. changes in sample size, method, unforeseen obstacles). Contributors agree to follow study protocols as closely as possible, and to report any protocol deviations. Data peeking (i.e. analyzing the results and adjusting sample size) is prohibited. Contributors agree to follow the ethical and legal standards of their institution, country of operation, and granting agencies, and obtain ethical approval accordingly. Laboratories operating in countries and/or institutions without clear processes for ethics training and/or research ethics approval processes ensure that they follow the declaration of Helsinki and adhere to local laws.

Project Contributors (design, materials, analysis, etc)

In addition to collecting data, contributions to a careful design, stimuli, presentation software, analysis code, and writing are welcome and necessary at ManyBabies projects. Contributors are listed as authors and it is the project leads’ responsibility to track contribution and ensure that all qualifying contributions lead to authorship.

**Role:** Depending on the type of contribution: Create stimuli and document the process carefully, specify protocols, contribute analysis code (ideally via a version control system, such as git), etc.

**Responsibilities:** All contributors are kindly asked to carefully document their contribution (e.g. comments in code, protocols for stimulus creating) and follow the general ManyBabies principles outlined in the beginning.

Spin-off Project Leads

ManyBabies projects can have parallel or follow-up studies (see ManyBabies 1B on bilingual infant directed speech preference / gaze following for an example of a parallel spin-off and ManyBabies 1 CDI data collection for a follow-up). These can be proposed by an independent leadership team.

**Role:** Similar to a Project Lead. We recommend coordination with the main project leadership team to facilitate protocol and data exchange, ensure that all necessary information is available, and to avoid any friction between related projects.
**Responsibilities:** In addition to the responsibilities of a Project Lead, Spin-off Project Leads additionally ensure that their project does not create conflicts. A spin-off project requires a separate preregistration (if appropriate), report, and (usually) design and analysis team. Project leads should communicate clearly what constitutes authorship-level contributions and how authorship order will be determined.

**Secondary Data Analysts and Analyses**

All ManyBabies data are shared with clear license attribution at a time point when the main project lead deems it appropriate (although see Project Leads for restrictions). Secondary analysts are required to adhere to this license when using ManyBabies data, and not share any results from secondary or subset data analyses of yet unreleased datasets (accessed for example in the context of local data collection or through analysis efforts).

For ManyBabies 1, the data from the main project were shared with an MIT license at https://github.com/manybabies/mb1-analysis-public. This license is very permissive and permits free re-use: https://github.com/manybabies/mb1-analysis-public/blob/master/LICENSE

It is recommended to use permissive license when possible to facilitate data re-use and avoid possibly conflicting licenses across sharing platforms.

As courtesy and to be able to track data re-use, we ask secondary analysts to cite the main manuscript. Providing the correct citation along with the data is the responsibility of the leadership team.

A few different types of secondary data analysis can be identified which may have different expectations.

**Internal secondary analyses:** The first type is "secondary" analyses that are conceived as part of the original planning for the project and conducted by members of the original team. For example, in ManyBabies 1, we had planned to examine "lab factors" in a separate analysis that was not part of the primary registered report. Because this analysis project emerged as a planned "spin-off" to the ManyBabies 1 primary study, it is expected that (as with any spin-off project) practices and management will conform to all ManyBabies guidelines and practices.

**External, independent secondary analyses:** On the other end of the spectrum are analyses conducted entirely by third parties without any direct affiliation with ManyBabies, using ManyBabies data that have been made publicly available. In such cases we strongly request that such analyses conform to the standards of Open and Collaborative Science and follow the principles articulated in this document. However, in the interest of Open Science we do not place formal restrictions on the data use other than those articulated in the original license on the data. As a courtesy we ask that the main ManyBabies report related to the dataset will be cited.

**Intermediate secondary analysis models:** In between these two types are a broad spectrum of other possibilities for secondary use of data, with groups of varying degrees of affiliation with ManyBabies. In general, where it is perceived that the group conducting the analysis is a representative of ManyBabies, they should conform to all ManyBabies Standards.

**Secondary analyses involving ManyBabies Governing Board members and Project Leads** are expected to conform to the standard of this document whenever they conduct analyses involving ManyBabies data, regardless of the status of the data within the public domain, and to ensure fairness in access to authorship on such analyses. In particular, Governing Board members and Project Leads will ensure that any secondary analysis projects they lead or participate in on a ManyBabies dataset that they are involved with (at any stage) will:
a) Be made known to all data contributors and project contributors (e.g. via an email announcement to the group)

b) Provide a specific mechanism for data contributors and project contributors to elect to become an active author-level participant in the project (e.g. a sign-up form with clear identification of contribution types and opportunities). Participation in the secondary analyses entails a productive contribution and is similar to the role of a project contributor defined above.

_Data Contributors and Project Contributors_ may also want to conduct secondary analyses that were not part of the original planning for the project. In such cases, _if the analysis is started prior to the dataset becoming public_, contributors are expected to follow the same rules as the Governing Board and Project Leads in making the planned analysis known and inviting participation (a and b above). This is in part to avoid duplicate efforts and to ensure that contributors who have made a substantial contribution to making the analysis possible (e.g. in data processing) are given the opportunity to join the team and be appropriately recognized for their efforts in an authorship role. Contributors are also expected to conform to ManyBabies Open Science standards in their management of the secondary analysis. _Once the data have been made public_, contributors are still strongly advised to conform to the ManyBabies standards of collaborative and open science, and encouraged to seek out collaborators within the ManyBabies framework, but contributors may at this point choose to conduct analyses on the ManyBabies dataset outside of the umbrella of ManyBabies and would therefore not be held to the standards of this document.
Timeline of a typical ManyBabies project

1. Converge on topic and leadership (process can vary)
   a. Recommendation: Converge on leadership team early, distribute tasks
   b. Optional: Call for contributions to the following steps
2. Define research question (which aspects of the topic should be studied in a ManyBabies setting?)
   a. Which aspects can/should be controlled across labs?
   b. What can vary?
   c. Are the above feasible?
3. Design study
   a. Recommendations:
      i. Create rich documentation for contributing labs
      ii. Document stimulus creation procedure
      iii. If possible, share stimuli at this point
4. Create analysis protocol
   a. Recommendations:
      i. Use an open source language
      ii. Simulate expected data (update with pilot data where appropriate)
      iii. Share at this point with the group, if possible
5. Pilot
   a. Goal: Test procedure / protocol
   b. Recommendations:
      i. Do not use pilot data to estimate effect sizes
6. Preregister (or submit a registered report, see cos.io/rr for journals accepting registered reports)
   a. How many participants should a contributing lab test?
      i. Recommendations:
         1. Consult a recent meta-analysis if available (see e.g. metalab.stanford.edu)
         2. Maximize for lab number if systematic between-lab variability is expected
7. In parallel: Open call for participation
   a. Send to mailing lists
   b. Recommendation:
      i. Send out about 3 months before data collection starts
8. Data collection
   a. Recommendations:
      i. Data Contributors will have many questions, so channels of communication and documenting the answers are useful (See ManyBabies 1 using Slack)
      ii. Ask for careful documentation of lab variation, ideally via a lab walkthrough video (including [mock] briefing, procedure, debriefing)
      iii. Ask contributors to sign an adapted cooperation agreement to ensure that the responsibilities mentioned in this document and specific to the project are clear to both the leadership team and the contributors
   b. Pre-specify time frame
9. Data analysis
   a. Recommendation: Ensure compliance to a template through validation
10. Write up
11. Share data (in line with legal and ethical standards)
ManyBabies Code of Conduct

Our Pledge
In the interest of fostering an open and welcoming environment, we as contributors and maintainers pledge to making participation in our project and our community a harassment-free experience for everyone, regardless of age, body size, disability, ethnicity, sex characteristics, gender identity and expression, level of experience, education, socio-economic status, nationality, political perspective, status within the community, personal appearance, race, religion, or sexual identity and orientation.

Our Standards
Examples of behavior that contributes to creating a positive environment include:

- Using welcoming and inclusive language
- Being respectful of differing viewpoints and experiences
- Gracefully accepting constructive criticism
- Focusing on what is best for the community
- Showing empathy towards other community members

Examples of unacceptable behavior by contributors include:

- The use of sexualized language or imagery and unwelcome sexual attention or advances
- Trolling, insulting/derogatory comments, and personal or political attacks
- Public or private harassment
- Publishing others’ private information, such as a physical or electronic address, without explicit permission
- Other conduct which could reasonably be considered inappropriate in a professional setting

Our Responsibilities

The governing board and project leads are responsible for clarifying the standards of acceptable behavior and are expected to take appropriate and fair corrective action in response to any instances of unacceptable behavior.

The governing board and project leads have the right and responsibility to remove, edit, or reject comments, commits, code, wiki edits, issues, emails, and other contributions that are not aligned to this Code of Conduct, or to ban temporarily or permanently any contributor for other behaviors that they deem inappropriate, threatening, offensive, or harmful.

Scope

This Code of Conduct applies both within project spaces (OSF, github, mailing lists) and in public spaces when an individual is representing the project or its community. Examples of representing a project or community include using an official project e-mail address, posting via an official social media account, or acting as an appointed representative at an online or offline event (e.g. during a conference presentation or at a meetup). Representation of a project may be further defined and clarified by the governing board.
Enforcement

Instances of abusive, harassing, or otherwise unacceptable behavior may be reported by contacting the ManyBabies governing board (manybabies-gb@mailman.stanford.edu) or any of its members (see manybabies.stanford.edu). All complaints will be reviewed and investigated and will result in a response that is deemed necessary and appropriate to the circumstances, ranging from discussion of these policies to expulsion from the group, including all mailing lists (either for a set time span or permanently). It is at the complainant’s discretion to pursue additional action through other legal or institutional mechanisms. The governing board is obligated to maintain confidentiality with regard to the reporter of an incident. In cases of complaints directly to project leads or governing board members, the reporter's identity will be known only to those directly contacted by the complainant, unless the complainant’s explicit permission is obtained.

Governing board members or project leads who do not follow or enforce the Code of Conduct in good faith may face temporary or permanent repercussions as determined by other members of the governing board or project's leadership.

Further details of specific enforcement policies may be posted separately.

Attribution