Semi-automated transcription and scoring of autobiographical memory narratives

Victoria Wardell[1] · Christian L. Esposito[1] · Christopher R. Madan[2] · Daniela J. Palombo[1]

Abstract

Autobiographical memory studies conducted with narrative methods are onerous, requiring significant resources in time and labor. We have created a semi-automated process that allows autobiographical transcribing and scoring methods to be streamlined. Our paper focuses on the Autobiographical Interview (AI; Levine, Svoboda, Hay, Winocur, & Moscovitch, *Psychology and Aging, 17*, 677–89, 2002), but this method can be adapted for other narrative protocols. Specifically, here we lay out a procedure that guides researchers through the four main phases of the autobiographical narrative pipeline: (1) data collection, (2) transcribing, (3) scoring, and (4) analysis. First, we provide recommendations for incorporating transcription software to augment human transcribing. We then introduce an electronic scoring procedure for tagging narratives for scoring that incorporates the traditional AI scoring method with basic keyboard shortcuts in Microsoft Word. Finally, we provide a Python script that can be used to automate counting of scored transcripts. This method accelerates the time it takes to conduct a narrative study and reduces the opportunity for error in narrative quantification. Available open access on GitHub (https://github.com/cMadan/scoreAI), our pipeline makes narrative methods more accessible for future research.

Keywords Autobiographical interview · Autobiographical memory · Narratives · Scoring

Research into the phenomenon of memory has interested psychologists for well over a century (Ebbinghaus, 1885; James, 1890). First popularized by Ebbinghaus, studies of human memory for stimuli generated in a laboratory, such as words and pictures, have produced myriad insights into the characteristics of memory. With tight control of encoding conditions, this approach continues to generate vital findings in the field. Yet, aspects such as deep personal relevance and entangled multi-sensorial environments cannot be captured within a laboratory setting. Accordingly, the 1970s saw a growing interest in exploring memory for real-world experiences, namely, autobiographical memory (AM). This prompted debate as to whether findings from laboratory-based studies could translate to AM and, in turn, how AM might be measured for scientific study (e.g., Banaji & Crowder, 1989; Neisser, 1978, 1982). The

works of Tulving (1972), Crovitz and Schiffman (1974), Rubin, Wetzler, and Nebes (1986), Neisser (1982), Kopelman, Wilson, and Baddeley (1989), and Conway and Bekerian (1987) brought forth new methods for quantifying AM, expanding the scope of memory research by integrating the systematic study of participant narratives (see Sheldon et al., 2018, for a recent review). Critically, research has shown that AM performance can be dissociated from performance on laboratory tests of memory, both in terms of behavioral and neural processes (Conway & Rubin, 1993; Diamond, Abdi, & Levine, 2020; Gilboa, 2004; Leport, Stark, Mcgaugh, & Stark, 2017; Mcdermott, Szpunar, & Christ, 2009; Palombo, Alain, Söderlund, Khuu, & Levine, 2015).

Accordingly, today AM is recognized as a vital field of study, armed with vast implications for understanding healthy individuals, aging, dementia, amnesia, depression, post-traumatic stress disorder (PTSD), and beyond. To briefly highlight a few examples, studies of AM have shown that, relative to younger adults, older adults produce a paucity of episodic details from events. Yet, they produce an augmented number of external (non-event-specific) details—a pattern attributed to compensatory processes to "fill in" for impoverished episodic detail (Levine et al., 2002; also see Gaesser, Sacchetti, Addis, & Schacter, 2011; Devitt, Addis, & Schacter, 2017; Addis, Musicaro, Pan, & Schacter, 2010).

✉ Daniela J. Palombo
daniela.palombo@ubc.ca

1   Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC V6T 1Z4, Canada

2   School of Psychology, University of Nottingham, Nottingham, UK

Springer

A similar pattern has been observed in PTSD (Brown et al., 2014). AM research in depression has revealed reduced retrieval of specific memories (e.g., Dalgleish et al., 2007; Williams et al., 2007), with more recent work showing that this may be a more pervasive deficit that extends to general (categorical) AMs as well (Hitchcock et al., 2019). Other work shows a deficit in the production of episodic AM details in depression (Söderlund et al., 2014). AM approaches have been particularly useful in shedding light on the nature of remote memory loss (which cannot be readily captured by laboratory approaches) in patients with certain forms of amnesia (e.g., Reed & Squire, 1998; Nadel, Samsonovich, Ryan, & Moscovitch, 2000; Irish et al., 2011).

To study AM, researchers rely on participants to narrate their personal past experiences[1]. However, as crucial as narrative studies are in providing real-world context to research, they can also be elaborate and time-consuming. This paper aims to facilitate narrative methodologies in AM research by providing a simple protocol for augmenting processing and scoring procedures. We have developed a semi-automated paperless transcribing and scoring protocol that employs computer programming to accurately and automatically summate data to bolster (but not replace) the efficiency of manual transcribing and scoring (Fig. 1). This paper provides a guide for running narrative studies using this approach (with documentation and code included), making this methodology more accessible for future AM research. An ancillary goal is to provide some "best practices" to further facilitate transcribing and scoring narratives (see Adler et al., 2017; Syed & Nelson, 2015).

Data collection

The current procedure was developed with data collected using the "Autobiographical Interview" (AI) protocol (see Levine et al., 2002; also see Addis, Wong, & Schacter, 2008). The AI is a standardized semi-structured interview and scoring method that has been used to examine autobiographical narratives in over 200 narrative studies (see AutobiographicalInterview.com).[2] Briefly,

Fig. 1 Stages of the semi-automated transcribing and scoring procedure

in the AI protocol, participants are asked to select events from their lives that are specific to a time and place, i.e., episodic memories, and then to describe these events in as much detail as possible (i.e., "Free Recall").[3]

The narratives that participants provide are aided by two stages of probes to elicit more mnemonic details: The "General Probe" prompts participants to recall any additional details or helps guide them towards recalling a specific event, if one was not selected in Free Recall. Finally, the "Specific Probe" consists of direct questions related to the experience of the event (for further information, see Levine et al., 2002). For a variety of reasons, some researchers opt not to administer the Specific Probe, although it can be very useful, particularly in clinical populations, wherein the additional probes provide scaffolding support for cueing memory recall. In the protocol described

---

[1] A similar approach is used in studies of autobiographical imagination, wherein participants are asked to narrate an event imagined in a specific context (e.g., "Imagine catching your grandchild getting into trouble twenty years from now"; e.g., Race, Keane, & Verfaellie, 2011; Addis et al., 2008), or in studies of counterfactual thinking, wherein participants are asked to think about what could have been (e.g., De Brigard et al., 2016).

[2] A wide array of other interview structures have been employed to capture a participant's narrative for analysis, such as the Autobiographical Memory Interview (Kopelman et al., 1989) or the TEMPau task (Piolino et al., 2003), and our method can be modified to these other structures as well. Moreover, although our focus is on narrative work in the context of specific events (i.e., situated in a specific time and place), our approach can be modified for studies examining broader autobiographical content, including life stories (e.g., Grilli, Wank, & Verfaellie, 2018), narrative meaning (McAdams & McLean, 2013) or self-referential processing (Kurczek et al., 2015; Verfaellie, Wank, Reid, Race, & Keane, 2019; also see Adler et al., 2017, for a discussion of other approaches).
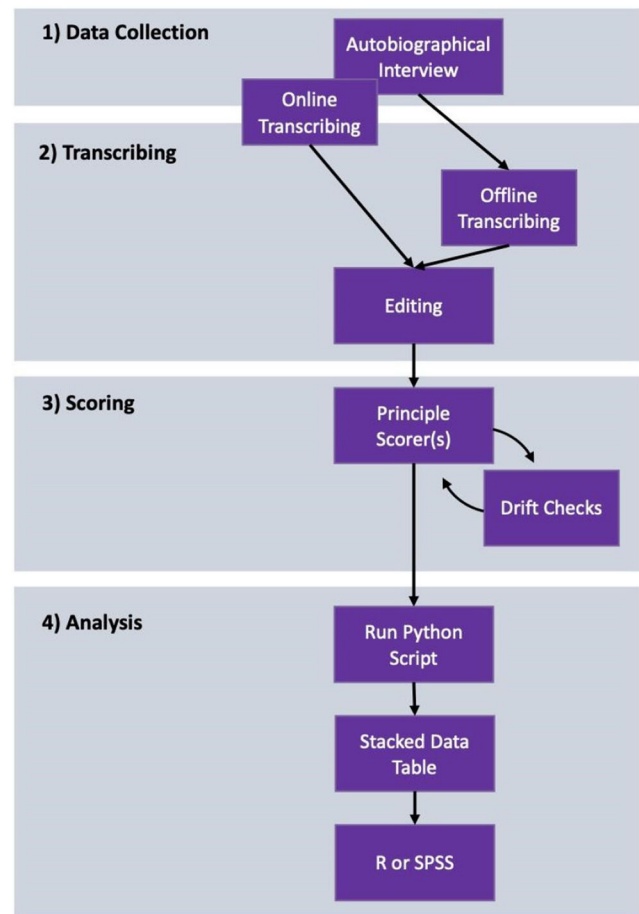
[3] The AI method can also be used with other event selection prescriptions, such as the use of single word cues to elicit specific events (see Addis et al., 2008; also see Crovitz & Schiffman, 1974). We note that more specific cues (e.g., "granddaughter's recital") have been shown to elicit more specific and detailed memories than cue words (e.g., "lemon"), particularly in patient populations. The former may afford greater organizational scaffold to augment memory search (Kwan, Kurczek, & Rosenbaum, 2016). Other work shows that the emotional nature of the retrieval cues can also impact the nature of recall (Sheldon & Donahue, 2017).

below, we do not include the Specific Probe. Although a review of AI findings is beyond the scope of this paper, we note that the AI has been used successfully to characterize patterns of memory performance in studies of healthy individuals, aging, development, patients with brain lesions, psychiatric populations, and neurodegeneration. Other studies have applied the AI to examine imaginative processes, including future thinking and the like (see Sheldon et al., 2018 for review).

The AI administration is captured using a digital recording device for subsequent transcription and scoring. We used a Sony PX370 Mono Digital Voice Recorder due to its user simplicity and long battery life. Placing the recording device between the experimenter and participant in a quiet laboratory room will result in a high-quality recording. Clear audio is paramount, as the recording will provide an original record of the narrative data that will be transcribed for analysis. After a testing session is complete, the audio file can be downloaded and saved to a secure server.

Transcribing

Following the interview, the narrative provided must be transcribed. As the transcripts developed during this process will directly impact how the data are scored, accuracy is paramount. Transcripts must be written verbatim, capturing the interview clearly enough that the written text is as representative of the interview as possible (also see Adler et al., 2017). The process can be tedious and very time-intensive. If you are using a large team of transcribers (as many labs do), developing clear protocols for the format of the final transcripts is crucial to ensure consistency. Whether using a pre-existing transcribing system (such as the Jefferson Transcription System, which captures not only *what* was said but *how* it was said; Jefferson, 1984) or developing your own (as we have done, see Appendix 1), a well-defined protocol outlining punctuation, filler words (e.g., "um" or "like"), and de-identification (e.g., personal names, addresses) relieves some of the difficulty of the transcribing process. Cementing a template document for final transcripts that marks speakers, data identification, and any other information pertinent to the study at hand simplifies the formatting of draft transcripts. As our template is formatted for the Python code used for counting scored details (see below), we recommend using this prescription (see Fig. 2 for an example; also see Appendix 1).

Our laboratory has opted to use Nuance's *Dragon NaturallySpeaking* transcription software (version 15; 2016) to further augment human transcribing, although there are a variety of other transcription software options available that

Fig. 2 Transcription of a mock Free Recall and General Probe formatted in our transcription template and scored with the Autobiographical Interview protocol (Levine et al., 2002; also see Fig. 3 for scoring legend)

one can choose from. Dragon translates audio files into draft transcripts and saves the transcript locally. (As Dragon does not rely on cloud services, it allows for the use of transcription software without added risks to confidentiality.) Notably, there are two ways in which one can employ Dragon. The first is during data collection, wherein Dragon transcribes speech in the moment using voice recognition software (hereafter referred to as "online transcribing") and writes the transcription as a text file such as .doc. To augment accuracy, prior to the interview, participants' voices can be trained on Dragon by selecting from a list of accents (by region, e.g., "US; English with Chinese accent" or "Canada; English with Spanish accent") and having the participant read a brief passage into a microphone. This training tailors Dragon to the individual's unique manner of speech, optimizing its ability to accurately transcribe the speaker.

Alternatively, when Dragon is not used initially, a prerecorded audio file can be fed through the software to generate a draft transcript from the recording (i.e., "offline transcribing"). Launching Dragon at this stage tends to generate a slightly less accurate transcription (even with training). When a recording is of poor quality (e.g., due to excessive background noise or when the participant does not speak clearly), Dragon will not perform well. When this occurs, employing online transcription is even more valuable, as writing a transcript from scratch with challenging audio can greatly impair the progress of the study. Notably, some use the "listen and repeat" technique, in which a researcher, who has trained the software with their own voice, listens to the prerecorded file and vocalizes what they hear, recreating a post-interview automated transcription more akin to our online transcription method (see Matheson, 2007). As Dragon is a self-learning software that improves with use as it "learns" your style of speaking, this approach is advantageous in that Dragon will become more proficient over time at understanding the individual employing this method.

Critically, depending on the method of data collection, dialogue will alternate between the experimenter and the participant at different frequencies throughout the interview. Dragon will produce a continuous block of text that does not differentiate between the participant and the experimenter (or between memories) for both online and offline approaches. Further, Dragon tailors to one voice at a time. This can present challenges in capturing dialogue between two people, i.e., the participant and the experimenter. If opting for online transcribing, one way to mitigate the blocked-text issue is to place your mouse in the appropriate place in a template file before the participant begins speaking (and again when the experimenter begins speaking) as this will separate the text prior to when the speaker changes (or when a new memory begins). However, doing so can interrupt the natural flow of the interview and is thus not encouraged for all studies. This approach may be useful in studies with a simplified protocol (e.g., studies that only employ one probe). Alternatively, we utilize an editing process in which researchers meticulously examine draft transcripts and split the text onto new lines when the speaker changes and when a new memory begins. We find that in doing so, transitioning the draft transcript into our final transcription template is simplified (see Appendix 1). Given the semi-structured format of the AI, the experimenter's speech is scripted and thus can be identified and deciphered with relative ease. In the case of interview procedures with less structured experimenter-participant interactions, the "listen and repeat" technique discussed above may provide the most efficient method, as it capitalizes on the limitation of Dragon to tailor to one voice at a time and allows for the researcher to separate dialogue as they repeat the interview.

Regardless of interview structure, and whether online or offline transcribing is employed, the transcript produced by Dragon must be reviewed for errors by comparing it to the audio recording of the interview. To do so, we opted to use Express Scribe Transcription Software (version 8; 2019) to play back the audio file, coupled with an Infinity IN-USB-2 foot pedal. A foot pedal allows transcribers to easily navigate time-position in the file by simply using their foot to pause, fast forward, and rewind the recording, leaving their hands free for typing. Researchers new to transcribing benefit from slowing the speed of the audio playback; however, with practice, transcriptions can easily be edited in real time. If at any point researchers are unable to discern the dialogue, they insert the word "inaudible", followed by a timestamp noting the point in the interview the stifled audio occurred (see Appendix 1). This allows senior transcribers to easily find unfinished portions of the transcript by searching "inaudible" via the "Find" function in Microsoft Word (also see Footnote 4) and navigate to the appropriate section of the audio file in order to decipher what was said. (Occasionally, the inaudible text cannot be recovered and is noted in the file.)

While editing and formatting raw transcripts is laborious, our experience suggests that editing the initial outputted draft from Dragon is still much less time-consuming than manually writing transcripts directly from the audio file. Conversely, some studies opt for narratives to be collected in written, rather than oral, format (e.g., Ison, 2009). In such cases, namely where transcription is not needed, our scoring protocol (described below) may still be useful.

Scoring

A brief overview of the Levine et al. (2002) scoring protocol

Once transcripts have been edited, data must be extracted from the narratives. By implementing scoring procedures, this qualitative data can be quantified for statistical analysis. Below, we first review key features of the AI scoring procedure (Levine et al., 2002), before turning to our pipeline for augmenting the scoring process. In the AI procedure, details are subdivided into two overarching groups: "internal" and "external." Internal tags

are given to any information pertaining to the event that the participant identified as the memory (e.g., "it was a sunny day in Vancouver"). That is, internal details represent episodic memory. Internal tags are further subdivided into five detail types to classify the content of the information provided regarding the memory: event, perception, emotion/thought, time, and place. External tags are given to any information that does not reflect the specific event (e.g., I always loved Vancouver). This group is further subdivided into detail types, including external events, semantic details, repetition, and other comments (such as metacognitions or clarifications; see Levine et al., 2002, for a full breakdown of detail types). More recent analyses of AI data have expanded these initial categories, for example by classifying semantic details as "personal" or "general" (Strikwerda-Brown, Mothakunnel, Hodges, Piguet, & Irish, 2019; Renoult et al., 2020). The dissociation between types of details has proved important for understanding memory performance in a variety of populations, particularly those discussed above (see Sheldon et al., 2018), but is not reviewed any further here for the sake of brevity. Finally, experimenter ratings are assigned to each narrative, including a rating of episodic richness, which captures the extent to which the participant was able to evoke a sense of re-experiencing an event that is specific in time and place (see Fig. 2). For simplicity, our protocol includes only the episodic richness rating, but the reader is encouraged to see Levine et al., 2002 for the full list of ratings.

Learning the scoring protocol is not trivial and requires practice. Briefly, our laboratory follows the scoring training protocol of the Levine Lab, wherein new scorers practice on an initial pool of memories and then move on to an established set of 20 additional memories for further training. To assess reliability, new scorers are compared to the established Levine training set (comprised of seasoned scorers' data) via intraclass correlations (see Syed & Nelson, 2015; also see Miloyan, Mcfarlane, & Echeverría, 2019, 2019, for a more detailed discussion of best practices for the AI specifically). Under this approach, it is not uncommon to observe the involvement of multiple primary scorers within a study who are randomly assigned memories from a pooled set of narratives. In another common approach, a primary scorer is identified, who scores all the memories from a study, while a second scorer randomly scores a subset of these memories (e.g., 10–20%) so that interrater reliability can be computed and reported.

Regardless of what approach is used, we also recommend performing "drift checks" on additional practice memories (i.e., having multiple scorers compare their scored memories with each other over time) to ensure that scorers within a laboratory are not implicitly deviating from the original AI scoring over time. Moreover, where possible, it is optimal for the scorers to be blind to experimental conditions or groups.

Paperless scoring pipeline Traditionally, scoring is often done on paper copies of the interview. However, our procedure offers researchers a paperless method of scoring, which not only reduces resources, but importantly, also minimizes the chance of error: by scoring in an electronic format, tallying what was scored can be accomplished automatically by computer software as opposed to by hand. Some software packages exist that automate the scoring and tallying procedure, such as the commercially available and general purpose NVivo software (version 12; 2018) or the freely available "Autobiographical Interview Scoring" (AIS) software. NVivo allows for themes to be coded in transcripts (e.g., each detail type from the AI can be coded as a theme) from which a report is produced, providing the total number of references to the theme in each transcript as well as the raw text that was initially coded. In contrast, the AIS is designed specifically for use with the AI (Wickner, Englert, & Addis, 2015) and allows for digital scoring and tallying of details that can be exported into a spreadsheet for analysis. Here, we have developed an additional pipeline for scoring that can be tailored to a range of AI procedures, dubbed "scoreAI" (Scoring the Autobiographical Interview). Our protocol is conceptually similar to the AIS, but is more extensive, as it spans the entire processing pipeline from transcription to analysis.

First, to insert a given detail tag within a narrative, we altered keyboard shortcuts in Microsoft Office 365 Word 2019[4] so that simple keystrokes would result in complete tags for detail types after the relevant to-be-scored clause (see Appendix 2 for complete instructions). That is, whenever a detail is identified, a tag would be inserted into the Word document via the tailored keyboard. For example, if the experimenter wanted to score a detail, such as "we were at the Cheesecake Factory", as an internal place detail, they would insert a tag (in this case "Int_PL") after the appropriate clause (see Fig. 2). We also created a keyboard cover using a keyboard skin protector cover to assist the scoring process (see Fig. 3 for a schematic).

Analysis

Finally, to automatically summate all the detail types, the scored transcripts were fed through a Python (2020) script and saved in a .csv file for subsequent analyses. This Python script is provided along with instructions for use (see Appendix 2). Briefly, this Python script uses the python-docx module (Canny, 2019) to read in the Microsoft Word document, based on the formatting indicated earlier (i.e., the template), and isolates the portions of the text associated with each transcribed and scored memory. The counts associated with each of the tags from the scoring procedure are then calculated, along with extracting the episodic richness rating value, and collated into a summary table. This procedure is then repeated for all available Word documents to generate a single summary table for all participants and all scored

---

[4] For free and open software alternatives, our protocol can be used in conjunction with Google Docs or OpenOffice.

Category	Detail Type	Keyboard Tag	Keyboard Shortcut
Internal	Event	Int_EV	Ctrl+F
	Perceptual	Int_PERC	Ctrl+D
	Emotional	Int_EMO	Ctrl+S
	Place	Int_PL	Ctrl+A
	Time	Int_TM	Ctrl+E
External	Event	Ext_EV	Ctrl+J
	Semantic	Ext_SEM	Ctrl+K
	Repetition	Ext_REP	Ctrl+L
	Other	Ext_OTH	Ctrl+;
Rating	Episodic Richness	[ER−]	Ctrl+O

Fig. 3 Top: Legend for keyboard shortcuts for internal and external details and rating from the Autobiographical Interview protocol (Levine et al., 2002). Bottom: Example layout for a keyboard cover with keyboard shortcuts designated for each detail type. Stickers can be used to label the keyboard

memories. We note that researchers differ in terms of whether they examine individual detail types or composite internal versus external scores. Moreover, depending on the goal of the study, the researcher may opt to control for verbal output by computing an internal-to-total ratio score (see Miloyan et al., 2019).

To capture nuances of different applications of the AI, or other narrative methodologies, modifications to the template or code may be needed. Both the template and code are easily adaptable to such modifications. For example, the "tags" used for scoring can be easily adjusted in the code to tailor output towards the intended measure (for additional information, see https://github.com/cMadan/scoreAI; Madan, 2020).

We provide three practice memories and an accompanying output file for the reader to run through the code to ensure the Python script is being used correctly (see Appendix 2, Fig. 5). (We note that these practice memories are scored based on our interpretation of the Levine et al., 2002, protocol and the instructions provided by the Levine laboratory.) We encourage the reader to perform "spot checks" on a small subset of

their actual data to ensure that the outputted Python results line up with manual counting.

Discussion

In the current paper we presented a novel, semi-automated, paperless transcribing and scoring procedure tailored to AM research, particularly research that employs the AI protocol (Levine et al., 2002). For transcribing, we presented two ways of applying automatic transcription software (in this case, Dragon) to aid the transcribing process of participant interviews. Transcribing software does not replace human labor but accelerates it considerably. We also provided some recommendations for editing transcriptions to ensure consistency across narratives.

We then introduced an electronic scoring procedure for AM details that incorporates basic keyboard shortcuts in Microsoft Word to facilitate the standard Levine et al., 2002 scoring procedure. We also introduced a simple Python script

(scoreAI) written by our group that performs automated detail counting and generates a user-friendly output file. The data in the output can then be easily analyzed with a variety of statistical procedures.

Although these procedures do not eliminate the time commitment and human labor required for AM narrative studies, they extricate and reduce error, making this methodology more accessible for future research.

Appendix 1: Transcription manual

The following is an example of our laboratory's protocol for transcribing narratives. We use a Microsoft Word.docx file as a template for each participant (as shown below). This document is also available for download in Supplementary Materials (template.docx).

General transcribing instructions

1 Open the transcript template and fill out the following information:

    a Participant ID
    b Transcriber (that's you)
    c Experimenter
    d Date of Testing

2 Open the participant's raw Dragon outputted transcript
3 Open the participant's MP3 file with Express Scribe software
4 Connect an Infinity pedal to the USB port (optional)
5 Edit the raw Dragon transcript (following the guidelines listed below under "editing transcripts")
6 Insert the dialogue into the appropriate section of the template

    a It is possible that there will be some dialogue between the experimenter and participant that requires you to add in additional 'Experimenter:' or 'Participant:' lines to the template

7 Go through the transcript and:

    a Bold everything said by the experimenter
    b Double-space everything said by the participant
8 Spell-check the transcript
9 The transcript is now ready for a second transcriber to perform a quality check and sign off on the cover sheet (when complete).

Editing transcripts

- Transcripts are verbatim. This means any ums, ahs, or stutters are documented. Dragon will not do this for you
- When the audio is unclear, type **(inaudible 00:00)**, where 00:00 is the time stamp for the words you could not hear
- For consistency, it is recommended that you use the following spelling for filler/shortened words:
    - Hyphens are used when a word or sentence is not finished

Fig. 4 Screenshot of the transcription template

Hmm	Um hmm	Wanna	Shoulda
Mm	Y'know	Lemme	Coulda
Uh	Yeah	Tryna	'em
Ah	Yep	Kinda	'cause
Um	Dunno	Gotta	Goin'
Uh huh	Gonna	Woulda	Doin'

Ex: We **were – Well** we didn't want to go

- Commas are used when words are repeated and around filler words

  Ex: **I, I, I** was so tired
  Ex: I, **um,** wondered what to do

- Ellipses ("…") are used when participants pause for an extended period of time. Be careful not to overuse this; it is only necessary for long pauses (e.g., more than 3 seconds)
- Round brackets are used to mark noises that are not words

  Ex: **(sighs)** or **(laughs)**

- Square brackets are used to conceal identifiers. It is important that we do NOT include anything that could identify the participant in the transcript

- Names:
  Include the names of public names such as celebrities or scholars
  Ex: I always loved poetry
  Do not include any names mentioned that have personal relationships with the participant
  Ex: **[participant's name]** or **[participant's boyfriend]**

- Places:
  Include the names of locations that might offer important context
  Ex: I grew up in Shanghai
  Do not include the names of places that might identify the participant
  Ex: I went for a run in **[name of park in Vancouver]** because it's so close to my place

- Quotation marks are used when someone says what they or someone else said but not for things the speaker thought to themselves

Ex: She was like**, "**Don't you think we should tell them**?"**
Ex: And in my head I was like, what are you talking about?

- Numerals follow APA format

- Spell numbers one to nine
- Use numerals for numbers greater than 10
- Use numerals for years and dates
- Use numerals for time. Only include a.m./p.m. if the speaker says it

Appendix 2: Keyboard configuration and Python code for detail counts

Part 1: creating the keyboard

A   Creating AutoText

Microsoft Word allows for common segments to be saved as 'AutoText'. Not only are characters saved in an AutoText, all formatting, including style and highlights, are saved as well. Create separate AutoTexts for each detail type you are scoring for.

1 Open a new Word document
2 Type out exactly what text/format you want to appear and highlight it
3 Go to the Insert tab
4 Select 'Quick Parts' drop down menu
5 Select 'Save selection to Quick Part Gallery'
6 The name will automatically be filled in with the text you have selected
7 Select 'AutoText' in the 'Gallery' tab
8 Assign the AutoText to the appropriate category, in this case 'AutoBio_Scoring'
9 Click 'OK'

B.   Creating Keyboard Shortcuts

ParticipantID	Memory	Time Period	Condition	Int_EV	Int_PERC	Int_EMO	Int_PL	Int_TM	Ext_EV	Ext_SEM	Ext_REP	Ext_OTH	ER
99	1	Remote	Negative	23	15	8	6	0	0	3	1	1	6
99	2	Recent	Neutral	11	5	2	4	1	0	2	1	2	4
99	3	Remote	Positive	4	2	2	3	1	0	0	0	5	2

Fig. 5 Screenshot of a study "master sheet" for three memories for one participant. The rows represent the detail counts for the scored memories provided in the Supplementary Material

To insert the AutoText efficiently into transcripts, you can use keyboard shortcuts. Assign a keyboard shortcut to each AutoText.

1. Open Word Options by pressing Alt+F+T
2. Select the 'Customize Ribbon' tab in the left hand menu
3. Select the 'Customize' button
4. In the 'Press new shortcut key' box enter the key you wish to use (e.g., Ctrl+F)
5. Scroll through the 'Categories' list and select 'Building Blocks'
6. Scroll through the 'Commands' list and select the AutoText you created in step one
7. Click the 'Assign' button to assign the shortcut
8. Click 'Close'

Once this process is complete, the newly created keyboard shortcuts can be used during scoring.

Part 2: running the Python code for counting details

The Python script for counting details in the Word documents can be downloaded from https://github.com/cMadan/scoreAI (current version is build 10). The script is comprised of five sections.

The first section requires the user to configure the code and should be adjusted on a case-by-case basis. The options to configure include specifying the directory that has the input Word documents, the folder to output the stacked data to, and the number of memories in each Word document. The specific filename of the Word documents does not matter, though the script will load the files in alphabetical order and assumes no other files are in this input directory. Each Word document is expected to have the number of memories configured and be formatted as specified in the template. For an example of a scored memory document, see example_scoring.docx in the Supplementary Materials.

The second section and onwards should not be modified unless changing the overall functionality of the script (e.g., using a different document template or changing the memory labels. The second section of the code loads several Python modules into the environment for the script to use in the processing of the documents. The only non-standard Python module that is required is python-docx, which can be installed using the pip program (see https://python-docx.readthedocs.io/en/latest/user/install.html).

The third section defines the memory scoring labels (e.g., Int_EV, Ext_SEM), looks up the list of files in the input directory, and includes additional 'under the hood' settings.

The fourth and fifth sections do most of the actual work. The fourth section defines several functions that will need to be used repeatedly, such as for extracting specific paragraphs of text from the document and counting the number of occurrences for each scoring label. The fifth section of code brings it all together, cycling through each document, first extracting the participant ID and episodic richness ratings. The code then goes through and finds the start of each memory within the document and then uses these to extract the related text and calculate the component memory scores. These scores are then converted into a single data record along with the participant ID and episodic richness values, such that each memory is its own row. This then continues until all of the documents are processed and iteratively merged together. The final section of code converts these records into a dataframe format and then outputs them as a CSV into the designated output folder, with the filename including the number of documents (i.e., number of participants) and current date. An example output file, corresponding to the example scored memory document, is provided as example_output.csv.

Figure 5 is an example of a filled out sheet, which is a formatted version of the output file from Python, displayed in Excel: The three rows represent the detail counts for the memories provided in the scoring example. Other variables (e.g., time period, condition) are shown for display purposes.

References

Addis, D. R., Musicaro, R., Pan, L., & Schacter, D. L. (2010). Episodic simulation of past and future events in older adults: Evidence from an experimental recombination task. *Psychology and Aging, 25*(2), 369.

Addis, D. R., Wong, A. T., & Schacter, D. L. (2008). Age-related changes in the episodic simulation of future events. *Psychological Science, 19*(1), 33–41.

Adler, J. M., Dunlop, W. L., Fivush, R., Lilgendahl, J. P., Lodi-Smith, J., Mcadams, D. P., ... & Syed, M. (2017). Research methods for studying narrative identity: A primer. *Social Psychological and Personality Science, 8*(5), 519–527.

Banaji, M. R., & Crowder, R. G. (1989). The bankruptcy of everyday memory. *American Psychologist, 44*(9), 1185.

Brown, A. D., Addis, D. R., Romano, T. A., Marmar, C. R., Bryant, R. A., Hirst, W., & Schacter, D. L. (2014). Episodic and semantic components of autobiographical memories and imagined future events in post-traumatic stress disorder. *Memory, 22*(6), 595–604.

Canny, S. (2019). Python-Docx Python Module. https://www.pypi.org/project/python-docx/

Conway, M. A., & Bekerian, D. A. (1987). Organization in autobiographical memory. *Memory & Cognition, 15*(2), 119–132.

Conway, M. A., & Rubin, D. C. (1993). The structure of autobiographical memory. In A. F. Collins, M. A. Conway, P. E. Morris (Eds.), *Theories of memory* (pp. 261-88). Hillsdale, New Jersey: Erlbaum.

Crovitz, H. F., & Schiffman, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomic Society, 4*(5), 517–518.

Dalgleish, T., Williams, J. M. G., Golden, A. M. J., Perkins, N., Barrett, L. F., Barnard, P. J., … Watkins, E. (2007). Reduced specificity of autobiographical memory and depression: The role of executive control. *Journal of Experimental Psychology: General, 136*(1), 23.

De Brigard, F., Giovanello, K. S., Stewart, G. W., Lockrow, A. W., O'Brien, M. M., & Spreng, R. N. (2016). Characterizing the subjective experience of episodic past, future, and counterfactual thinking in healthy younger and older adults. *The Quarterly Journal of Experimental Psychology, 69*(12), 2358–2375.

Devitt, A. L., Addis, D. R., & Schacter, D. L. (2017). Episodic and semantic content of memory and imagination: A multilevel analysis. *Memory & Cognition, 45*(7), 1078–1094.

Diamond, N. B., Abdi, H., Levine, B. (2020). Different patterns of recollection for matched Real-World and Laboratory-Based Episodes in Younger and Older Adults. *Cognition 202*, 104309.

Dragon Naturally Speaking Voice Recognition Software (Version 15) [Computer Software]. (2016). Burlington, MA: Nuance Communications.

Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology.* New York: Dover Publications.

Express Scribe Transcription Software (Version 8) [Computer Software]. (2019). Greenwood Village, CO: NCH Software.

Gaesser, B., Sacchetti, D. C., Addis, D. R., & Schacter, D. L. (2011). Characterizing age-related changes in remembering the past and imagining the future. *Psychology and Aging, 26*(1), 80.

Gilboa, A. (2004). Autobiographical and episodic memory—One and the same?: Evidence from prefrontal activation in neuroimaging studies. *Neuropsychologia, 42*(10), 1336–1349.

Grilli, M.D., Wank, A.A., & Verfaellie, M. (2018). The life stories of adults with amnesia: Insights into the contribution of the medial temporal lobes to the organization of autobiographical memory. *Neuropsychologia, 110*, 84–91.

Hitchcock, C., Rodrigues, E., Rees, C., Gormley, S., Dritschel, B., & Dalgleish, T. (2019). Misremembrance of things past: Depression is associated with difficulties in the recollection of both specific and categoric autobiographical memories. *Clinical Psychological Science, 7*(4), 693–700.

Irish, M., Hornberger, M., Lah, S., Miller, L., Pengas, G., Nestor, P. J., … Piguet, O. (2011). Profiles of recent autobiographical memory retrieval in semantic dementia, behavioural-variant frontotemporal dementia, and Alzheimer's disease. *Neuropsychologia, 49*(9), 2694–2702. https://doi.org/10.1016/J.Neuropsychologia.2011.05.017

Ison, N. L. (2009). Having their say: Email interviews for research data collection with people who have verbal communication impairment. *International Journal of Social Research Methodology, 12*(2), 161–172.

James, W. (1890). Attention. *The Principles of Psychology, 1*, 402-458.

Jefferson, G. (1984). Transcript notation. In Atkinson, J. M., Heritage, J., & Oatley, K. (Eds.), *Structures of social action: Studies in conversation analysis* (pp. ix-xvi). Cambridge: Cambridge University Press.

Kopelman, M. D., Wilson, B. A., & Baddeley, A. D. (1989). The autobiographical memory interview: A new assessment of autobiographical and personal semantic memory in amnesic patients. *Journal of Clinical and Experimental Neuropsychology, 11*(5), 724–744.

Kurczek, J., Wechsler, E., Ahuja, S., Jensen, U., Cohen, N. J., Tranel, D., & Duff, M. (2015). Differential contributions of hippocampus and medial prefrontal cortex to self-projection and self-referential processing. *Neuropsychologia, 73*, 116–126. https://doi.org/10.1016/J.Neuropsychologia.2015.05.002

Kwan, D., Kurczek, J., & Rosenbaum, R. S. (2016). Specific, personally meaningful cues can benefit episodic prospection in medial temporal lobe amnesia. *British Journal of Clinical Psychology, 55*(2), 137–153.

LePort, A. K., Stark, S. M., McGaugh, J. L., & Stark, C. E. (2017). A cognitive assessment of highly superior autobiographical memory. *Memory, 25*(2), 276–288.

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging, 17*, 677–89. https://doi.org/10.1037//0882-7974.17.4.677

Madan, C. R. (2020). scoreAI Python Code. https://github.com/cMadan/scoreAI

Matheson, J. L. (2007). The voice transcription technique: Use of voice recognition software to transcribe digital interview data in qualitative research. *Qualitative Report, 12*(4), 547–560.

McAdams, D. P., & McLean, K. C. (2013). Narrative identity. *Current Directions in Psychological Science, 22*(3), 233–238.

McDermott, K. B., Szpunar, K. K., & Christ, S. E. (2009). Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia 47*, 2290–2298.

Miloyan, B., Mcfarlane, K., & Echeverría, A. V. (2019). The adapted autobiographical interview: A systematic review and proposal for conduct and reporting. *Behavioural Brain Research 370:*111881

Nadel, L., Samsonovich, A., Ryan, L., & Moscovitch, M. (2000). Multiple trace theory of human memory: Computational, neuroimaging, and neuropsychological results. *Hippocampus, 10*(4), 352–368. https://doi.org/10.1002/1098-1063(2000)10:4<352::AID-HIPO2>3.0.CO;2-D

Neisser, U. (1978). Memory: What are the important questions?. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 3–24). London: Academic Press.

Neisser, U. (Ed.) (1982). *Memory observed: Remembering in natural contexts.* San Francisco: Freeman.

NVivo (Version 12) [Computer Software]. (2018). Tehachapi, CA: QSR International.

Palombo, D. J., Alain, C., Söderlund, H., Khuu, W., & Levine, B. (2015). Severely Deficient Autobiographical Memory (SDAM) in healthy adults: A new mnemonic syndrome. *Neuropsychologia, 72*, 105–118.

Piolino, P., Desgranges, B., Belliard, S., Matuszewski, V., Lalevée, C., De La Sayette, V., & Eustache, F. (2003). Autobiographical memory and autonoetic consciousness: Triple dissociation in neurodegenerative diseases. *Brain, 126*(10), 2203–2219.

Python [Computer Software]. (2020). Retrieved from https://www.python.org

Race, E., Keane, M. M., & Verfaellie, M. (2011). Medial temporal lobe damage causes deficits in episodic memory and episodic future thinking not attributable to deficits in narrative construction. *Journal of Neuroscience, 31*(28), 10262–10269.

Reed, J. M., & Squire, L. R. (1998). Retrograde amnesia for facts and events: Findings from four new cases. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 18*(10), 3943–3954. https://doi.org/10.1523/JNEUROSCI.18-10-03943.1998

Renoult, L., Armson, M. J., Diamond, N., Fan, C., Jeyakumar, N., Levesque, L., … St. Jacques, P. L. (2020). Classification of general

and personal semantic details in the autobiographical interview. *Neuropsychologia, 107501.*

Rubin, D. C., Wetzler, S. E., and Nebes, R. D. (1986). Autobiographical memory across the adult lifespan. In DC Rubin (Ed.), *Autobiographical memory* (pp. 202–221). Cambridge: Cambridge University Press

Sheldon, S., Diamond, N., Armson, M. J., Palombo, D. J., Selarka, D., Romero, K., … Levine, B. (2018). Autobiographical memory: Assessment and neuroimaging in health and disease. The Stevens' handbook of experimental psychology and cognitive neuroscience, fourth edition.

Sheldon, S., & Donahue, J. (2017). More than a feeling: Emotional cues impact the access and experience of autobiographical memories. *Memory & Cognition, 45*(5), 731–744. https://doi.org/10.3758/S13421-017-0691-6

Söderlund, H., Moscovitch, M., Kumar, N., Daskalakis, Z. J., Flint, A., Herrmann, N., & Levine, B. (2014). Autobiographical episodic memory in major depressive disorder. *Journal of Abnormal Psychology, 123*(1), 51–60.

Strikwerda-Brown, C., Mothakunnel, A., Hodges, J. R., Piguet, O., & Irish, M. (2019). External details revisited – A new taxonomy for coding 'non-episodic' content during autobiographical memory retrieval. *Journal of Neuropsychology, 13*(3), 371–97. Doi:https://doi.org/10.1111/Jnp.12160

Syed, M., & Nelson, S. C. (2015). Guidelines for establishing reliability when coding narrative data. *Emerging Adulthood, 3*(6), 375–387.

Tulving, E. (1972). Episodic and semantic memory. *Organization of Memory, 1*, 381–403.

Verfaellie, M., Wank, A. A., Reid, A. G., Race, E., & Keane, M. M. (2019). Self-related processing and future thinking: Distinct contributions of ventromedial prefrontal cortex and the medial temporal lobes. *Cortex; A Journal Devoted to the Study of The Nervous System and Behavior, 115*, 159–171. https://doi.org/10.1016/J.Cortex.2019.01.028

Wickner, C., Englert, C., Addis, D.R. (2015). Developing a tool for autobiographical interview scoring. Kiwicam Conference, Wellington, New Zealand. https://github.com/scientific-tool-set/scitos

Williams, J. M. G., Barnhofer, T., Crane, C., Herman, D., Raes, F., Watkins, E., & Dalgleish, T. (2007). Autobiographical memory specificity and emotional disorder. *Psychological Bulletin, 133*(1), 122.