Deconstructing Accurate and Inaccurate Recall in the DRM Paradigm: A Phenomenological and Behavioral Exploration

Jaša Černe • University of Ljubljana, Slovenia • jasa.cerne/at/pef.uni-lj.si
Urban Kordeš • University of Ljubljana, Slovenia • urban.kordes/at/pef.uni-lj.si

> Context • Reliably differentiating between memories of events that happened and those that did not (commonly referred to as "false memories" in the literature) is essential for both personal and societal well-being, yet studies increasingly show that this is nearly impossible to achieve. **> Problem** • Significant support for this conclusion comes from the research using the Deese-Roediger-McDermott (DRM) paradigm, which has been limited in terms of behavioral and phenomenological measures employed, particularly concerning recall memory (as opposed to recognition memory). **> Method** • We used a combination of disciplined first-person methodology and behavioral measures to explore the process of recall in the DRM paradigm. Participants were interrupted during word recall and underwent phenomenological interviews about their recall processes. Reaction times and performance accuracy were measured for each recalled word. **> Results** • We identified a diverse selection of categories pertaining to the phenomenology of recall, describing the mechanisms of attention, structural and processual characteristics, mood and motivation, and other experiences indirectly related to recall. These categories were used to construct the phenomenological indicators of accurate and inaccurate recall. The most coherent accurate recall indicators were related to the experience of strong certainty or the emergence of a word matching the critical lure during encoding as a word that should be avoided during recall. The most coherent inaccurate recall indicators were linked to a deliberate, inferential type of recall or an automatic and fluid recall. Preliminary joint analysis of phenomenological and behavioral data indicated support for the most coherent inaccurate recall indicators, showing varying reaction times associated with different indicators and noticeable differences in the serial position of recalled words corresponding to the indicators. **> Implications** • This study sheds light on the complexity of recall in the DRM paradigm, highlighting the need for a phenomenologically driven mixed-methods approach in a quest to better understand the nature of both accurate and inaccurate memories. **> Constructivist content** • The research design for this study draws heavily on Varela's neurophenomenological framework. **> Key words** • **Constructive memory, DRM paradigm, empirical phenomenology, false memory, memory accuracy, neurophenomenology, reaction times, recall**.

Introduction

« 1 » The ability to rely on our own and others' memories is essential for both personal and societal well-being. Accurate memories play a vital role in the construction of trustworthy personal narratives (Gallo 2010) and contribute to just resolutions of legal disputes (Loftus 1996). Thus, the question of how to separate accurate memories from inaccurate ones (commonly referred to as "false memories" in the literature), has attracted a lot of interest in cognitive science (e.g., Lampinen, Neuschatz & Payne 1997; Heaps & Nash 2001; Kim & Cabeza 2007; Bernstein & Loftus 2009; Marche, Brainerd & Reyna 2010).[1] However, a growing consensus is emerging that, due to the ample contextual variability of inaccurate

1| While researchers typically understand the term "false memories" to mean memories of events that did not happen (Brainerd & Reyna 2005), it is important to note that this definition is rooted in the perspective of the observer (e.g., a researcher) rather than the rememberer (e.g., a participant). A more epistemologically precise definition is as follows: "false memories" are instances when a participant's experiential account of their past encounter with a stimulus significantly deviates from the researcher's perspective. Therefore, in this article, we decided to use the term "inaccurate memories" instead of "false memories" to emphasize (a) that memories should be viewed on a continuum ranging from less accurate to more accurate rather than in binary fashion, where some are "true" and some are "false"; and (b) that, when talking about memory or any other psychological construct, it is not possible to adopt a God's eye view from nowhere, as is assumed in the "true/false" distinction, but one always has to adopt a certain point of view that may or may not be shared with the view of other individuals.

memory occurrences, it is extremely difficult to reliably predict their onset (Brainerd & Reyna 2005; Jou & Flores 2013). In our study, we aimed to challenge this consensus by examining the relationship between phenomenology and behavior via a novel methodological approach that has not yet been used in this setting.

« 2 » The phenomenon of memory inaccuracy has been approached through a diverse array of research paradigms ranging from highly suggestive, where participants are directly exposed to inaccurate narratives, to less suggestive, where inaccurate narratives are implied through the meaning of the study material (Brainerd & Reyna 2005). So far, little work has been done on the latter, i.e., inaccurate memories that are constructed largely by the participants themselves, without help from the environment. By focusing on this type of inaccurate memory, we strove to delve deeper into the endogenous processes of memory, providing a more comprehensive understanding of its associative and constructive nature.

« 3 » One of the most reliable methods for inducing inaccurate memories without explicit external suggestion is the Deese-Roediger-McDermott (DRM) paradigm (Deese 1959; Roediger & McDermott 1995). Typically, the DRM paradigm involves presenting participants with different lists of words where the words from each list (henceforth list items) bear strong associative relatedness to the one common concept (henceforth critical lure), which is not included on the list. For example, a list of words with the critical lure *blossom* may include *flower*, *bee*, *fragrant*, *spring*, *daisy*, *anther*, *meadow,* etc. When participants are tested on their ability to remember list items from a given list, they often report the critical lure as being part of the list (for a review, see Gallo 2006).

« 4 » There are two ways researchers conduct the memory test in the DRM paradigm: with a recall test and a recognition test. The recall test is especially valuable as it allows us to tap into inaccurate memories that are mostly self-produced; instead of asking participants to *recognize* list items among the set of words that also include critical lures, participants are asked, usually after the presentation of each list, to *recall* list items from their memory without receiving any external cues.

« 5 » Research on the phenomenological and behavioral characteristics of recall memory in the context of the DRM paradigm has been scarce and selective in terms of methods employed (for reviews, see Brainerd & Reyna 2005; Gallo 2006; Jou & Flores 2013). In terms of behavioral measures, studies focused almost exclusively on performance accuracy. They showed that the recall rate of critical lures is typically higher than that of unrelated lures (i.e., incorrect reports that are unrelated to the theme of the list), and that, for some lists, the recall rate of critical lures is similar to that of list items (e.g., Roediger & McDermott 1995; Payne et al. 1996; Toglia, Neuschatz & Goodwin 1999; Anastasi, Rhodes & Burns 2000; Iacullo & Marucci 2016).

« 6 » Although speed, fluency, and ease of retrieval occupy a prominent place in several theoretical proposals related to memory inaccuracy (Jacoby & Whitehouse 1989; Koriat 1993; Jou 2008), so far, only two studies have compared recall reaction times (RTs) of accurate and inaccurate memories: Jou (2008) reported slightly faster RTs for list items than for critical lures, while Černe and Kordeš (2023) demonstrated similar RTs for both.

« 7 » To examine the characteristics of accurate and inaccurate recall at the phenomenological level, studies used confidence scales and the remember/know paradigm (Tulving 1985).[2] They showed that list item recall is usually rated as more confident and more frequently involving specific episodic details about a prior event (e.g., sounds, feelings, position in the list) than critical lure recall, which often relies on just a belief in a prior event's occurrence (for a review, see Gallo 2006). However, they also showed that critical lure recall is sometimes rated with similar or even higher confidence, and judged as involving episodic details about a past event as well (Roediger & McDermott 1995; Payne et al. 1996; Read 1996; Toglia, Neuschatz & Goodwin 1999), suggesting that some inaccurate memories are indistinguishable from accurate memories.

« 8 » While confidence scales and the remember/know paradigm allow for fast and uniform data acquisition, their capacity to provide subject-centered and in-depth data is limited (Berkovich-Ohana et al. 2020). For example, the remember/know paradigm reveals only *whether* a participant relived episodic details about a past event, but not which details, how exactly they were experienced, and in what order they appeared. This type of data could prove useful for developing reliable phenomenological indicators of accurate and inaccurate recall, which highlights the need for methods that are sensitive enough to detect richer and more nuanced phenomenology.

« 9 » In the past few decades, various such methods have been proposed in the field of contemporary first-person research (Varela 1996; Depraz, Varela & Vermersch 2003; Petitmengin 2006; Hurlburt 2011; Kordeš & Klauser 2016; Kordeš & Demšar 2021), also known as empirical phenomenology (Kordeš 2016). These methods strive to suspend everyday beliefs, judgments, and explanations about the experience, adopt in-depth phenomenological interviews, and require practice on the side of both researchers and participants. By doing so, they allow for the acquisition of more elaborate and complex phenomenological data (Varela 1996; Petitmengin 2006; Hurlburt 2011; Petitmengin, Remillieux & Valenzuela-Moguillansky 2019). While no published research has yet examined the DRM paradigm using phenomenological methods,[3] researchers

2| The remember/know paradigm maintains that the experience of recall can be divided into either a "remember" response (i.e., recollection), which is commonly associated with accurate memory, or a "know" response (i.e., a feeling of familiarity without recollection), which is commonly associated with inaccurate memory (for reviews, see Brainerd & Reyna 2005; Gallo 2006).

3| Although no published study has so far investigated the phenomenology of recall in the DRM paradigm by using a first-person research method, a similar idea was presented on 4 September 2018 at the seminar "A micro-phenomenological investigation into memory and confabulation" at the Interacting Minds Centre, Aarhus University, Denmark, by Emily Hammond, Katrin Heimann, Chris Allen, Terje Sparby and Maria Gyemant, https://interactingminds.au.dk/events/single-events/artikel/how-to-trust-ourselves-empirical-assessments-of-confabulation-in-introspection.

have already demonstrated their benefits in similar settings (Petitmengin et al. 2013).

« 10 » Inspired by Francisco Varela's neurophenomenological research proposal (Varela 1996; Lutz & Thompson 2003), we conducted an exploratory study of accurate and inaccurate recall in the DRM paradigm by pairing a more sensitive and rigorous first-person method with behavioral measures.[4] Two research questions guided the study:

RQ1 What is the phenomenology of recall within the context of the DRM paradigm?

RQ2 Can we uncover any novel differences between accurate and inaccurate recall by employing a neurophenomenological research design?

In this article, we are going to focus on RQ2 only, while RQ1 will be explored in a separate publication.

Material and methods

Overall procedure

« 11 » The study was conducted entirely online. Zoom (https://zoom.us) or Skype (https://www.skype.com) was used as a medium of online communication between the researcher and the participants. Drawing from the neurophenomenological research program (Varela 1996), contemporary first-person research guidelines (e.g., Petitmengin 2006; Hurlburt 2011; Petitmengin et al. 2019) and the grounded theory approach (Charmaz 2014; Flick 2014), we adopted an exploratory, bottom-up, and iterative approach. This involved:

- bracketing preconceived notions about the phenomenon (Hurlburt 2011; Flick 2014);
- studying participants with an interest and expertise in systematically examining their experience (Varela 1996; Hurlburt 2011; Kordeš & Klauser 2016);
- conducting qualitative analysis and data collection simultaneously, allowing insights from early interviews to inform subsequent ones (Charmaz 2014; Petitmengin et al. 2019);
- conducting qualitative analysis inductively, grounding insights in the concrete reports of participants rather than imposing them based on current knowledge of the phenomenon (Charmaz 2014; Petitmengin et al. 2019); and
- analyzing phenomenological and behavioral data not only independently of each other, but also jointly, by mutually informing each other (Varela 1996; for details, see the *Data analysis* section).

« 12 » Each participant underwent multiple sessions over several days, each of which consisted of listening to and recalling words as part of the DRM paradigm, a phenomenological interview about their experience of recall, and a post-interview questionnaire. RTs needed to recall each word and performance accuracy were measured. To limit potential bias (e.g., suggestive questioning in interviews, researcher/participant confirmation bias), data collection and most of the qualitative analysis were conducted in accordance with the "double-blind" criterion, meaning that neither the researchers nor the participants knew which target words were presented.[5] Pilot studies were conducted prior to the study to inform the final design and provide practice for the researcher in conducting the phenomenological interviews. Figure 1 illustrates the overall research design, which will be explained further in the following subsections.

Participants

« 13 » Eleven younger adults (four females, seven males; mean age = 26.18; *SD* = 1.89) were recruited via email according to non-probabilistic purposive sampling. Only native Slovenian speakers who were experienced with and interested in first-person research were recruited. At the time of conducting the study, eleven participants were postgraduate students in the field of cognitive science, and one participant was a working professional.

« 14 » The study was carried out in compliance with the standards of the Ethical Committee of the Faculty of Education, University of Ljubljana. The participants provided written and oral consent before taking part in the study. To avoid potential influence on the results, the participants were informed that the study aimed to investigate the experience of recall, but not that the study focused on memory accuracy. The participants were fully debriefed upon completion. Phenomenological interviews were conducted with utmost respect for participants' willingness to share their experiences. The participants were assigned codes (e.g., D-S01) to protect their privacy, and their personal details were kept separate from the study results on an encrypted external drive.

The DRM paradigm

« 15 » An open-source tool for designing experiments, PsychoPy v3.0 (Peirce et al. 2019), and an online repository for running PsychoPy experiments, Pavlovia (https://pavlovia.org/), were used to develop and run an online, Slovenian version of the DRM paradigm (Černe & Kordeš 2023). Below, we outline the main characteristics of the stimuli, online application, and task-completion procedure, with further details provided in *Supplementary Material A* at https://constructivist.info/data/19/1/cerne.

« 16 » Stimuli consisted of 36 DRM lists in the form of audio recordings (duration from 48 to 53 seconds, spoken by a male, native Slovenian speaker). Each list started with a beep, followed by 15 words uttered in descending order of association strength. There was a two-second pause between each word, and the list ended with another beep. Eight pseudo-random lists were selected per session for each participant, and their order of presentation was additionally randomized. The participants had the option to delete an already submitted word.

« 17 » The application guided the participants through the following steps: reading the instructions, listening to the audio recording of one list, solving mathematical

4| We use Varela's definition of neurophenomenology, in which "neuro" refers to the "entire array of scientific correlates which are relevant in cognitive science" (Varela 1996: 330). This definition contrasts first-person (phenomenological) data with third-person data, which includes data that is available to public scrutiny and does not rely on subjective reports, including neural (e.g., EEG), psychophysiological (e.g., electro-dermal activity), and behavioral (e.g., RTs) data.

5| While the researcher was familiar with the basic idea of the DRM paradigm (i.e., that each list has a common theme), all but one of the participants (see Footnote 8), were not.

problems for two minutes, and recalling words for another two minutes.[6] The participant pressed the *return* key to write another word, or the *right* key to move to the next list. Each recalled word from a given list appeared on the left side of the screen. The text was presented in white, uppercase font against a gray background.

« 18 » The process of listening to lists, solving mathematical problems, and recalling words, was repeated until the application pseudo-randomly interrupted (i.e., interview cue) the process of recall after the participant entered and confirmed a specific word. The application was designed to distribute interruptions evenly between list items and critical lures throughout the entire recall phase, and to minimize interruptions during the early recall stages, since critical lure recall was seldom observed at that point (Roediger & McDermott 1995). Within one session, the participant listened to a maximum of eight lists, but the interruption typically occurred earlier.

« 19 » RTs were measured for each recalled word without the participants' awareness. The RT of the individual word was defined as the time that elapsed from the confirmation of the previous word by pressing the *return* key (or, for the initial word, when the recall test started) to the confirmation of the current word and so on until two minutes had passed or until the *right* key was pressed, indicating the end of recall.

Phenomenological interview

« 20 » A phenomenological interview inspired by the micro-phenomenological interview technique (Petitmengin 2006) was used to elicit a detailed description of the experience of accurate and inaccurate recall. The focus of the interview was primarily on the experience of recalling the very last word from the previously presented list. All interviews were recorded.

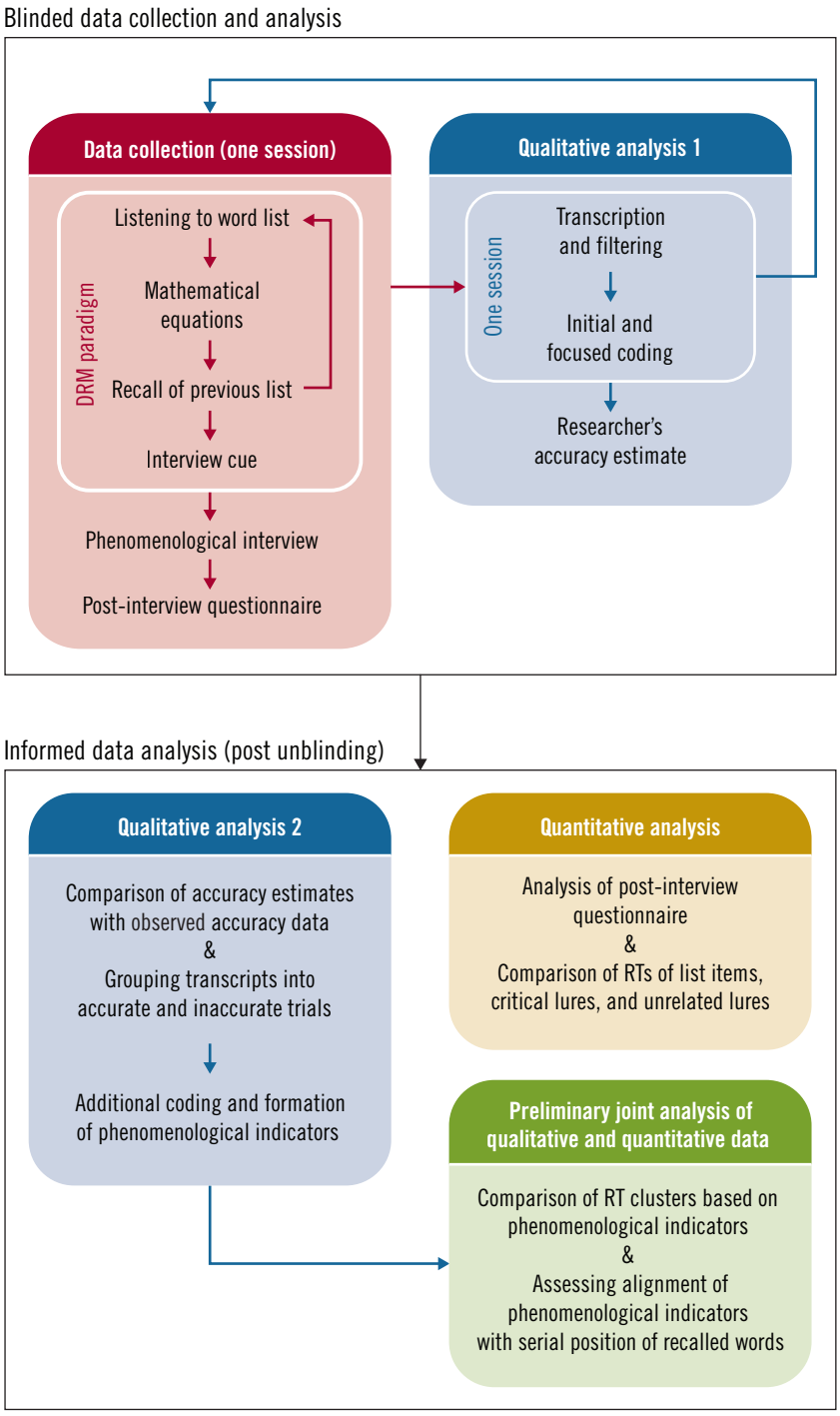

Figure 1 • The overall procedure of data collection and data analysis.

6| The participants were told that they had two minutes to recall as many words as possible. Mathematical problems were introduced based on the findings gathered from the pilot studies and included elementary multiplication and division equations (e.g., 14 × _ = 70). They were used as a distractor task meant to provoke more cases of genuine recall as opposed to just reciting words from short-term memory.

« 21 » When the interruption occurred, a brief pre-interview discussion took place to mutually define the beginning of the experiential episode for later exploration in the interview. The interview began with the researcher suggesting that participant return in their memory to the first moment of the selected experiential episode. The interviewer used open questions aimed not at the content of experience but at its structure. He guided the participant to reflect on and thoroughly describe the diachronic (i.e., the succession of phases and subphases of the experiential episode) as well as the synchronic characteristics (i.e., characteristics co-occurring within a single phase or sub-phase, such as the level of clarity of an inner image or the shape of a bodily sensation) of the selected episode. Whenever the participant started to explain, comment, or theorize about the experience, the researcher redirected the participant's attention back to describing the structure of the experience. When the participant seemed to have difficulty describing the experience, the researcher briefly reformulated what they had already reported.

Post-interview questionnaire

« 22 » A Google Forms questionnaire was used as an additional quantitative measure of how much participants' reports in the interview could be trusted from their own subjective perspective (see *Supplementary Material B*). This questionnaire was developed partly based on the study by Michel Bitbol and Claire Petitmengin (2013) and partly based on our own experience of conducting phenomenological interviews. The questions asked about different aspects of how accurately participants recalled and reported on their experiences during the phenomenological interview. Questions were answered using a seven-point Likert scale, ranging from one (strongly disagree) to seven (strongly agree), plus an open question field for optional comments on each response.

Data analysis

Qualitative analysis

« 23 » The qualitative analysis followed the grounded theory approach (Glaser & Strauss 1967; Corbin & Strauss 2008; Charmaz 2014) and micro-phenomenological analysis method (Petitmengin, Remillieux & Valenzuela-Moguillansky 2019). Although our focus in this article was on identifying recall patterns in accurate and inaccurate recall (RQ2), we summarize below the analysis for both RQs, as they were tightly intertwined.

« 24 » The analysis unfolded in two phases: blinded analysis for RQ1, and informed analysis for RQ2.[7]

« 25 » In the blinded analysis (see the rectangle "Qualitative analysis 2" in Figure 1), we adhered to the double-blind criterion. After each interview, we transcribed verbatim and selected relevant parts focusing on the structure of experience as opposed to the descriptions of content, context, or comments, judgments, and theories about the experience (Petitmengin, Remillieux & Valenzuela-Moguillansky 2019). Using an inductive initial coding (Charmaz 2014), we identified salient characteristics and their relationships, as well as more abstract patterns (i.e., categories), both along the synchronic and diachronic dimensions of experience (Petitmengin, Remillieux & Valenzuela-Moguillansky 2019). While continuing to use initial coding for each newly gathered transcript, we additionally utilized focused coding (Charmaz 2014) to screen these (and previous) transcripts for the most relevant categories. We used a constant comparative analysis (Charmaz 2014) to refine categories. We stopped coding when further analysis yielded no new insights.

« 26 » Before unblinding ourselves, we provided a subjective accuracy estimate for target words by recording whether we believed they were accurate or inaccurate, how certain we were in the estimation, and the reasons for making the estimation.

« 27 » In informed analysis (see the rectangle "Qualitative analysis 2" in Figure 1), we unblinded ourselves, and compared our accuracy estimates with the observed accuracy data to see if researcher blinding was preserved. Using knowledge of which target words were presented, we grouped transcripts as accurate or inaccurate, and, building on the insights from the blinded analysis, carried out additional coding to identify distinct categories and relationships for each group. We organized the results into a codebook, assigning names, descriptions, subordinate/superordinate categories, and examples.

« 28 » Based on the codebook, we assembled phenomenological indicators of accurate and inaccurate recall as follows (see Figure 2). We identified categories clearly featuring more strongly in either accurate or inaccurate recall, called distinctive categories. We used a small subset of these distinctive categories that were most frequent and most exclusive either to accurate or inaccurate recall as grouping variables for assembling phenomenological indicators. Starting from these grouping variables, we recorded the presence of distinct categories in each recall instance and determined their minimal configurations that were *completely* exclusive to either accurate or inaccurate recall. In some cases, we had to include less selective categories in the list of distinctive categories to achieve exclusivity. We defined the resulting configurations as phenomenological indicators. Subsequently, as some of these configurations were linked to multiple other categories, we further assembled second-order indicators, which included one or more additional categories.

Quantitative analysis

« 29 » In the post-interview questionnaire, scores below 4.0 (the middle point) on items 1, 2, 3, and 5, or above 4.0 on item 4, were considered as an indication that some parts of the interview were judged less valid from the perspective of the participant. These were thoroughly inspected and, if necessary, problematic parts of the interview transcript were excluded from the analysis. Results for each participant's session for each of the questionnaire items were additionally averaged to provide an overall estimate of the participant's confidence in their report.

« 30 » In our RT analysis, we compared participants' average RTs of list items, critical lures, and unrelated lures among all recalled words. Additionally, we compared

7| The rationale for blinded analysis was twofold: (a) as data collection unfolded simultaneously with the first phase of qualitative analysis, both had to adhere to the double-blind criterion; and (b) as the first phase of qualitative analysis aimed primarily at identifying the experience of recall *irrespective of the accuracy of target words* (RQ1), we wanted to minimize prior knowledge about the *accuracy* to maximize the validity of the results.

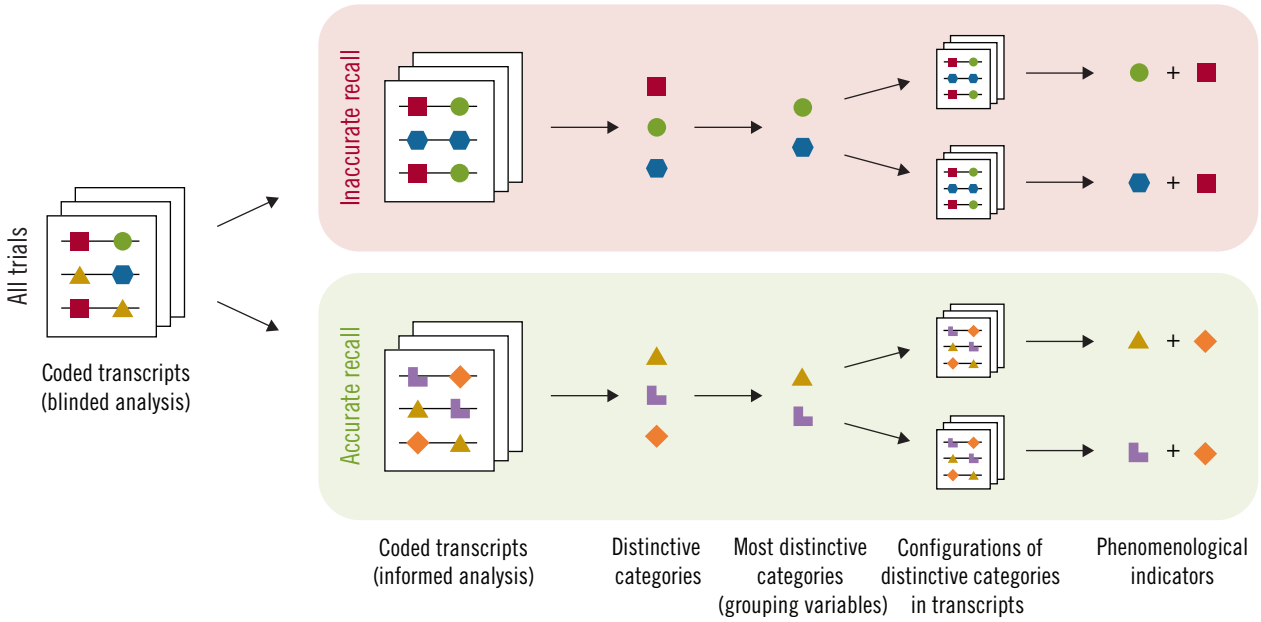

Figure 2 • The process of developing phenomenological indicators of accurate and inaccurate recall. Colored shapes represent categories, and their dissimilarity between accurate and inaccurate recall in terms of frequency and exclusiveness.

participants' average RTs of list items and critical lures for target words only, excluding unrelated lures as they occurred only once. We used the open-source software environment R (https://www.r-project.org). The level for assessing the significance of the statistical results was .05. Because the RT data violated the assumptions of common parametric procedures, we used robust statistical methods from the R package WRS2 (Mair & Wilcox 2020) and bootstrapped version of effect size (Algina, Keselman & Penfield 2005). As per recommendations (Field & Wilcox 2017), we used 20% trimming for all trimmed means procedures and 2,000 samples for all bootstrap-based procedures.

Preliminary joint analysis of phenomenological and behavioral data

« 31 » To mutually analyze phenomenological and behavioral data, we integrated phenomenological indicators with RTs of target words for each trial. It is important to note that, due to small sample sizes (ranging from three to eight RTs per group), we refrained from using inferential statistical procedures. Hence, the results are preliminary and should be interpreted with caution.

« 32 » We analyzed RTs based on phenomenological data. We used accurate recall indicators to cluster RTs of list items, and inaccurate recall indicators to cluster RTs of critical lures, subsequently comparing these clusters separately for accurate and inaccurate recall. In addition, we assessed whether the phenomenological indicators aligned with words recalled earlier or later during the recall phase by separately comparing the average serial positions of recalled words corresponding to accurate and inaccurate recall indicators.

Results

« 33 » We conducted two to five sessions with each participant (on average 4.27 sessions), amounting to 47 sessions altogether. The average time to complete the DRM paradigm was 16 min (*SD* = 10 min) and the average interview length was 98 min (*SD* = 29 min). Overall, the participants recalled 1,332 list items, 26 critical lures, and 23 unrelated lures, of which 20 list items, all 26 critical lures, and one unrelated lure were target words that we further explored in the interview.[8]

« 34 » In the remainder of this section, we present the results of the qualitative and quantitative analysis, as well as the results of the preliminary joint analysis of phenomenological and behavioral data.

8 | We omitted or lost some target word data for the following reasons: (a) due to exclusiveness, we used the data from the single unrelated lure instance (the first session with D-S06) only in the first phase of qualitative analysis, and completely excluded it from the quantitative and joint analysis; (b) we excluded one list item's data (the third session with D-S06) because the participant reported prior experience with a task resembling the DRM paradigm, which likely influenced their recall, as indicated by corresponding phenomenological data; (c) technical issues led to the loss of RT data for two critical lures (the fifth session with D-S02 and the first session with D-S10) and one list item (the first session with D-S03).

Results of the qualitative analysis

« 35 » The qualitative analysis yielded a large set of categories describing the phenomenology of recall in the DRM paradigm. As this article focuses on RQ2, we first provide a brief overview of the complete set of categories and then focus mainly on distinctive categories, i.e., those that were most relevant for distinguishing accurate from inaccurate recall. An exhaustive description of the remaining categories and specifics regarding the diachronic dimension of recall will be explored in a separate publication. In the subsequent sections, we define distinctive categories and provide descriptions of the phenomenological indicators of both accurate and inaccurate recall. Additional insights and concrete interview examples related to these distinctive categories can be found in *Supplementary Material C*.[9]

Phenomenological landscape of the recall process

« 36 » The categories that emerged were grouped into five overarching categories: *attention*, *elements of recall*, *phases of recall*, *mood and motivation*, and *other* (Figure 3).

« 37 » *Attention* constitutes descriptions of attentional direction, gestures, states, and tendencies in recall. When searching for memories, the participants reported that their attention was oriented either inward or outward, and that they performed gestures, such as focusing on, moving away from, or maintaining in awareness different experiential entities. Attention also denotes reports of being more or less deeply immersed in the activity, as well as reports of feeling as though attention was being pulled towards somewhere.

« 38 » *Elements of recall* refers to reports of structural characteristics present in different phases of recall. Among them, we find essential memory entities, such as meta-knowledge or knowledge about the existence and accessibility (but not content) of the relevant elements for recall; memory trace,[10] which refers to a reconstruction of the experience from the encoding phase (i.e., the time of listening to the list of words); and concept, which is a fully developed memory of a particular word. We also find temporal experiences, such as a sense of speed or a sense of sequence in which words were (or should have been) recalled, and various experiences that were particularly distinctive to one of the phases of recall.

« 39 » *Phases of recall* describes four main phases of a recall experience and their types. These phases are: localizing an appropriate memory trace, either by searching the inner mental space or by reading on the computer screen the words that were already recalled; processing that memory trace, be it more deliberate or more automatic; the emergence of a concept, which can be more or less spontaneous and more or less independent from the emergence of other concepts; and an optional phase of verification or testing the accuracy of the concept if the emergence of the concept is accompanied by uncertainty.

« 40 » Next, *mood and motivation* captures reports of more subject-related experiences, such as the participants' tendencies in relation to the task or affective states. The participants reported adopting different attitudes toward the task, varying in terms of motivation, earnestness, self-assurance, and trust. They also reported being in different moods, such as feeling detached from their own bodies or being in a generally bad state of mind.

« 41 » Lastly, *other* groups reports that were important for distinguishing between accurate and inaccurate recall but not directly related to the experience of recall. These include unknowingly conflating two similar concepts during the interview, where one is a critical lure, but the other is not; thinking about the critical lure already in the encoding phase; and experiencing various difficulties during the encoding phase.

Phenomenological indicators of accurate recall

« 42 » Phenomenological indicators of accurate recall consist of the smallest possible configurations of one or more categories distinctive of accurate recall that (i.e., the configurations) emerged within at least one instance of accurate recall and never in inaccurate recall. We present these distinctive categories in Table 1, along with the corresponding phase of the task, accurate and inaccurate recall rate, and the difference between them (see *Supplementary Material C* for additional details and interview excerpts).

« 43 » Some of these categories were more indicative of accurate recall than others. For example, *strong certainty* is the category that emerged most frequently in accurate recall and very rarely in inaccurate recall, and the category *complete critical lure recognition* emerged only in accurate recall and never in inaccurate recall. Accordingly, we used these two categories as grouping variables and organized the phenomenological indicators of accurate recall in a threefold structure: (a) indicators related to *strong certainty*, (b) indicators related to *complete critical lure recognition*, and (c) *other* indicators (see Figure 4). Below, we describe indicators (a) and (b), as they are the most coherent, while the description of indicators (c) can be found in *Supplementary Material D*.

« 44 » The first group of phenomenological indicators of accurate recall is structured around the category *strong certainty*. It encompasses those instances of accurate recall where the participants have explicitly reported high levels of certainty in their memories. The group includes indicators 1 and 2.

« 45 » Indicator 1 describes recall instances in which the participants relied on a

9| Throughout this document and supplementary material, the names of the phenomenological categories are written in italics (e.g., *meta-knowledge*). The descriptions of the indicators sometimes include references to the words that the participants recalled. Because the words were in Slovenian, we structured the initial reference of each word in the following way: *EnglishWord* [*SlovenianWord*; WordType], where *SlovenianWord* stands for the originally recalled word, *EnglishWord* for the English translation of that word, and WordType for the type of the recalled word, which can be either list item, critical lure, or unrelated lure. For example, "*blossom* [*cvet*; list item]" would mean that the list item *cvet* was recalled, which translates to *blossom* in English.

10| In contrast to its frequent usage in reference to the neurological processes underpinning memory functions (e.g., Robins 2018), within this article, we employ the term "memory trace" to specifically refer to observations situated within the domain of the experiential level of analysis. Hence, if not stated otherwise, any claims that we make regarding memory traces should be understood as referring to this domain.

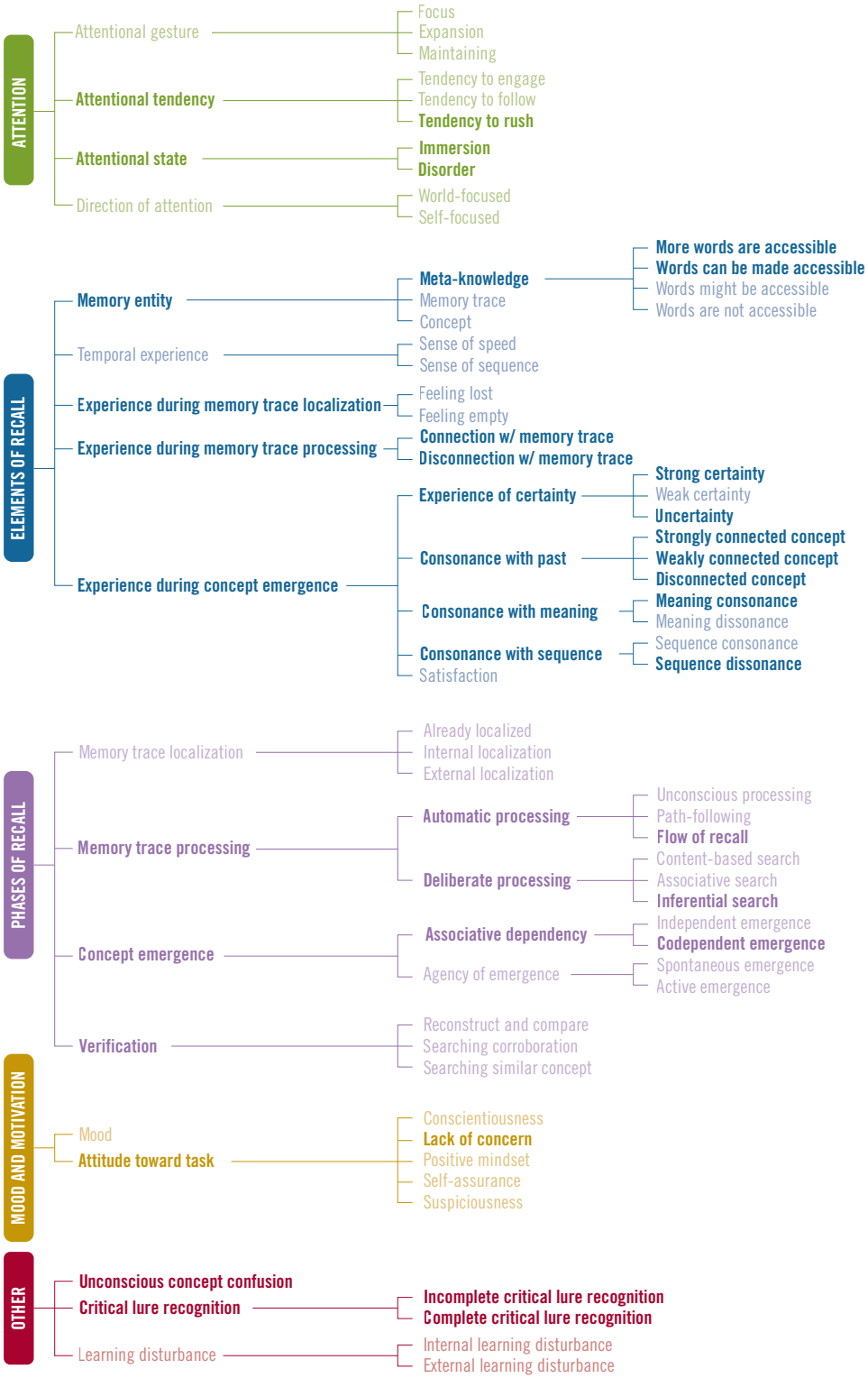

Figure 3 • Complete set of categories describing the experience of recall in the DRM paradigm. The vertically aligned boxes at the left-hand side denote the highest-order categories, serving as the overarching themes for the subordinate categories to their right. Subordinate categories marked in bold represent distinctive categories that were used to assemble the phenomenological indicators of accurate and inaccurate recall.

Category	Task phase	Short description	Accurate recall rate	Inaccurate recall rate
Strong certainty	Concept emergence	A feeling of unequivocal certainty in the recalled word that is explicitly reported by the participant	.58	.04
Connection with memory trace	Memory trace processing	A participant relies on a relevant memory trace while engaged in the phase of memory trace processing	.89	.46
Complete critical lure recognition	Encoding	A critical lure is experienced during the encoding phase, identified as the concept that unites some of the words from the list, and extra care is taken not to recall it in the recall phase	.32	.00
Strongly connected concept	Concept emergence	A recalled word is experienced as distinctively situated in the corresponding memory trace	.42	.12
Words can be made accessible	Memory trace processing	Meta-knowledge indicating that more words can be recalled by active engagement	.47	.19
Weakly connected concept	Concept emergence	A recalled word is experienced as weakly connected to a corresponding memory trace	.37	.12
Unconscious concept confusion	Unclear	A confusion of two similar concepts during the interview, where one is a critical lure	.00	.08
Uncertainty	Concept emergence	A feeling of uncertainty or doubt in the recalled word	.11	.42

Table 1 • Distinctive categories used to assemble phenomenological indicators of accurate recall, organized in descending order based on the difference between accurate and inaccurate recall rates. Task phase = the phase of the DRM paradigm (encoding phase or particular recall phase) that is most distinctly connected to the category; Accurate/Inaccurate recall rate = the rate of occurrence of the category in accurate/inaccurate recall. Rates were calculated based on 19 reports of list item recall and 26 reports of critical lure recall.

memory trace to find a concept (*connection with memory trace*), which was then experienced as at least weakly connected to that memory trace (*weakly connected concept* or *strongly connected concept*) and perceived with absolute certainty (*strong certainty*). For example, in the fifth session, participant D-S01 reported searching for a new word by holding in the periphery of their awareness a mental image of a *tiger* [*tiger*; list item]. This resulted in the emergence of the word *roar* [*rjovenje*; list item] in the form of inner speech, accompanied by a sense of familiarity and a feeling of rightness.

« 46 » Overall, indicator 1 applied to as much as 58% of accurate recall instances and never in inaccurate recall, which makes it the best indicator of accurate recall. Some of these recall instances additionally reflected the presence of *meta-knowledge* indicating a belief that more words could be recalled by active engagement (*words can be made accessible*).

« 47 » In some other recall instances, captured by indicator 1, the participants reported thinking about the word that was also the critical lure already in the encoding phase and recognizing it as important for the task (*complete critical lure recognition*). This scenario suggests that the participants could have been particularly immune to critical lure recall, so we decided to treat the category *complete critical lure recognition* independently of the first group of indicators (see section *Phenomenological indicators of accurate recall related to complete critical lure recognition*).

« 48 » Still other recall instances, captured by indicator 1, did not fit any other distinctive category.

« 49 » Indicator 2 applied to a single instance of accurate recall where the participant did not report experiencing any other characteristic that we deemed indicative of accurate recall apart from being very certain of the recalled word (*strong certainty*). Although this scenario was present in one inaccurate recall instance, the two scenarios differ in that the inaccurate recall instance was additionally characterized by the participant unknowingly mixing two similar concepts during the interview, where one was a critical lure, and the other was not (*unconscious concept confusion*). For example, in the second session, participant D-S01 mentioned first the word *baby* [*dojenček*; list item] and a few minutes later the word *child* [*otrok*; critical lure], both times referring to the same memory entity (a child/baby that was a part of their mental image reconstructed from the encoding phase). Since the confusion went unnoticed by the participant, it might be that it contributed to inaccurate recall. Together with *strong certainty*, indicator 2 therefore includes an additional negative criterion – a lack of *unconscious concept confusion*.

« 50 » There is only one phenomenological indicator of accurate recall related to *complete critical lure recognition* – indicator 3. It is derived from the observation that the participants sometimes reported thinking about the concept that was consistent with the critical lure already during the encoding phase and paying special attention not to recall it. However, according to their reports, they were unaware that this concept constituted a central manipulation of the task. Instead, they

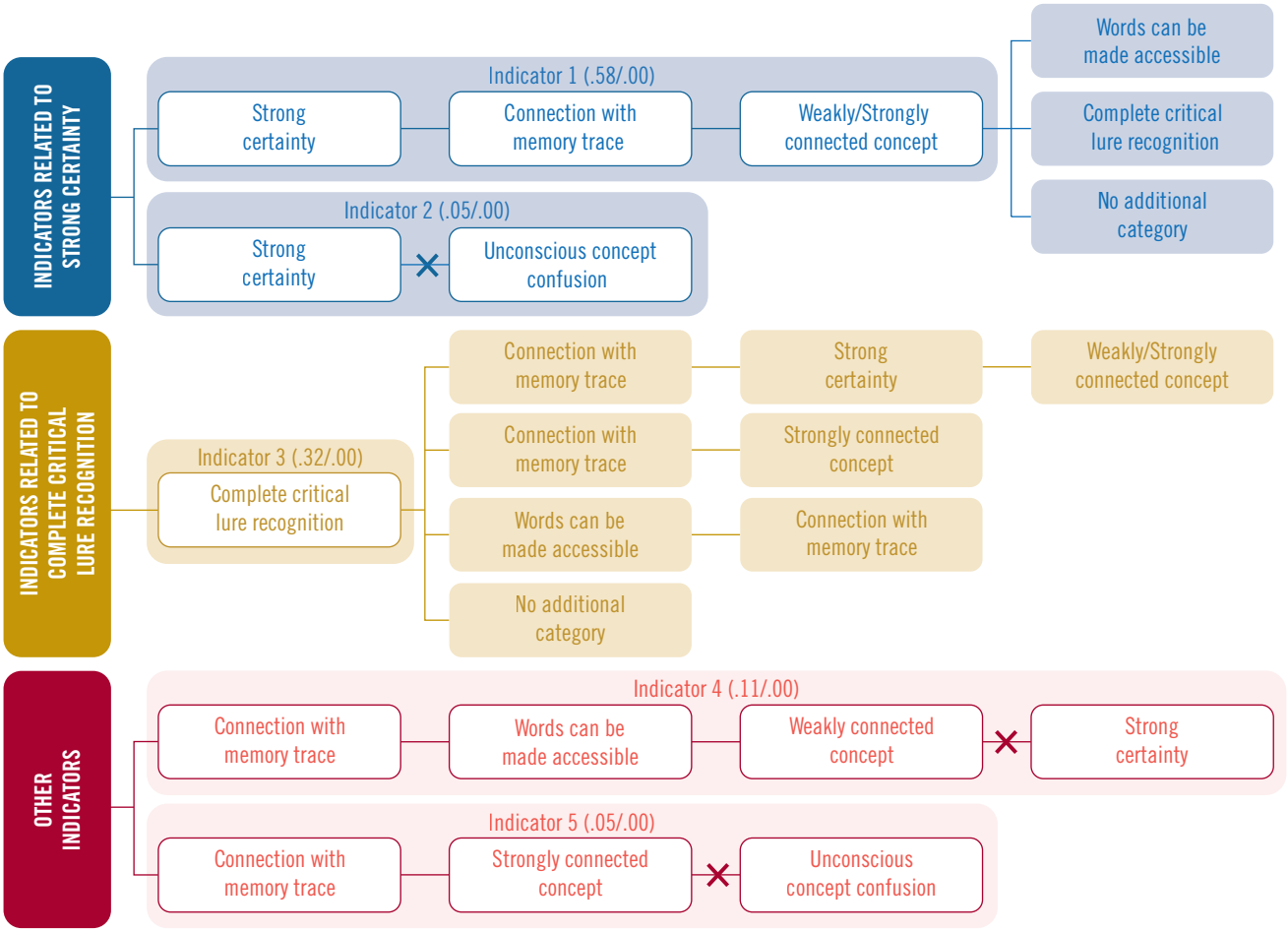

Figure 4 • Phenomenological indicators of accurate recall. The rates of occurrence of indicators in both accurate and inaccurate recall, respectively, are provided in parentheses. The configuration of the indicator's categories is displayed using white-background rectangles, while colored-background rectangles show categories that occurred alongside the indicator's categories less frequently. A crossed line means that the following category must not be present in the configuration of the indicator. The "No additional category" rectangle means that the configuration of categories to its left sometimes appeared without any additional distinctive category. Apart from that, the order of the categories holds no significance.

perceived it as a common theme among the words being encoded that was significant either because it could aid recall through association or lead to confusion. As an example, in the fifth session, participant D-S09 identified *cabbage* [*zelje*; critical lure] as a unifying concept and used it as a means to recall some other words from the previously presented list by inwardly saying *cabbage* together with a new word (e.g., by saying *sour* [*kislo*; list item] – *cabbage*, when the word *sour* was presented but *cabbage* was not).

« 51 » Indicator 3 applied to six (32%) instances of accurate recall, three of which overlap with the instances captured by indicator 1 (characterized by the categories *connection with memory trace, weakly or strongly connected concept*, and *strong certainty*). However, the other three recall instances lacked features of indicator 1. In one of them, the participant's report suggested the presence of the category *complete critical lure recognition* without any other category indicative of accurate recall. In another, the participant reported actively examining the memory trace prior to *concept emergence* (*connection with memory trace*) and distinctly recollecting the experience of the emerged concept from the encoding phase (*strongly connected concept*). In the third one, the participant's report involved parts that are consistent with the categories *connection with memory trace* and *words can be made accessible*, a *meta-knowledge*.

« 52 » The above-listed indicators (together with *other* indicators of accurate recall, described in *Supplementary Material D*) account for 89% (17 out of 19) of instances of accurate recall. For two of the remaining ones, we were unable to identify a convincing distinguishing criterion.

Category	Task phase	Short description	Accurate recall rate	Inaccurate recall rate
Disconnected concept	Concept emergence	A recalled word is not experienced as being connected to (or emerging from) a corresponding memory trace	.11	.65
Meaning consonance	Concept emergence	A feeling that the recalled word fits the general meaning of the set of words being currently recalled	.05	.58
Tendency to rush	All phases	A strong urge to proceed with the next step of the task	.11	.62
Flow of recall *	Memory trace processing	A recall is characterized by deep immersion in the task, automaticity, rapid and fluid unfolding, merging of different concept emergences, a belief that more words will arrive by themselves, and a strong tendency to move on with the next part of the task (with characteristics often extending over multiple recall phases)	.05	.42
Inferential search	Memory trace processing	Deliberate memory trace processing based on making inferences rather than on examining a memory trace	.05	.38
Uncertainty **	Concept emergence	A feeling of uncertainty or doubt in the recalled content	.11	.42
Disconnection with memory trace	Memory trace processing	Memory trace is not clearly present as a subject of examination during the phase of memory trace processing	.05	.35
More words are accessible	Concept emergence	Meta-knowledge indicating that more words are readily accessible for recall and will arrive fluently	.26	.54
Verification	Verification	An optional phase of recall where the accuracy of the concept is tested due to experienced uncertainty	.11	.35
Incomplete critical lure recognition	Encoding	A critical lure is experienced during the encoding phase, identified as the concept that unites some of the words from the list, but no special attention is dedicated not to recalling it in the recall phase	.00	.23
Codependent emergence	Concept emergence	An emergence of one concept is experienced as seamlessly coupled with the emergence of another concept	.37	.58
Disorder	All phases	Broader and scattered attention accompanied by feelings of confusion or distraction	.00	.19
Lack of concern	All phases	Nonchalant approach toward the task with minimal effort put in and disregard for the instructions to only report the previously heard words	.00	.19
Sequence dissonance	Concept emergence	A feeling of dissonance between the sequence in which the word is recalled and the sequence in which it is presented	.00	.19
Immersion	All phases	Intense focus on the task	.21	.38
Unconscious concept confusion **	Unclear	A confusion of two similar concepts during the interview where one is a critical lure	.00	.08

Table 2 • Distinctive categories used to assemble phenomenological indicators of inaccurate recall, organized in descending order based on the difference between accurate and inaccurate recall rates. Task phase = the phase of the DRM paradigm (encoding phase or particular recall phase) that is most distinctly connected to the category; Accurate/Inaccurate recall rate = the rate of occurrence of the category in accurate/inaccurate recall. Rates and differences were calculated based on 19 reports of list item recall and 26 reports of critical lure recall.
* The label "flow of recall" does not directly align with Csikszentmihalyi's concept of flow (e.g., Nakamura & Csikszentmihalyi 2002) but rather reflects the concrete reports provided by participants, which may or may not have been shaped by this idea..
** The categories *uncertainty* and *unconscious concept confusion* are presented in Table 1 and Table 2, as they are found in both accurate and inaccurate indicators.

Phenomenological indicators of inaccurate recall

« 53 » Phenomenological indicators of inaccurate recall consist of the smallest possible configurations of categories distinctive of inaccurate recall that (i.e., the configurations) emerged within at least one instance of inaccurate recall and never in accurate recall. As in the previous section, we first provide a short overview of these distinctive categories and then proceed to describe the phenomenological indicators of inaccurate recall.

« 54 » Table 2 shows distinctive categories used to assemble inaccurate recall indicators, along with the corresponding phase of the task, accurate recall rate, inaccurate recall rate, and the difference between them (see *Supplementary Material C* for additional details and interview excerpts).

« 55 » Some of these categories were more indicative of inaccurate recall than others. For instance, *disconnected concept* and *meaning consonance* appeared very frequently in inaccurate recall and seldom in accurate recall. Similarly, *incomplete critical lure recognition*, *disorder*, *lack of concern*, *sequence dissonance*, and *unconscious concept confusion* appeared only in inaccurate recall and never in accurate recall. Most notably, however, *inferential search* and *flow of recall* exhibited a high frequency of occurrence in inaccurate recall and a low frequency of occurrence in accurate recall, and they occurred in the same recall instance only once. Hence, we used them as grouping variables and organized the inaccurate recall indicators in a threefold structure: (a) indicators related to *inferential search*, (b) indicators related to *flow of recall*, and (c) *other* indicators (see Figure 5). Below, we describe indicators (a) and (b), as they are the most coherent, while the description of indicators (c) can be found in *Supplementary Material D.*

Phenomenological indicators of inaccurate recall related to inferential search

« 56 » The first group of phenomenological indicators of inaccurate recall describes recall instances where the participants engaged in a more deliberate type of *memory trace processing* involving actively inferring words rather than letting them emerge by themselves (*inferential search*). It includes indicators 1, 1.1, 1.2, and 1.3.

« 57 » Indicator 1 denotes inaccurate recall instances that fit the categories *meaning consonance*, and *disconnected concept*. The reports of these instances demonstrated that the participants always based their inferences on the gist of previously recalled words (*meaning consonance*) and experienced emerged concepts as detached from *memory traces* (*disconnected concept*). For example, in the second session, participant D-S03 reported recalling the word *window* [*okno*; critical lure] by logically inferring it from the recognized common theme, "observation," rather than letting themself immerse in the corresponding memory. While listening to the list of words, D-S03 was constructing a mental scene and simultaneously narrating what was going on in the form of inner speech ("I'm looking through something, so it's observation"). In recall, instead of accessing the memory of the scene, D-S03 recalled the narrative and said, "the theme is observation, I was looking out – of course, I was looking through the *window*."

« 58 » Indicator 1.1 applies to seven out of the nine recall instances covered by indicator 1, which additionally displayed the presence of the category *uncertainty*. As an example, in the first session, participant D-S07 reported feeling uncertain about the emerged word because of the inability to inwardly hear the emerged word in the same voice that it was presented in the encoding phase. In 6 out of 7 such instances, the phase of *concept emergence* was typically followed by an additional phase of verifying the concept's accuracy (*verification*). Whether *uncertainty* was followed by *verification* or not, such recall instances were sometimes additionally characterized by the categories *tendency to rush*, *disorder*, *disconnection with memory trace*, *lack of concern*, or *incomplete critical lure recognition*.

« 59 » Two recall instances covered by indicator 1 did not reflect the presence of *uncertainty*. One such instance is described by indicator 1.2, involving a situation where the participant additionally experienced a strong urge to keep moving on with the task even before the concept that was currently being recalled had an opportunity to develop fully (*tendency to rush*). Another such instance is described by indicator 1.3, involving a situation where the participant thought about the critical lure already in the encoding phase but did not consider it important (*incomplete critical lure recognition*); knew, even before the recall of the current word was concluded, that more words will arrive by themselves (*more words are accessible*); and did not rely so much on the *memory trace* in the phase of *memory trace processing* (*disconnection with memory trace*).

Phenomenological indicators of inaccurate recall related to flow of recall

« 60 » The second group of phenomenological indicators of inaccurate recall includes indicators 2 through 5. The group is structured around the category *flow of recall*, which is defined by an amalgam of categories found in different phases of recall. Namely, *flow of recall* refers to a recall where a participant is deeply focused on the task (*immersion*) and the words typically arise one after another, automatically, quickly, and fluidly, without a clear division between them (*codependent emergence*). Importantly, before one concept emerges fully, there is already a sense that more concepts are readily available (*more words are accessible*) with a strong tendency to move on to the recall of another concept (*tendency to rush*).

« 61 » Indicator 2 describes those recall instances where the participants' reports additionally displayed a lack of connection between the emergent concept and the corresponding memory of the experience from the encoding phase (*disconnected concept*). For example, in session one, participant D-S02 reported experiencing the emerged word *needle* [*igla*; critical lure] as a mere part of the recall process without any additional modality (e.g., inner image or inner speech).

« 62 » Indicator 2.1 describes those recall instances, covered by indicator 2, where the participants additionally reported not relying very much on the recollection of experience from the encoding phase while recalling words (*disconnection with memory trace*). In one of these instances, the participant also reported thinking about the concept that was the critical lure already in the encoding phase without acknowledging it as clearly dubious (*incomplete critical lure recognition*). More concretely, in the first session, participant D-S08 reported forming a mental image of the words being presented during the encoding phase and trying to include the word *needle* [*igla*; critical lure] into it but with limited success ("it was as if the word lingered somewhere in the air"). In the recall phase, the word spontaneously appeared through a process that did not depend on the reconstruction of the mental image from before. The word itself lacked an independent existence in the participant's

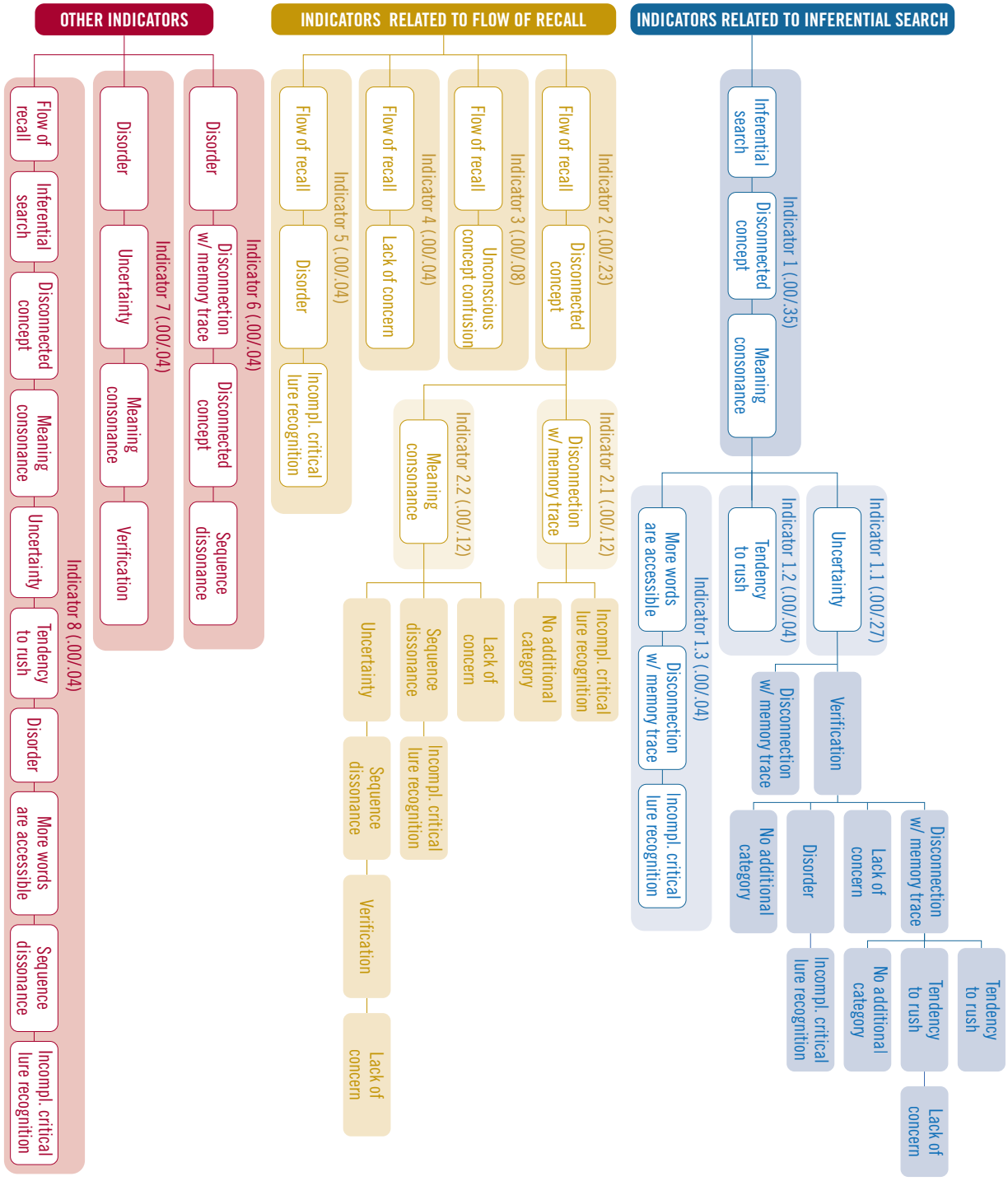

Figure 5 • Phenomenological indicators of inaccurate recall. A decimal point (e.g., Indicator 1.1) designates a second-order indicator, which includes the categories from the first-order indicator and additional ones. The rates of occurrence of indicators in both accurate and inaccurate recall, respectively, are provided in parentheses. The configuration of first- and second-order indicator's categories is displayed using white-background rectangles, while colored-background rectangles show additional categories that occurred alongside the first- and second-order indicator's categories less frequently. The "No additional category" rectangle means that the configuration of categories to its left sometimes appeared without any additional distinctive category. Apart from that, the order of the categories holds no significance.

experience; rather, it came together with a promise that more words would follow, which was accompanied by slight excitement.

« 63 » Indicator 2.2 describes those recall instances, covered by indicator 2, where the participants reported experiencing the emerged words as consistent with the general meaning of the set of words being currently recalled (*meaning consonance*). The reports of these instances additionally reflected the presence of different configurations of the categories *uncertainty*, *sequence dissonance*, *verification*, *lack of concern*, and *incomplete critical lure recognition*. For example, in session two, participant D-S02 reported repeating the word *girl* [*dekle*; critical lure] in their mind after its presentation in the encoding phase and later recalling it as part of the swift, automatic, and fluid process. Fully engrossed in this process, the participant knew that many words were accessible and that they would emerge by themselves. D-S02 was more concentrated on the words that had not yet emerged than on the word that was currently emerging, experiencing the latter with less clarity and individuality. When the word *dekle* emerged, D-S02 experienced it as semantically similar to some other words from the list, but also felt that it arrived too early compared to their memory of the order in which the words were presented.

« 64 » Finally, indicators 3 through 5 are three distinct types of *flow of recall* indicators that did not include the category *disconnected concept*. Each applies to a single inaccurate recall instance. Indicator 3 describes an instance where the participant's report reflected the category *unconscious concept confusion*; indicator 4 describes an instance where the participant's report contained a description corresponding to the category *lack of concern*; and indicator 5 describes an instance where the participant reported feeling generally confused during recall (*disorder*) and experiencing the concept that was also the critical lure already in the encoding phase (*incomplete critical lure recognition*). As an example, in session two, participant D-S07 visualized a *window* [*okno*; critical lure] in the encoding phase and even acknowledged it as a concept that should not be recalled. However, during the subsequent recall phase, overwhelmed by the numerous events happening simultaneously, D-S07 completely forgot this acknowledgment and recalled the concept.

Participant	Number of interviews	Item 1: Memory of the experience	Item 2: Consistency with the experience	Item 3: Completeness of the report	Item 4: Constructing/ inventing things	Item 5: Being understood by the interviewer
D-S01	5	5.0	5.8	5.4	2.2	5.4
D-S02	5	5.4	5.8	5.2	2.0	5.8
D-S03	3	5.7	5.3	5.0	1.7	6.3
D-S04	5	5.4	5.2	6.0	2.2	6.0
D-S05	2	4.5	5.0	6.0	2.5	5.0
D-S06	4	6.0	6.0	4.0	1.0	6.7
D-S07	5	5.0	5.2	4.8	3.4	6.2
D-S08	4	5.5	6.0	5.0	1.5	5.3
D-S09	5	4.4	4.8	5.6	3.6	6.2
D-S10	4	4.3	4.8	4.0	2.0	4.5
D-S11	5	4.8	5.2	4.8	1.8	6.0
M		5.1	5.4	5.1	2.2	5.8
SD		0.9	0.8	1.1	1.0	0.8

Table 3 • Average results of the post-interview questionnaire for each participant for each of the questionnaire items. *Note*: Answers were accepted on a scale from one (strongly disagree) to seven (strongly agree). Number of interviews = number of interviews conducted with each participant.

« 65 » The above-listed indicators (together with *other* indicators of inaccurate recall, described in *Supplementary Material D*) explain 85% (22 out of 26) of instances of inaccurate recall. For the remaining 15% (4 out of 26), we were unable to find any distinguishing criteria at all.

Researcher's accuracy estimate

« 66 » We successfully predicted the accuracy of the target word in 42 out of 47 cases (89% success rate). However, we were partly uncertain in all but one of our estimations (*sleep* [*spanje*; critical lure]) with an average certainty of 87%. Broadly speaking, the reasons stemmed from either our background knowledge of the DRM paradigm (e.g., knowing that the words from each list have a common theme), the content of the participant's report (e.g., the level of recall fluency reported by the participant), or the way that report was delivered during the interview (e.g., the certainty level implicitly or explicitly expressed by the participant while delivering the report).

Results of the quantitative analysis

« 67 » In terms of quantitative data, we first present the results of the post-interview questionnaire and then continue with presenting the results of the independent RT analysis.

Post-interview questionnaire

« 68 » Table 3 shows the results of the post-interview questionnaire averaged for each participant. Overall, not a single result was in the range that would indicate a lower validity of the interviews from the perspective of the participant (i.e., below 4.0 on items 1, 2, 3, and 5, or above it on item 4), suggesting that the participants generally judged their interviews to be valid. However, the results of the individual sessions (see *Supplementary Material E*) revealed that some participants were less confident in the reports given in some of their sessions (particularly the third session with D-S07, and the fourth session with D-S09 and D-S10).

Reaction times

« 69 » We used 42 participants' sessions in the analysis. As can be seen in Table 4,

	N	M	SD	Range	Skewness	Kurtosis
List items (target)	18	7030.9	7622.2	18675.9	0.83	–1.31
List items (all)	1331	4972.7	1957.1	6945.6	0.99	0.24
Critical lures (all/target)	24	3639.0	2561.3	8822.3	1.09	0.19
Unrelated lures (all)	23	14354.6	21409.7	64239.5	1.71	1.34

Table 4 • Descriptive statistics for RTs of list items, critical lures, and unrelated lures.
Note. The numbers of list items and critical lures selected for the interview (target words) are lower because RT data for two critical lures and one list item was lost due to technical issues. The statistics are based on averaged data for each participant.

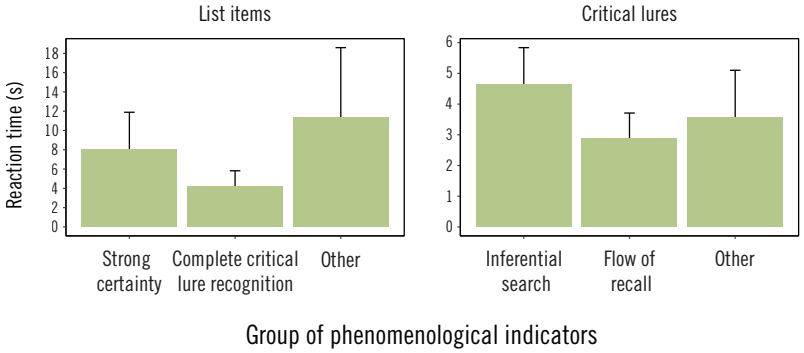

Figure 6 • Mean RTs for each group of phenomenological indicators of accurate and inaccurate recall. Error bars depict the standard error of the mean.

considering all recalled words, the participants recalled critical lures slightly faster than list items, which they recalled faster than unrelated lures. However, these differences were not statistically significant according to a bootstrap-based repeated measures ANOVA for trimmed means (rmanovab), $F_t = 3.21$, $F_{crit} = 9.27$, $p > .05$.

« 70 » Considering target words only, the participants similarly recalled critical lures faster than list items, with a robust paired Yuen's test (yuend) showing that the difference was not significant, $M_{diff} = 2049.57$, 95% CI [–7584.75, 11683.88], $Y_t(4) = 0.59$, $p = .587$, $\delta_t = 0.26$.

Results of the preliminary joint analysis of phenomenological and behavioral data

« 71 » In Figure 6, we can see that recall instances linked to various phenomenological indicators, especially the most coherent ones, exhibited different speeds, suggesting potential alignment between phenomenological and behavioral classifications. Considering accurate recall, instances of *complete critical lure recognition* indicators ($N = 5$; $M = 4201.4$; $SD = 3625.0$) were the fastest, followed by slower instances of *strong certainty* indicators (excluding instances of *complete critical lure recognition*; $N = 8$; $M = 8026.5$; $SD = 10930.3$), and even slower instances of *other* indicators ($N = 3$; $M = 11400.1$; $SD = 12456.7$). Considering inaccurate recall, instances of *flow of recall* indicators ($N = 10$; $M = 2870.8$; $SD = 2649.6$) were the fastest, followed by slightly slower instances of *other* indicators ($N = 3$; $M = 3579.3$; $SD = 2636.4$), and even slower instances of *inferential search* indicators ($N = 8$; $M = 4640.9$; $SD = 3387.8$).

« 72 » The analysis of average serial position of recalled words linked to main indicators showed a negligible difference for the indicators related to *strong certainty* ($M = 7.88$; $SD = 2.17$) and *complete critical lure recognition* ($M = 6.20$; $SD = 3.03$), but a more noticeable difference for indicators related to *flow of recall* ($M = 2.90$; $SD = 1.97$) and *inferential search* ($M = 7.13$; $SD = 3.68$), providing further support for inaccurate recall indicators. While the former occurred mainly at the beginning of the recall, the latter appeared more frequently in the later stages of the recall.

Discussion

« 73 » The present study aimed to explore the phenomenological and behavioral characteristics of accurate and inaccurate recall in the DRM paradigm. We interrupted participants while they were recalling the words from the previously presented list and conducted in-depth phenomenological interviews about their recall processes. Throughout the task, we measured the RTs needed to recall each word and performance accuracy. We conducted a separate analysis of phenomenological and behavioral data, as well as a preliminary joint analysis of both.

« 74 » The qualitative analysis revealed new insights into the complexity of the phenomenology of recall as induced by the DRM paradigm, which can be organized into five overarching categories:

- different states, gestures, and directions of attention;
- phases of recall;
- experiences that can be found throughout the recall or in a particular phase;
- mood and motivation; and
- other relevant observations.

« 75 » These categories gave rise to the phenomenological indicators of accurate and inaccurate recall, which is the main contribution of the present article. The most coherent accurate recall indicators were related to the experience of great assurance in the recalled word (*strong certainty*) or the emergence of a word matching the critical lure during encoding as a word that should be avoided during recall (*complete critical lure recognition*). In contrast, the most coherent inaccurate recall indicators were linked to a deliberate, inferential type of recall (*inferential search*), or an automatic and fluid recall (*flow of recall*). Preliminary joint analysis of phenomenological and behavioral data indicated mild support for the inaccurate recall indicators. This was evident in varying RTs associated with different indicators and noticeable differences in

the serial position of recalled words associated with these indicators.

« 76 » In the remainder of this section, we try to integrate our main findings into prior research, evaluate the reliability and validity of our approach, and suggest future research directions.

(In)consistency with prior research on memory accuracy

« 77 » Our observations align with some general findings from the research on inaccurate memory. First, the distribution of *disconnected concept*, *weakly connected concept*, and *strongly connected concept* in accurate and inaccurate recall, as well as the predominance of *inferential search* in inaccurate recall (38% as opposed to 5% in accurate recall) support the idea that inaccurate memories are less specific and include more cognitive processes and fewer perceptual details than accurate memories (Mather, Henkel & Johnson 1997; Slotnick & Schacter 2004). Second, the finding that *strong certainty* is the most prevalent category in accurate recall and the key characteristic of one of the groups of accurate recall indicators agrees with the common finding that accurate memory is rated as higher in confidence than inaccurate memory (for a review, see Gallo 2006). Third, we showed a similar frequency of recalled list items, critical lures, and unrelated lures to the one that is usually found in the DRM studies: the participants recalled the list items the most, fewer critical lures, and even fewer unrelated lures (e.g., Toglia, Neuschatz & Goodwin 1999; Anastasi, Rhodes & Burns 2000; Iacullo & Marucci 2016).

« 78 » Some of our observations are inconsistent with prior research on memory accuracy. Our finding that the experience of the recalled word can be separated into the categories *disconnected concept*, *weakly connected concept*, and *strongly connected concept* according to the degree of its connectedness with a memory trace challenges the core assumption of the remember/know paradigm (explained in Footnote 2). As *disconnected concept* was most frequently associated with inaccurate memory (it occurred in 65% of inaccurate recall and 11% of accurate recall), we would expect it to match the "know" response and feeling of familiarity. Instead, it was never described by using words, such as "familiarity"; this description was captured by *weakly connected concept*, which was more typical of accurate than inaccurate recall (it occurred in 37% of accurate recall and 12% of inaccurate recall). This discrepancy questions the remember/know paradigm's validity, suggesting that participants may select "remember" or "know" responses to adhere to study instructions, even when their experiences do not align. To capture recall's intricate phenomenology, researchers should adopt more complex data-collection methods.

« 79 » Contrary to previous research (Jou 2008), our study found no significant difference in recall RTs between list items and critical lures. While this may be caused by our small sample size of only 24 instances of critical lures, larger studies in the field of memory accuracy also yield inconclusive results regarding the existence of RT differences (Jou 2008). This suggests that inconsistency likely arose from variations in research design, including differences in stimuli, participants, and environmental factors (Brainerd & Reyna 2005; Jou & Flores 2013).

Multifaceted support for the distinction between inaccurate recall indicators related to *inferential search* and *flow of recall*

« 80 » As shown below, indicators related to *inferential search* and *flow of recall* not only garnered notable support in our study but also align well with prior research. First, among the 20 recall instances that we categorized as *inferential search* or *flow of recall*, only one displayed features of both, while the rest demonstrated features of one or the other.

« 81 » Second, *inferential search* indicators were more frequently associated with doubt, slower RTs,[11] and the later stages of the recall, while *flow of recall* indicators were more often associated with a lack of uncertainty, faster RTs, and the beginning of the recall.

« 82 » Third, this distinction also agrees with prior research, showing that critical lures are recalled in either an early or a later stage, where early-stage critical lures are constructed in the encoding phase, characterized by faster RTs, and experienced as indistinguishable from accurate recall, while later-stage critical lures are constructed in the recall phase, characterized by slower RT, and experienced as different from accurate recall (Jou 2008).[12] This fits the differences in RTs and serial position of both types of indicators, as well as the observation that *inferential search* indicators were less frequently linked to specific events during encoding (e.g., thinking about the critical lure or experiencing disturbances from the environment) compared to *flow of recall* indicators (33% vs 50%, respectively).

« 83 » Finally, the distinction also aligns well with dual-process theories of higher-level cognition, such as there being a dichotomy between System 1 and System 2 (Evans & Stanovich 2013; Kahneman 2003). Like *flow of recall*, System 1 is typically described as fast, intuitive, automatic, and effortless, and like *inferential search*, System 2 is considered to be slow, effortful, deliberative, and rule-governed (Kahneman 2003; Morewedge & Kahneman 2010). System 2 is also responsible for meta-cognitive evaluation (in other words: "doubt is a phenomenon of System 2"; Kahneman 2003: 8), which fits our observation that *flow of recall* was almost never experienced as uncertain, whereas *inferential search* almost always included a bit of uncertainty. Additionally, System 1 was connected to automatic operations of associative memory, such as spontaneous activation of related and easily accessible memories when the target memory is unavailable (also called attribute substitution; Morewedge & Kahneman 2010). Given that critical lures are strongly related to list items and processed with a similar degree of fluency (Karpicke, McCabe & Roediger 2008), it is conceivable that System 1 might retrieve a memory of one of these critical lures. This retrieved memory can then be accepted as accurate, either because the meta-cognitive System 2 is dormant (as in *flow of recall*) or because it fails at evaluation or ignores possible alternatives (as in inferential recall; Evans 2007).

11| It is important to note that the pattern observed may be a mere coincidence rather than a genuine trend, given the small sample size and high variability.

12| The terms "early stage" and "later stage" refer to different stages of the whole recall part of the task, not to different stages within a single recall instance.

« 84 » The multifaceted support suggests the suitability of *inferential search* and *flow of recall* indicators as the basis for developing operationalized measures within research that employs "thin" phenomenological methods (Berkovich-Ohana et al. 2020).

Why is *flow of recall* associated with inaccurate recall?

« 85 » Our finding that *flow of recall* is integral to one group of inaccurate recall indicators also causes some tension with prior research: two of the basic characteristics of *flow of recall* are fluency and immersion (focus), which are typically associated with performance accuracy, not inaccuracy (e.g., Whittlesea & Leboe 2000; Wulf 2013).

« 86 » In terms of fluency, cognitive psychologists Bruce Whittlesea and Jason Leboe summarized the prevailing position clearly: "In remembering, thoughts that one has had before are usually more fluently processed now than thoughts that one has not. Similarly, accurate identifications will usually be performed more smoothly than erroneous ones" (Whittlesea & Leboe 2000: 100). However, research has also shown that relying on recall fluency can lead to inaccurate performance (e.g., Karpicke, McCabe & Roediger 2008). More importantly, fluency has been suggested to play an important role in the DRM effect (Roediger & McDermott 1995). According to the activation-monitoring theory (McDermott & Watson 2001), as critical lures are strongly associated with the list items, they are repeatedly activated in memory when list items are presented. As a result, during a subsequent memory test, critical lures are experienced much like list items, leading to an inability to differentiate between the two. Although this may be the case with *flow of recall*, the evidence for the above explanation comes primarily from the studies at the cognitive level of analysis, rather than the phenomenological level. As *flow of recall* is an intricate phenomenological category, assessing the consistency is challenging.

« 87 » Recently, however, a viewpoint, comparable to Whittlese and Leboe's, has emerged from the debate on how to assess the authenticity of retrospectively collected first-person reports (Bitbol & Petitmengin 2013; Kordeš & Demšar 2018; Sparby, Edelhäuser & Weger 2020). Based on the examination of the process of recall that contributed to the report, two cues of authenticity were recently proposed: (a) "spontaneous emergence of experiential elements" (Bitbol & Petitmengin 2013: 276) and (b) "activity with response," described as "a form of flow […], where one feels like one is in continued resonance with the experience […]" (Sparby, Edelhäuser & Weger 2020: 5). As both cues resemble *flow of recall* and are indicative of accurate recall, while we considered *flow of recall* as being central to inaccurate recall, this suggests that the spontaneity of the recall process is either not generalizable in terms of context, or not a reliable indicator of accurate or inaccurate recall.

« 88 » In terms of focus, the tension with prior research relates to the apparent contradiction within the constellation of *flow of recall*'s characteristics: how could it be that a deep focus (*immersion*) is associated with a strong tendency to move on with the task (*tendency to rush*) and being caught in-between two different concept emergences (*codependent emergence*)?[13] *Immersion* implies attending to a single task, which has been shown to facilitate retrieval functions (e.g., Gazzaley & Nobre 2012), whereas *tendency to rush* and *codependent emergence* imply scattered attention, which has been shown to do the contrary (e.g., Jacoby, Woloshyn & Kelley 1989). Although we were able to temporally situate *codependent emergence* in the phase of *concept emergence*, we have not been able to do so for *tendency to rush* and *immersion*, as participants were speaking about them as though they were present throughout the whole recall process with varying intensity. The problem is partly related to the phenomenon under investigation – recall within the DRM paradigm happens so quickly that it was extremely challenging for participants to disentangle the subtleties of its temporal progression. This proved to be particularly relevant for *flow of recall*, as, somewhat ironically, one of the problems was precisely that the participants were deeply immersed in the task, with experiential content unfolding automatically and quickly. Nevertheless, future studies should further unpack its temporal progression to see if these characteristics coincide or not.

13| We thank an anonymous reviewer for pointing this out.

« 89 » Another potential cause for this apparent contradiction is that concepts, like focus, immersion, and attention, do not have a universally accepted definition (e.g., Eysenck & Keane 2015) and may have been used by the participants differently. Two types of attention are especially relevant for this point: (a) selective attention, defined as attention to a single stimulus; and (b) divided attention, defined as attending to multiple tasks simultaneously (Eysenck & Keane 2015). While selective attention maps quite nicely to our category *immersion*, the mapping of divided attention is less straightforward. It could be mapped to the category *disorder*, but also to the categories *tendency to rush* and *codependent emergence*, despite apparent differences between them. Namely, the former assumes an almost complete lack of selective attention, while the latter two do not; although they demonstrate a micro-level dispersion, they still occur in a very narrow space of task processing. Hence, further research on attentional micro-structure is required to settle this debate.

« 90 » One possible way to navigate the tension related to *flow of recall* is to understand fluency and focus as being a part of a complex interdependent network of factors present in the recall situation as opposed to treating them in a vacuum. *Flow of recall* is a highly abstract category that encompasses not only immersion and fluency, but also characteristics, such as automaticity, rapid unfolding, and melding of different concept emergences. While some of these characteristics, in other constellations, can also be integral to accuracy,[14] our data suggests that this specific constellation of characteristics and their relationships are more indicative of inaccuracy.

14| Indeed, the participants in our study sometimes reported experiencing spontaneous emergence of concepts when they did not experience other characteristics of *flow of recall*, showing that its characteristic may also arise independently or in a different constellation. It is important to note that *flow of recall* (as a category and not as a group of indicators) by itself was not unique to inaccurate recall. It was once followed by an accurate recall, and some participants' reports suggested they experienced something like *flow of recall* when recalling words that *were* presented but *were not* selected as target words.

JAŠA ČERNE
is a cognitive scientist with a passion for phenomenologically grounded interdisciplinary research of the human mind. He holds a BSc in Computer Science and an MSc from the Middle European interdisciplinary master's programme in Cognitive Science. Presently, he serves as an assistant lecturer and researcher at the Center for Cognitive Science, University of Ljubljana. His research focuses on the neurophenomenological investigations of consciousness, psychopathology, psychedelics, contemplative practices, and social interactions.

URBAN KORDEŠ
is a professor of cognitive science and first-person research at the University of Ljubljana, where he is currently heading the cognitive science program. He holds a bachelor's degree in mathematical physics and a doctorate in philosophy of cognitive science. His research interests include in-depth empirical phenomenological research, neurophenomenology, second-order cybernetics, collaborative knowledge creation, as well as epistemic and methodological issues in the research of non-trivial systems. Urban believes that training in the skill of introspection and subsequent first-person reporting should become one of the essential cognitive science research techniques.

Demand characteristics

« 91 » Our study underscores the role of demand characteristics (Corneille & Lush 2022; Orne 1962; see also Kordeš & Demšar 2023 for an exploration of this phenomenon in the context of phenomenological studies) within the domain of memory accuracy research. We captured one apparent case of demand effects in the category *lack of concern*. Namely, in five instances of inaccurate recall and never in accurate recall, the participants approached the task with indifference or disinterest, which was at least partly due to their interpretation of the research context and not the instructions themselves. One possible explanation is the absence of incentives, which may have influenced the participant's (un)willingness to dedicate the required cognitive resources to the task. Such an indifferent attitude agrees with the idea of a "reactant subject," which refers to a participant who understands the instructions but does not comply with them willingly (Corneille & Lush 2022). Furthermore, since the participants conducted the study without the presence of the researcher, they might have acted differently from how they typically would (Černe & Kordeš 2023). Although previous research has examined demand characteristics within the DRM paradigm (Lampinen, Neuschatz & Payne 1999), thus far, the effects of participant attitudes and the researcher's presence/absence remain unexplored (see Oblak et al. 2022 for an attempt to address this gap in a similar context).

Everyday implications

« 92 » Our findings show that proposed indicators explain as much as 87% (39 out of 45) of all recall instances that we investigated in phenomenological interviews. This resembles a conclusion that was reached in a recent choice blindness study by Petitmengin and colleagues (2013). Put simply, participants were shown pairs of different faces and asked which they preferred. Unbeknownst to them, the pictures were switched, and participants were questioned about the reasons for their choices. Those who underwent phenomenological interviews before being questioned detected the manipulation in 80% of cases, while those who did not only noticed it in 33% of cases. Admittedly, unlike Petitmengin et al. (2013), in our study, we interviewed participants after they had already made a memory report, and we did not rely on external manipulation to induce inaccurate recall. Nevertheless, both studies seem to suggest that understanding and critically evaluating the phenomenological structure of a given cognitive process (e.g., recall or decision-making) can improve the identification of inaccurate judgments about corresponding mental states (e.g., memories or decisions), whether from a first-person perspective (as in Petitmengin et al. 2013) or a third-person point of view (as demonstrated in our study).

« 93 » This also raises the question of whether these effects can apply to everyday life of the general population. Our study investigated artificially induced memories in an online laboratory, so the extent of generalizability to everyday remembering is questionable. However, as many inaccurate memories rely on gist processes (Brainerd & Reyna 2002), which is precisely the focus of the DRM paradigm, our study's results should not be dismissed as ecologically invalid (cf. Gallo 2010). More importantly, finding a skilled empirical phenomenology practitioner willing to explore one's private memory processes in a series of in-depth phenomenological interviews can be challenging. Yet, Terje Sparby (2023) has recently introduced a method known as micro-

https://constructivist.info/19/1/038.cerne

phenomenological self-inquiry, which seeks to diminish the need for a skilled researcher and empowers individuals to undertake such exploration independently.

Constructivist perspective on memory

« 94 » With neurophenomenology, Varela aimed to establish an enactivist framework for empirical research that takes the middle ground between the realist belief that science is investigating the external world from the impersonal perspective and the idealist position that everything is a mere projection of our minds (Vörös, Froese & Riegler 2016). As a potential solution, he proposed seeking areas where "phenomenological accounts of the structure of experience and their counterparts in cognitive science relate to each other through reciprocal constraints" (Varela 1996: 343). Our study contributes to this endeavor within memory accuracy research by offering detailed structural accounts of both accurate and inaccurate recall experiences while pointing to potential intersections with behavioral data.

« 95 » Our study also provides compelling support for the constructivist perspective on memory (e.g., Bartlett 1920; Riegler 2005; Schacter 2012), which rejects the notion of memory as a passive retrieval of objective information and underscores its dynamic nature that is intertwined with other cognitive processes. The central feature of *inferential search*, one of the main phenomenological anchors for inaccurate recall indicators, is making inferences or guesswork, which points to the inseparability between memory and decision-making processes. More importantly, out of 22 distinctive categories used to develop phenomenological indicators, 16 were found in instances of both accurate and inaccurate recall, while the remaining ones would likely emerge in both with a longer study duration. This shows that the phenomenological landscapes associated with both accurate and inaccurate recall exhibit remarkable similarities and suggests that all memories rely on a similar underlying mechanism that is not meant to re-instantiate our past experiences perfectly (cf. Riegler 2005; Schacter 2012).

« 96 » Moreover, our phenomenological investigation showed that not a single characteristic that we observed was independent of the complex network of other characteristics of recall, participants' attitudes toward the task, moods, beliefs, expectations, current context, etc. It also showed that the relationships between these characteristics are crucial for making sound conclusions about memory accuracy. This resonates with the constructivist position, which emphasizes that a typical psychological way of conceiving cognitive faculties as independent entities in a sterile environment (such as conceptualizing memory in terms of processes that encode, store, and retrieve information; Eysenck & Keane 2015) is not only an oversimplification but also misleading for scientific pursuits (Riegler 2005).

Problems and limitations

« 97 » Our study did not unfold without challenges. First, the implementation of the double-blind criterion has proved to be problematic, since we were not allowed to come into direct contact with the material while organizing the whole study up until the second part of the analysis. Despite the 89% success rate of our accuracy estimates, the slight uncertainty in most estimations suggests that the blinding was maintained, making the interviews and analysis less biased.

« 98 » Second, while conducting the entire study remotely without the presence of the researcher minimized the impact of the researcher's bias (e.g., Friesen et al. 2020), it also limited our ability to control the environment, which occasionally influenced the task or interview.

« 99 » Third, interpreting the results of the post-interview questionnaire was challenging as they reflected subjective generalizations of a 98-minute experiential episode (cf. Hurlburt 2011). A more specific approach would be needed to accurately map quantitative results to corresponding parts of the experience.

« 100 » Fourth, a relatively small sample size coupled with lax environmental control prevented us from drawing firm statistical conclusions.

« 101 » Finally, a closer look at the remaining 13% of instances that were not explained by our indicators showed that in five out of seven instances, the participants faced difficulties providing detailed accounts of their experience during the interviews. Some of them were uncertain about the object of investigation, some had trouble retrieving the experience, and others were interrupted by external factors during the task. This indicates that distinguishing characteristics might have existed in these instances as well, but their discovery resisted our method of investigation. To overcome these limitations, future studies should consider:

- providing training to participants to improve their skills in observing and reporting on experience (although we recruited participants with experience in first-person research, we did not test their expertise or provide any additional training in systematically examining experience);[15]
- offering thorough assistance in evoking past experiences through a dialogical setting; and
- conducting a study in a controlled environment without external disturbances.

Conclusion

« 102 » In summary, by adopting a neurophenomenological research design, we uncovered new insights about accurate and inaccurate recall in the DRM paradigm that support, extend, and challenge commonly accepted research approaches and findings in the realm of memory accuracy research. We identified phenomenological indicators of accurate and inaccurate recall, which have the potential to inform our understanding of everyday cases of remembering. If our proposed indicators withstand further scrutiny, they hold the promise of paving the way for more robust investigations across neuroscientific, practical, and clinical domains. Although our study does not settle the debate on whether and how accurate and inaccurate memories differ, it suggests that better understanding and critically evaluating one's recall experience can improve the identification and potentially prevent the undesirable consequences of inaccurate memories.

15| One way of conducting such training is captured in the descriptive experience sampling method, introduced by Russell Hurlburt (2011).

Acknowledgments

We extend our gratitude to Florian Klauser and James Sanders for their insightful comments and suggestions on the earlier drafts of the article. We also thank the participants for their unwavering commitment and cooperation throughout the study.

Funding

The authors declare that they did not receive any funding for the study.

Competing interests

The authors declare that they have no competing interests.

References

Algina J., Keselman H. J. & Penfield R. D. (2005) An alternative to Cohen's standardized mean difference effect size: A robust parameter and confidence interval in the two independent groups case. Psychological Methods 10(3): 317–328.

Anastasi J. S., Rhodes M. G. & Burns M. C. (2000) Distinguishing between memory illusions and actual memories using phenomenological measurements and explicit warnings. American Journal of Psychology 113(1): 1–26.

Berkovich-Ohana A., Dor-Ziderman Y., Trautwein F. M., Schweitzer Y., Nave O., Fulder S. & Ataria Y. (2020) The hitchhiker's guide to neurophenomenology – the case of studying self boundaries with meditators. Frontiers in Psychology 11: 1680. ▸ https://cepa.info/6666

Bartlett F. C. (1920) Some experiments on the reproduction of folk-stories. Folklore 31(1): 30–47. https://doi.org/10.1080/001558 7X.1920.9719123

Bernstein D. M. & Loftus E. F. (2009) How to tell if a particular memory is true or false. Perspectives on Psychological Science 4(4): 370–374.

Bitbol M. & Petitmengin C. (2013) A defense of introspection from within. Constructivist Foundations 8(3): 269–279. ▸ https://constructivist.info/8/3/269

Brainerd C. J. & Reyna V. F. (2002) Fuzzy-trace theory and false memory. Current Directions in Psychological Science 11(5): 164–169.

Brainerd C. J. & Reyna V. F. (2005) The science of false memory. Oxford University Press, Oxford.

Černe J. & Kordeš U. (2023) Slovenian version of the Deese–Roediger–Mcdermott paradigm: A normative and exploratory study. Psihološka Obzorja 32: 32–45. https://doi.org/10.20419/2023.32.572

Charmaz K. (2014) Constructing grounded theory. Second edition. SAGE, Los Angeles.

Corbin J. & Strauss A. (2008) Basics of qualitative research: Techniques and procedures for developing grounded theory. Third edition. SAGE, Los Angeles.

Corneille O. & Lush P. (2022) Sixty years after Orne's American Psychologist article: A conceptual framework for subjective experiences elicited by demand characteristics. Personality and Social Psychology Review 27(1): 83–101.

Deese J. (1959) On the prediction of occurrence of particular verbal intrusions in immediate recall. Journal of Experimental Psychology 58(1): 17–22.

Depraz N., Varela F. J. & Vermersch P. (2003) On becoming aware: A pragmatics of experiencing. John Benjamins, Amsterdam.

Evans J. S. B. T. (2007) Hypothetical thinking: Dual processes in reasoning and judgement. Psychology Press, London.

Evans J. S. B. T. & Stanovich K. E. (2013) Dual-process theories of higher cognition: Advancing the debate. Perspectives on Psychological Science 8(3): 223–241.

Eysenck M. W. & Keane M. T. (2015) Cognitive psychology: A student's handbook (7th ed.). Psychology Press. https://doi.org/10.4324/9781315778006

Field A. P. & Wilcox R. R. (2017) Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers. Behaviour Research and Therapy 98: 19–38.

Friesen K. B., Zhang Z., Monaghan P. G., Oliver G. D. & Roper J. A. (2020) All eyes on you: How researcher presence changes the way you walk. Scientific Reports 10(1): 1–8. https://doi.org/10.1038/s41598-020-73734-5

Gallo D. A. (2006) Associative illusions of memory: False memory research in DRM and related tasks. Psychology Press, New York.

Gallo D. A. (2010) False memories and fantastic beliefs: 15 years of the DRM illusion. Memory and Cognition 38(7): 833–848. https://doi.org/10.3758/MC.38.7.833

Gazzaley A. & Nobre A. C. (2012) Top-down modulation: Bridging selective attention and working memory. Trends in Cognitive Sciences 16(2): 129–135. https://doi.org/10.1016/j.tics.2011.11.014

Glaser B. & Strauss A. (1967) The discovery of grounded theory: Strategies for qualitative research. Aldine Transaction, New Brunswick.

Heaps C. M. & Nash M. (2001) Comparing recollective experience in true and false autobiographical memories. Journal of Experimental Psychology: Learning Memory and Cognition 27(4): 920–930.

Hurlburt R. T. (2011) Investigating pristine inner experience: Moments of truth. Cambridge University Press, Cambridge MA.

Iacullo V. M. & Marucci F. S. (2016) Normative data for Italian Deese/Roediger–McDermott lists. Behavior Research Methods 48(1): 381–389. https://doi.org/10.3758/s13428-015-0582-3

Jacoby L. L., Woloshyn V. & Kelley C. (1989) Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. Journal of Experimental Psychology: General 118(2): 115–125. https://doi.org/10.1037//0096-3445.118.2.115

Jacoby L. L. & Whitehouse K. (1989) An illusion of memory: False recognition influenced by unconscious perception. Journal of Experimental Psychology: General 118(2): 126–135.

Jou J. (2008) Recall latencies, confidence, and output positions of true and false memories: Implications for recall and metamemory theories. Journal of Memory and Language 58(4): 1049–1064.

Jou J. & Flores S. (2013) How are false memories distinguishable from true memories in the Deese-Roediger-McDermott paradigm? A review of the findings. Psychological Research 77(6): 671–686.

Kahneman D. (2003) A perspective on judgment and choice: Mapping bounded rationality. American Psychologist 58(9): 697–720.

Karpicke J. D., McCabe D. P. & Roediger H. L. (2008) False memories are not surprising: The subjective experience of an associative

memory illusion. Journal of Memory and Language 58: 1065–1079.

Kim H. & Cabeza R. (2007) Trusting our memories: Dissociating the neural correlates of confidence in veridical versus illusory memories. Journal of Neuroscience 27(45): 12190–12197. https://doi.org/10.1523/JNEUROSCI.3408-07.2007

Kordeš U. (2016) Going beyond theory: Constructivism and empirical phenomenology. Constructivist Foundations 11(2): 375–385. ▸ https://constructivist.info/11/2/375

Kordeš U. & Demšar E. (2018) Excavating belief about past experience: Experiential dynamics of the reflective act. Constructivist Foundations 13(2): 219–229. ▸ https://constructivist.info/13/2/219

Kordeš U. & Demšar E. (2021) Being there when it happens: A novel approach to sampling reflectively observed experience. New Ideas in Psychology 60: 100821. ▸ https://cepa.info/6940

Kordeš U. & Demšar E. (2023) Horizons of becoming aware: Constructing a pragmatic-epistemological framework for empirical first-person research. Phenomenology and the Cognitive Sciences 22: 339–367. ▸ https://cepa.info/7367

Kordeš U. & Klauser F. (2016) Second-person in-depth phenomenological inquiry as an approach for studying enaction of beliefs. Interdisciplinary Description of Complex Systems 14(4): 369–377. ▸ https://cepa.info/4334

Koriat A. (1993) How do we know that we know? The accessibility model of the feeling of knowing. Psychological Review 100(4): 609–639.

Lampinen J. M., Neuschatz J. S. & Payne D. G. (1997) Memory illusions and consciousness: Examining the phenomenology of true and false memories. Current Psychology 16(3–4): 181–224.

Lampinen J. M., Neuschatz J. S. & Payne D. G. (1999) Source attributions and false memories: A test of the demand characteristics account. Psychonomic Bulletin and Review 6(1): 130–135. https://doi.org/10.3758/BF03210820

Loftus E. F. (1996) Memory distortion and false memory creation. Bulletin of the American Academy of Psychiatry and the Law 24(3): 281–295.

Lutz A. & Thompson E. (2003) Neurophenomenology: Integrating subjective experience and brain dynamics in the neuroscience of consciousness. Journal of Consciousness Studies 10(9): 31–52. ▸ https://cepa.info/2363

Mair P. & Wilcox R. (2020) Robust statistical methods in R using the WRS2 package. Behavior Research Methods 52(2): 464–488. https://doi.org/10.3758/s13428-019-01246-w

Marche T. A., Brainerd C. J. & Reyna V. F. (2010) Distinguishing true from false memories in forensic contexts: Can phenomenology tell us what is real? Applied Cognitive Psychology 24(2): 1168–1182.

Mather M., Henkel L. A. & Johnson M. K. (1997) Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaire compared. Memory & Cognition 25(6): 826–837.

McDermott K. B. & Watson J. M. (2001) The rise and fall of false recall: The impact of presentation duration. Journal of Memory and Language 45(1): 160–176. https://doi.org/10.1006/jmla.2000.2771

Morewedge C. K. & Kahneman D. (2010) Associative processes in intuitive judgment. Trends in Cognitive Sciences 14(10): 435–440.

Nakamura J. & Csikszentmihalyi M. (2002) The concept of flow. In Snyder C. R. & Lopez S. J. (eds.) Handbook of positive psychology. Oxford University Press: 89–105.

Oblak A., Slana Ozimič A., Repovš G. & Kordeš U. (2022) What individuals experience during visuo-spatial working memory task performance: An exploratory phenomenological study. Frontiers in Psychology 13(811712): 1–19. https://doi.org/10.3389/fpsyg.2022.811712

Orne M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. American Psychologist, 17(11): 776–783.

Payne D. G., Elie C. J., Blackwell J. M. & Neuschatz J. S. (1996) Memory illusions: Recalling, recognizing, and recollecting events that never occurred. Journal of Memory and Language 35(2): 261–285.

Peirce J., Gray J. R., Simpson S., MacAskill M., Höchenberger R., Sogo H., Kastman E. & Lindeløv J. K. (2019) PsychoPy2: Experiments in behavior made easy. Behavior Research Methods 51(1): 195–203. https://doi.org/10.3758/s13428-018-01193-y

Petitmengin C. (2006) Describing one's subjective experience in the second person: An interview method for the science of consciousness. Phenomenology and the Cognitive Sciences 5(3–4): 229–269. ▸ https://cepa.info/2376

Petitmengin C., Remillieux A., Cahour B. & Carter-Thomas S. (2013) A gap in Nisbett and Wilson's findings? A first-person access to our cognitive processes. Consciousness and Cognition 22(2): 654–669. ▸ https://cepa.info/931

Petitmengin C., Remillieux A. & Valenzuela-Moguillansky C. (2019) Discovering the structures of lived experience: Towards a micro-phenomenological analysis method. Phenomenology and the Cognitive Sciences 18: 691–730. ▸ https://cepa.info/6664

Read J. D. (1996) From a passing thought to a false memory in 2 minutes: Confusing real and illusory events. Psychonomic Bulletin & Review 3(1): 105–111.

Robins S. K. (2018) Memory and optogenetic intervention: Separating the engram from the ecphory. Philosophy of Science 85(5): 1078–1089. https://doi.org/10.1086/699692

Roediger H. L. & McDermott K. B. (1995) Creating false memories: Remembering words not presented in lists. Journal of Experimental Psychology: Learning, Memory, and Cognition 21(4): 803–814.

Riegler A. (2005) Constructive memory. Kybernetes 34(1/2): 89–104. ▸ https://cepa.info/1779

Schacter D. L. (2012) Constructive memory: Past and future. Dialogues in Clinical Neuroscience 14(1): 7–18.

Slotnick S. D. & Schacter D. L. (2004) A sensory signature that distinguishes true from false memories. Nature Neuroscience 7(6): 664–672.

Sparby T. (2023) Micro-phenomenological self-inquiry. Phenomenology and the Cognitive Sciences 22(1): 247–266. https://doi.org/10.1007/s11097-022-09815-9

Sparby T., Edelhäuser F. & Weger U. W. (2020) The sense of certainty and the reliability of first person reports: An investigation using micro-phenomenological self-inquiry. Scandinavian Journal of Psychology 61(6): 775–783. ▸ https://cepa.info/8532

Toglia M. P., Neuschatz J. S. & Goodwin K. A. (1999) Recall accuracy and illusory memories: When more is less. Memory 7(2): 233–256.

Tulving E. (1985) Memory and consciousness. Canadian Psychology 26(1): 1–12.
Varela F. J. (1996) Neurophenomenology: A methodological remedy for the hard problem. Journal of Consciousness Studies 3(4): 330–349. ▸ https://cepa.info/1893
Vörös S., Froese T. & Riegler A. (2016) Epistemological odyssey: Introduction to special issue on the diversity of enactivism and neurophenomenology. Constructivist Foundations 11(2): 189–203. ▸ http://constructivist.info/11/2/189
Whittlesea B. W. A. & Leboe J. P. (2000) The heuristic basis of remembering and classification: Fluency, generation, and resemblance. Journal of Experimental Psychology: General 129(1): 84–106. https://doi.org/10.1037/0096-3445.129.1.84
Wulf G. (2013) Attentional focus and motor learning: A review of 15 years. International Review of Sport and Exercise Psychology 6(1): 77–104. https://doi.org/10.1080/1750984X.2012.723728

Received: 7 June 2023
Revised: 18 September 2023
Revised: 28 September 2023
Accepted: 2 October 2023

This target article is part of a bigger picture that encompasses several open peer commentaries and the response to these commentaries. You can read these accompanying texts on the following pages (or, if you have only the target article, follow the embedded link that takes you to the journal's web page from which you can download these texts).

Open Peer Commentaries

on Jaša Černe & Urban Kordeš's "Deconstructing Accurate and Inaccurate Recall in the DRM Paradigm: A Phenomenological and Behavioral Exploration"

Between Naming and Knowing: The Possibility of Critique in Working with Cognitive-Psychological Paradigms of Memory

Davood Gozli
Independent Scholar, Canada
d.ghara/at/gmail.com

> Abstract • The phenomenological-constructivist perspective can offer insights into the hidden assumptions of empirical research methods, including those related to the Deese-Roediger-McDermott (DRM) paradigm, because it makes it possible to detect acts of construction underlying the assertions and the decisions within research paradigms. While attempting to align their efforts with conventional cognitive science, phenomenological researchers might risk neglecting the critical potential of their own perspectives.

Handling Editor • Alexander Riegler

« 1 » In the fifth of his *Logical Investigations*, Edmund Husserl (2001) provides an analysis of explicit and implicit judgments, emphasizing their distinctions as (parts of) meaningful acts. He observes that some acts of judgment are so implicitly embedded in our speech that they function primarily as acts of naming. To illustrate, he notes:

> "If we have found out or seen that the town Halle is on the Saale, or that π is a transcendent number, we may go on to talk of Halle-on-the-Saale or the transcendent number π, but we shall not be judging any longer, or at least we need not be doing so, and such a judgment, should it arise in the side, makes no contribution to our act of nominal reference." (Husserl 2001: 153)

« 2 » Pertaining to our discussion, Husserl's insight can be related to experimental procedures and how they have been deemed valid tools for assessing specific psychological functions. Just as we transition from "π is a transcendent number" to "the transcendent number π," we shift from the assertion, "The Deese-Roediger-McDermott (DRM) paradigm can assess memory," to an act of naming or nominal reference such as "The DRM memory test." In doing so, we are led to overlook the underlying assumptions of the DRM paradigm.

« 3 » The value of a constructivist-phenomenological stance, especially in the context of the present study, lies in its capacity to discern how a paradigm, or an experimental