Standard Operating Procedure for Training New and Existing Team Members for Narrative Coding

Author: Dr. Jordan A. Booker

Version: 0.8

Date last revised: June 20, 2025

Change Log

- Version 0.8
  - Additional details on scheduling teams for training and reliability-building, and expected timelines for training
- Version 0.1
  - Initial Operating Plan

Purpose of Narrative Coding

Narratives are life stories, and in human research, narratives can be a source of data that are collected during either structured or open-ended tasks. Narratives can be autobiographical (e.g., stories about previous high points in life), vicarious (e.g., stories about parents' upbringing), or even imaginary (e.g., children making up their own fairy tales). Narratives can be used as sources of different kinds of data: deeper *qualitative* themes that emerge from the people's experiences; certain kinds of *content and linguistic* themes that are represented with certain coding or rating approaches; and forms of *quantified, numbers-based* ratings of different narrative dimensions. Our lab most often uses quantified approaches, grounded in *narrative-* and *personality-based* sciences that are common in areas of psychology. Specifically, our team often rates narratives, and often autobiographical or vicarious narratives (stories that count as **life stories**) for expressions of what is called **narrative identity**.

Narrative identity involves the ways people organize a complex and well-integrated life story; situate emotion and motivation into their life story; and draw meaning and reasoning from their life stories. We train our lab members to take what start as open-ended narrative responses from research study participants–whether these responses are spoken or written–and use established coding systems to form new scores about certain dimensions of narrative identity; the expressions of structure, emotion and motivation, and life reasoning that emerge in people's narratives. Other examples of our team's work in this research space are below:

1. We consider the ways people incorporate structure, like *elaborative details*, in their autobiographical narratives

   > Schroeder, E., Ell, M. A., & Booker, J. A. (2024). Storytelling elaboration and hope during COVID-19 shutdowns: Ties with college adults' psychosocial adjustment. *Journal of Happiness Studies, 25*(23). https://doi.org/10.1007/s10902-024-00734-x

2. We consider the ways people express *agency and communion* as motivational themes in people's life stories

   > Booker, J. A., Manson, J. I., Zanger, L., Froese, S., & Kulesa, B. (2025). Adolescents' agency and communion in life storytelling with mothers: Addressing descriptive and inferential questions. *Journal of Adolescence*. https://doi.org/10.1002/jad.12499

3. We consider the ways people express reasoning through *personal growth* in their autobiographical narratives

> Booker, J. A., Brakke, K., & Pierre, N. (2022). It's time to make more goals so I can keep pushing: Hope, growth, and well-being among young Black women. *Emerging Adulthood*, 10(4), 876–890. https://doi.org/10.1177/21676968221089179

Our research team uses scores of expressed narrative dimensions to answer additional research questions. Do expressions of narrative dimensions, like agency or growth, tend to differ by development, or given gender identification, or across different points of time (e.g., before COVID impacts and during COVID-related shutdowns)? Or, are differences in how people express narrative dimensions related to other reports of psychological well-being or identity development or mental health? Because these narrative scores are often key measures for addressing important research questions and sharing research findings professionally, **it is critical that we have a dependable and consistent approach for training teams to gather new data**.

Purpose of Team Training and Reliability-Building

Unlike some kinds of data, like questionnaires and surveys, there are additional, typically *human-dependent*, steps that are needed to successfully derive scores about narrative dimensions; scores that will serve as a new source of data for our projects. **This is one of the ways our research team gathers and analyzes data**.

Thus, we have a set of steps where we want to *consistently* and *dependably* train people to identify narrative dimensions and characteristics, and to do so in a similar fashion across team members. This contributes to the idea of **inter-rater reliability** and consistency between multiple people who are reviewing the same content for scoring.

> It is important to note that *reliability* is not the same thing as *validity*, and whether we are capturing what we think we are with our coding approaches. Though, we do have a theoretical basis and intentionally designed coding manuals we train people with.

Purpose of this Standard Operating Procedure

This document has been created to create guidance and a consistent set of training activities and standards for the conduct of narrative coding as a means of data analysis in the Milestones Lab at the University of Missouri. **Lab members should ensure they are consistently following these steps in every new attempt for team-based narrative coding**. New narrative coding emerges under the following circumstances:

1. There are **new coding members** who are being trained to code narratives for a dimension of narrative expression or narrative identity
2. Existing coding team members are beginning coding after an extended break (we are operating in a **new phase of time**), such as following Thanksgiving or Spring Break, or a return from a change in semesters.
3. Existing coding team members are beginning coding on a **new narrative dimension(s)**, such as moving from coding agency to communion, or moving from coding contextual coherence to thematic coherence. Even with similar narrative dimensions, training should be conducted to ensure high inter-rater agreement and reliability of coding.
4. Existing coding team members are coding on narrative dimensions they are already trained and reliable on, but are working with a **new data source**. Here, a new data source may come from:
    - a *different narrative prompts* from the same sample (e.g., coding high point events after coding low point events from the same individuals);
    - narratives captured at *different points* in time from the same sample (e.g., moving from narratives of turning point events collected at baseline to turning point events collected at one-point follow-up);
    - narratives captured from a *different samples* responding to the same narrative prompts (e.g., moving from self-defining memories collected from students at the University of Missouri to self-defining memories collected from students at the University of Houston)

- narratives captured from *different projects* (e.g., moving from stories about parents' upbringing, collected from a Fall 2021 sample to stories about parents' upbringing, collected from a Spring 2022 sample)

Aside from well-experienced lab members (e.g., Dr. Booker, graduate students, undergraduate students with multiple years of experience), coding team members should only be involved with training and coding of **one narrative dimension at a time**.

Steps of Narrative Coding Training, Reliability-Building, and Individual Coding given Reliability

Narrative coding is an activity that occurs after we have collected research data from a human sample. Narratives might be in a spoken (audio or video recorded) format. Narratives can come from one person sharing stories after receiving a story prompt, from two or more people sharing stories together (e.g., parent-adolescent storytelling), or from people responding to interview prompts with a researcher.

The kinds of narrative prompts we use for data collection and the kinds of narrative dimensions our teams will code to form new data are often determined during the project design and planning phases. Hence, we are usually selecting narrative dimensions to train people on given this earlier planning and given interest in addressing specific research questions (e.g., is agency related to measures of depressive problems concurrently and at a four-month follow-up?) and hypotheses (e.g., agency will be negatively related to depressive problems at baseline; agency will predict a decrease in depressive problems at four-month follow-up)?

Once we have *narratives* and *narrative dimensions* of interest, we begin training people to code these dimensions. We usually organize teams of 2-to-4 people to contribute to a coding team. This team size balances feasibility for team logistics (e.g., finding a common meeting time) and statistical progress (e.g., building evidence of consensus across the entire team). Part of why we need to initially begin with multiple people coding our narratives is because it helps reduce concerns of *error* and *bias* in our data findings–if there are multiple people, trained on the same information and criteria, who are interpreting findings in similar ways, we have stronger evidence that we are capturing something in the intended manner. This is not the same thing as guaranteeing that what we capture is valid and truly represents the idea or construct of interest, but that is a separate theoretical and methodological concern. This approach accounts for what we can best control at this phase of the research process.

The overall steps for our narrative teams include the following:

1. **Establishing a meeting schedule for the team**. Teams should aim to meet at least twice a week when possible. Meetings should be at least two days apart, to give adequate time for reviewing materials and contributing effort to coding work.
    - If a team cannot identify multiple meeting times per week, especially with less experienced coding team members, a different team might need to be organized with better scheduling fit.
2. **Reviewing a coding manual** together for addressing our narrative dimension of interest.
3. Engaging in **practice** with this coding system, as a team, using sample narratives.
    - This is an important time to consider whether there are gaps or needed changes for the coding system. Sometimes, with a new sample (e.g., a different age group than usual) or different type of narrative (e.g., vicarious rather than autobiographical narratives), slight adjustments are needed for consistent and high-quality coding. This is acceptable, so long as we work to establish needed changes early and then stick to them.
4. Once we have an agreed approach for the coding system, we will begin **building consensus** with each other. This is a *team-based approach* to developing *inter-rater reliability*.
    - Each team member will receive a set of the *same* narratives to review and give narrative dimension scores *on their own*. Team members should continue to reference the coding manual during this process and later coding efforts. *You should not exclusively rely on memory*.
    - The team will meet and discuss each person's scores.

        - If everyone already has the same scores for a narrative–great, no further discussion is needed unless there are specific questions people would like to discuss.
        - If there are any disagreements between team members' scores, the team will take the time to discuss the narrative together and will move toward a final group score. This is called a **consensus score**. That consensus score is the final data used in later analyses.
        - If the team is having difficulty reaching consensus for a specific narrative, a designated lead coder–typically the person who led training–has the authority to set a consensus score for the team. That person should provide their reasoning to other team members so similar decisions can be reached by the team with ongoing coding.
        - Each team member's original scores should always be saved and should not be changed following consensus. *These original scores are valuable sets of data in their own right*. They are the information used to test evidence of inter-rater reliability, or how much similarity team members showed *before* they discussed their scores.
    - After consensus meetings, the lead coder should organize all original scores and should conduct a statistical analysis to determine inter-rater reliability. The score tested for reliability will differ given the nature of codes (e.g., ordinal, categorical, etc.).
5. After inter-rater reliability has been established, **individual coding assignments** will be given to each team member. This means that each person is coding a separate set of narratives on their own and that their scores will be treated as *final*.
    - Individual coding can occur at the timing of each team member's choosing, but should stick to a reasonable and agreed upon timeline with the team leader (e.g., "coding assignments should be completed within the next week")
    - Coding should be done from a **private** and **secure** location.
        - This means that others should not be able to see or interact with your work. Coding team members should be mindful if completing work at home where others live and should *avoid* doing coding work in public spaces (e.g., coffee shops). Again, we are working with private information from research participants and want to take all reasonable steps to avoid breaching participants' confidentiality and anonymity in participating in our projects.
        - Because we use password-protected systems like Microsoft Teams for managing files and data entry, work that you complete in Teams should already be secure. You should not need to fully download narrative files or data entry spreadsheets for work. There are steps for opening files with desktop apps that are only saved back to online servers rather than locally on a computer. Please use these steps rather than downloading files fully.
            - If you do download any files for temporary use, please make sure to completely delete and remove these files from your personal computer once done with coding and coding entry. Do **not** ever download our data files (e.g., narratives, data entry spreadsheets) onto public computers (e.g., library computers). We do have dedicated lab computers in McReynolds Hall that are private and secure and (e.g., only used for research purposes; in a locked room when not in use). You are always welcome to use these computers as well for lab work.
    - There will be dedicated data entry spreadsheets for your work. Please only place your scores in these dedicated files, so that we have a dependable space to work with them. Again, this is a source of data for other professional data sharing efforts.
    - **If you believe there is an error in your codes that you did not catch initially, please let your coding team leader and/or Dr. Booker know, as soon as possible**. Life happens and we will not be upset that something needs to be corrected, but we depend on these scores for professional work and we need to know if there is an issue, prefably before we publicly share our findings and work.
6. Even after inter-rater reliability is established, we often complete ongoing consensus meetings to **minimize possible coder drift** or ways that people may begin scoring differently from each other if enough time passes without team discussion.
    - Consensus meetings often continue *alongside* other individual assignments, so that we can maintain productivity while ensuring data is robust.
7. **Both consensus and individual codings will continue** until such time as all relevant data for the project is coded, or such a time that a change in semester / team availability would necessitate the end of coding

team work.

- Again, if there is a major change in team or timing, please begin with fresh training and consensus work before moving into new individual assignments.

Note: There are some projects or coding systems where building inter-rater reliability and support for individual coding is not feasible. For example, some instance-based coding systems are very complex, and they require constant conversation among all team members. Rather than moving toward individual work, teams should plan for an entire schedule of consensus meetings and resolving differences through discussion with those projects. This is a rarer case for our lab and would likely be discussed with Dr. Booker in advance of coding in this way.

Additional Readings

Guidance for this standard operating procedure is based on expertise and experiences from our team members, as well as widely accepted guidance from other experts in narrative science. Additional readings from experts in narrative science speak to these considerations as well:

Adler, J. M., Dunlop, W. L., Fivush, R., Lilgendahl, J. P., Lodi-Smith, J., McAdams, D. P., McLean, K. C., Pasupathi, M., & Syed, M. (2017). Research methods for studying narrative identity: A primer. *Social Psychological and Personality Science, 8*(5), 519–527. https://doi.org/10.1177/1948550617698202 Syed, M., & Nelson, S. C. (2015). Guidelines for establishing reliability when coding narrative data. *Emerging Adulthood, 3*(6), 375–387. https://doi.org/10.1177/2167696815587648

Additional research on the theories and broad evidence of the nature and importance of narrative identity are below:

Booker, J. A., Fivush, R., & Graci, M. E. (2022). Narrative identity informs psychological adjustment: Considering three themes captured across five time points and two event valences. *Journal of Personality, 90*(3), 324–342. https://doi.org/10.1111/jopy.12668

Fivush, R., Booker, J. A., & Graci, M. E. (2017). Ongoing narrative meaning-making within events and across the life span. *Imagination, Cognition and Personality, 37*(2), 127–152. https://doi.org/10.1177/0276236617733824

McAdams, D. P., & McLean, K. C. (2013). Narrative identity. *Current Directions in Psychological Science, 22*, 233–238. https://doi.org/10/f42tmw

McLean, K. C., Syed, M., Pasupathi, M., Adler, J. M., Dunlop, W. L., Drustrup, D., Fivush, R., Graci, M. E., Lilgendahl, J. P., Lodi-Smith, J., McAdams, D. P., & McCoy, T. P. (2020). The empirical structure of narrative identity: The initial Big Three. *Journal of Personality and Social Psychology, 119*(4), 920–944. https://doi.org/10.1037/pspp0000247