Received: 11 April 2022 | Revised: 2 January 2023 | Accepted: 16 January 2023

DOI: 10.1002/tea.21851

RESEARCH ARTICLE

JRST | WILEY

Narrating science: Can it benefit science learning, and how? A theoretical review

Sara Soares[1,2] | **Matilde Gonçalves**[3] | **Rita Jerónimo**[1] | **Régine Kolinsky**[2]

[1]Iscte-Instituto Universitário de Lisboa, CIS-Iscte, Lisboa, Portugal

[2]Unité de Recherche en Neurosciences Cognitives (Unescog), Center for Research in Cognition and Neurosciences (CRCN), Université Libre de Bruxelles (ULB), Bruxelles, Belgium

[3]Centro de Linguística da Universidade NOVA de Lisboa, Faculdade de Ciências Sociais e Humanas, Universidade NOVA de Lisboa, Lisboa, Portugal

Correspondence
Sara Soares, Centro de Investigação e de Intervenção Social, Iscte – Instituto Universitário de Lisboa, Av. das Forças Armadas, 1649-026 – Lisboa, Portugal.
Email: sapss@iscte-iul.pt

Funding information
Fundação para a Ciência e a Tecnologia, Grant/Award Numbers: PD/BD/135345/2017, COVID/BD/152481/2022

Abstract

Narrative texts have been advocated as tools to tackle science learning challenges, and there is even the proposal of a "narrative effect" on learning. We believe it is necessary to examine previous evidence on this effect, as well as to characterize the process of learning through science narrative texts more broadly. In this article, we offer a theoretical review drawing on three frameworks, namely on pedagogical aspects of text learning, linguistic features of texts, and cognitive aspects of text comprehension. Based on that, we analyzed two complementary questions. First, we reviewed 36 studies to ask if science narrative texts can benefit learning and memory outcomes at different educational levels (i.e., the "If" question). We found encouraging evidence for the use of science narrative texts at various educational levels, especially in delayed assessments and longer-lasting interventions. Second, we gathered and linked ideas, hints, and evidence on how the process of learning with science narrative texts takes place, namely on conditions and underlying processes (i.e., the "How" question). There are many features from conditions (texts, learners, activities, wider context) and underlying processes (integration with prior knowledge, affective dispositions, and cognitive abilities) that can help to account for variability in outcomes; yet, ideas and evidence are not always tightly connected. We suggest that education and research should focus on specific narrative effects, that specify with what (texts), with whom (learners), when and where (activities and wider context) these effects occur, as well as "why" (underlying processes). We believe the proposed framing can help both make sense of previous evidence and inform future educational practices and research and provide some recommendations in this regard.

KEYWORDS

learner features, narrative effects, science learning, science narrative texts, text comprehension

The universe is made of stories, not of atoms (Rukeyser, 1968, p. 111).

1 | INTRODUCTION

Many authors contend that the challenges of science learning should be addressed by improving language and literacy processes (e.g., Klein, 2006; Morais & Kolinsky, 2016; Norris & Phillips, 2003; Webb, 2010). At the same time, these challenges are thought to stem from a fundamental difference between everyday and scientific modes of thinking (e.g., Bruner, 1986; Egan, 1997; Phillips & Norris, 2009). For the latter reason, narrative texts, which are generally viewed as temporally organized actions or events (e.g., Adam, 2011; Norris et al., 2005; Strube, 1994), have been advocated as effective tools to tackle the challenges of science learning, which is commonly based on expository texts (e.g., Arya & Maul, 2021; Olson, 2015; Solomon et al., 1992).

The idea that narrative materials can improve the understanding and retention of information is sometimes termed the *narrative effect* (e.g., Norris et al., 2005). Yet, it comes from theoretical (e.g., Bruner, 1986) and empirical (Graesser et al., 1980; Kintsch & Young, 1984; Zabrucky & Moore, 1999) works based on nonscience narrative texts. It is therefore relevant to ascertain if narrative materials actually consistently benefit science learning.

Additionally, learning occurs through the combination of different elements (Snow, 2002). Namely, texts have specific features (e.g., Adam, 1997) and are cognitively processed by readers in specific ways during learning activities (e.g., Kintsch, 1998), all these aspects interacting within and with a wider context (e.g., Adam, 1997; Snow, 2002). To give a few examples, qualitatively different contents have been used in science narrative texts (e.g., fiction, Banister & Ryan, 2001; nonfiction, Hong & Lin-Siegler, 2012), as well as different activity goals (e.g., studying, Wolfe & Mienko, 2007; evaluating text quality, Arya & Maul, 2012). Science narrative texts have also been claimed to connect to readers' social and cultural identities (e.g., Mutonyi, 2016), and to engage differently processes such as integration with prior knowledge (e.g., Maria & Johnson, 1989), emotions (e.g., Murmann & Avraamidou, 2014), and attention (e.g., Hadzigeorgiou et al., 2012). Thus, it is also important to characterize how the process of learning science through narrative texts takes place.

In short, providing a theoretically grounded examination of whether science narrative texts consistently improve memory and learning outcomes, as well as insights into the characteristics of the learning process that can lead to such outcomes, is an important step toward better understanding this science educational tool.

2 | A THEORETICAL REVIEW ON NARRATING SCIENCE FOR LEARNING

The goal of the current paper is to analyze by means of a theoretical review two questions pertaining to the topic of learning science through narrative texts that we believe to be of relevance for educators and researchers in education. To the best of our knowledge, such a review has not yet been provided.

Our first question (henceforth, the "If" question) is whether there is evidence that narrative texts have consistently benefited retention and learning from science at different levels of education. Learning and retention are relevant cognitive and pedagogical outcomes (e.g., Kintsch, 1994) whose conditions may depend on learners' educational level. Although the use of narrative educational materials is conventionally associated to young children, it has been claimed that these materials may benefit older learners as well (e.g., Klassen, 2006; Olson, 2015). This question will be examined by using a set of studies chosen on the basis of specific criteria (see Section 4).

A follow-up question (henceforth, the "How" question) concerns the characteristics of the learning process that may lead to the aforementioned educational outcomes. We will explore the conditions involved in this process, as well as the mechanisms that may underlie it, by establishing connections with a broader literature.

To our knowledge, there has been no strong theoretical framework guiding the interpretation of previous studies and the planning of future interventions. In the present theoretical review, we draw on a set of theories from relevant disciplines to accommodate the different aspects that our questions touch on.

As we aim at connecting science learning to literacy processes (e.g., Morais & Kolinsky, 2016; Norris & Phillips, 2003), we draw on pedagogical aspects pertaining to learning through reading. We chose the framework outlined by the Reading for Understanding (RAND) Reading Study Group (Snow, 2002) for several reasons. First, because it subscribes to the idea that improving learners' literacy skills promotes content learning. Second, because it recognizes the multifaceted nature of learning, providing a backbone of the conditions that should be considered when planning interventions or analyzing its outcomes. These elements combine research traditions (e.g., Pearson & Cervetti, 2015) that we see as highly relevant and complementary for the question at hands. Indeed, RAND builds its proposal around three elements that have been very present in cognitive models (i.e., text, reader, and activity), but whose interactions are context-dependent, which has been given more attention in sociocultural models. Finally, RAND summarizes and structures these elements within a policy-context, aiming at providing guidelines for research and development that focus on text-and discipline-specific reading practices. These aims are in agreement with the aims of the current review.

Because textual materials are a key aspect of our theoretical review, we also draw on a framework that describes its features. We chose text linguistics (TL), particularly the francophone line (e.g., Adam, 1997; Bronckart, 1997), because its conception of texts as social objects made up of textual and contextual features (Gonçalves, 2019) provides us with a better grasp of the features that make up science narrative texts, as well as with an important sociocultural lenses, as it brings wider sociodiscursive practices into play.

Finally, we also integrated cognitive aspects of text comprehension, as they provide valuable insights on memory and learning (i.e., the outcomes under examination) and on how readers cognitively process science narrative texts (i.e., the processes underlying the examined outcomes). We selected the Construction-Integration (C-I) model (e.g., Kintsch, 1988, 1998) because it is regarded as the most comprehensive cognitive model of text comprehension (McNamara & Magliano, 2009), dedicates special attention to science texts, and has been the predominant paradigm when conceptualizing basic processes and pedagogical practices for comprehension (e.g., Pearson & Cervetti, 2015).

Drawing on this multidisciplinary effort, we will examine the two questions outlined in the beginning of this section. In the following section, we will present the selected theoretical frameworks, underlining both their shared and specific (and thus complementary) aspects.

3 | LEARNING SCIENCE FROM TEXT: A PROPOSAL DRAWING ON THREE FRAMEWORKS

For the RAND model (Snow, 2002; Sweet & Snow, 2003), text comprehension is always a specific combination between features from the text, the reader, and the activity. These elements are highly permeable to each other's influences (dashed lines in Figure 1) and interact both within a wider sociocultural context (the wider circle surrounding these elements in Figure 1) and also with it (middle circle in Figure 1). Reading and learning by reading occur at the interfaces of these elements, and the process of learning is deemed as important as its content.

TL and RAND conceive literacy as a cultural and historical activity (e.g., Adam, 1997; Snow, 2002). Texts are viewed by TL as social objects that connect to expectations and practices

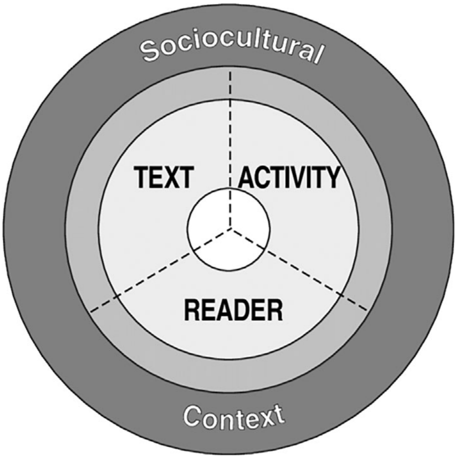

FIGURE 1 The RAND framework proposal (taken from Snow, 2002).

of the wider sociocultural context. For instance, a text with fantastic elements it expected to entertain, hence the fairy tale genre is not commonly found in science education. In other words, texts are governed by textual genres (Adam, 1997; Bakhtin, 1984; van Dijk & Kintsch, 1983), such as fable, scientific report, or cookbook, which are abstract models of what is to be expected and adopted in specific communicative situations (Bronckart, 1997). As genres have conventionalized structures, purposes and target audiences, the latter influence how texts are perceived and processed (e.g., Hidi et al., 1982; Rastier, 2001; Zwaan, 1994), connecting the textual and pragmatic components of a text (Adam, 1997).

According to TL, the textual component has different levels (see the arrows in Figure 2), namely clauses, sequences and text plans. Sequences belong to five prototypical categories: narrative, descriptive, explicative, argumentative, and dialogical. Crucially, these heterogeneous sequences are intermingled in most texts (Adam, 1997, 2011). Sequences are organized by a common text plan, which determines the global configuration of the text. In a narrative text, sequences are temporally organized; in expository texts, they are organized by topics and ideas.

Also according to TL, contextual aspects, or the pragmatic component, influence the configuration and processing of texts as well. Pragmatic aspects include enunciative features, content (or "semantic representation" in the original terminology), and communication aims (or "illocutive-argumentative"). Narrative texts often include fictional contents and have the aim of entertaining; science texts tend to stick to factuality and aim at instructing. The pragmatic and textual components interact constantly (the dashed lines in Figure 2).

What makes a text being perceived as a narrative text is therefore not homogeneous narrative sequences, but a combination of textual and pragmatic features. As, overall, texts are highly variable and heterogeneous, by definition all texts are, to some extent, "mixed" or "hybrid" (e.g., Hidi et al., 1982; Norris et al., 2005). Materials only containing prototypical narrative features have likely been extracted from larger heterogeneous texts, or carefully built that way

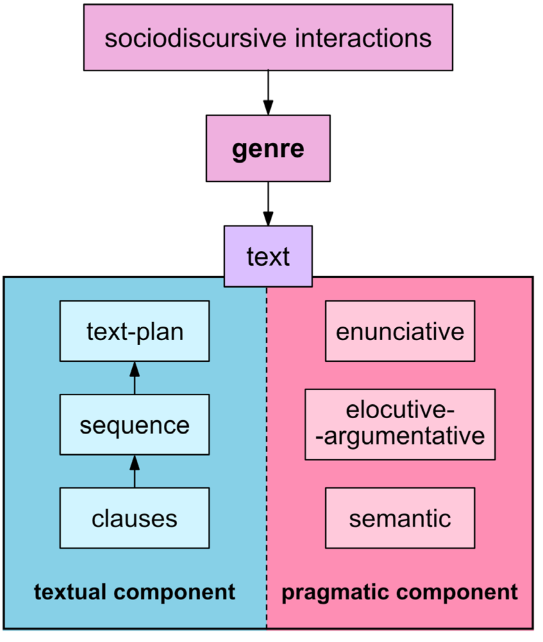

FIGURE 2 A text Linguistics' view of text features (adapted from Adam, 1997; 2008).

for a specific pedagogical or experimental purpose. They would more aptly be defined as sequences than as texts; yet, for simplicity's sake, we will refer to all materials as “texts.”

Both a novel about fictional characters' lives and a science discovery narrative have mixed features, yet the latter will more likely be perceived as “mixed”. Factual science contents (in contrast with science fiction) are not usually associated to an overall narrative structure, as this combination is present in few textual genres. This difference sets studies with science narrative texts apart from the studies on narrative comprehension they frequently draw on. Nonetheless, some science educational materials incorporate narrative features. To name a few examples, some texts add narrations about scientists thoughts and actions to provide a personal and social context to the science contents (e.g., hybrid adapted primary literature, Shanahan, 2012). Others present science contents as part of a story with characters and events (e.g., secondary literature or popular fiction, Baram-Tsabari & Yarden, 2005), or of a historical narrative that explains how concepts were discovered or developed, as to make more explicit how knowledge is constructed (e.g., epistemologically considerate texts, Kloser, 2013).

The communication aim of the text, whether externally communicated or inferred, has a crucial impact on how the activity is perceived and executed (e.g., to entertain vs. to learn), and hence on its comprehension outcomes. This notion is shared by all the theoretical models, and illustrates an interaction between the text, the activity, and wider sociocultural conventions.

Just as texts have different levels of organization, so do readers' representations of them. Drawing on a similar body of research, the three frameworks acknowledge that readers build various representations of the text's information (e.g., Kintsch & van Dijk, 1975; van Dijk & Kintsch, 1983). However, it is in the C-I model that these representation levels are thoroughly developed (e.g., Kintsch, 1998; Kintsch & Rawson, 2005). This model proposes that readers construct (i.e., represent meanings) and integrate (create a coherent representation) information into representations through an interactive interplay of text-driven and reader-driven processes, and focus on three levels of representation (see Figure 3). At the surface level, linguistic

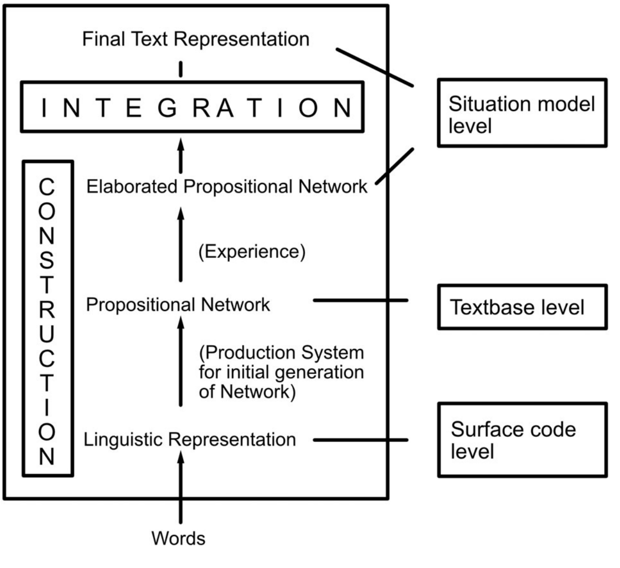

FIGURE 3 Schematic overview of the C-I model (adapted from Wharton & Kintsch, 1991).

information is represented literally (i.e., exact wording and phrasing), which is generally assumed to have little effect on comprehension (McNamara & Magliano, 2009). At the textbase level, the explicit meaning of the text is represented by a propositional network of interrelated idea units, based on words, their syntactic relationships, and inferences generated for text cohesion. These idea units are organized into higher meaning units according to global topics and their interrelationships that link larger portions of the text, and often follow the conventions of familiar schemata. Schemata are mental structures containing knowledge (e.g., elements; rules; strategies) on specific genres or discourses (e.g., fairy tale vs. informational piece), that orient and facilitate information processing and comprehension (Adam, 2011; Kintsch & Rawson, 2005; van Dijk & Kintsch, 1983). This is another illustration of how wider contextual conventions influence activities' aims and readers' cognitive processing of the text.

In addition to deriving relations between information explicitly mentioned in the text, which is fairly shallow, readers elaborate on this propositional network, generating inferences and integrating their own experience and world knowledge. Integration with prior knowledge cannot take place when this knowledge is inadequate or absent, but texts containing only already known information are also useless for learning (Wolfe et al., 1998). This deeper level of representation is called the situation model, and it involves a set of knowledge, affective dispositions and cognitive abilities, including previous experiences, motivation, memory and visual imagery (e.g., Kintsch, 1998; Xu et al., 2005). The extent to which these reader's features are engaged would depend on the interactions with features from the text, the activity, and the wider context. The C-I model acknowledges that the situation model is influenced by contextual features such as the genre, discipline, and goals (Kintsch, 1998; van Dijk & Kintsch, 1983). Yet, these features have a more supporting role in the model, which places greater emphasis on reader and text features.

The representation of a text's explicit meaning (i.e., textbase level) and its deeper integration with extra-text information (i.e., situation model) occur simultaneously and interdependently (Kintsch, 1988). Yet, for conceptual and pedagogical purposes, they are treated as memory and learning outcomes, respectively (Kintsch, 1994, 2012).

We believe that these theoretical frameworks provide a complementary multidisciplinary view of the different elements we wish to examine. Using this theoretical background, we will now describe the methods applied to tackle the two questions addressed in the present review.

4 | METHODS

4.1 | The "If" question

The aim of the "If" question was to examine evidence on memory and learning outcomes at different educational levels, having the outlined theoretical frameworks as theoretical grounding. We did this by selecting research papers that evaluated these outcomes according to a set of four inclusion criteria listed in Table 1 and analyzing their findings.

As we were interested in the role of narrative texts as science education tool, (1) texts had to be the central learning activity, and had to be read, listened to or written, as these activities are integral to literacy (e.g., Morais & Kolinksy, 2020); in addition, (2) text materials must be described as having narrative, story or novel features (the latter two convey central narrative

TABLE 1 Inclusion and exclusion criteria for selecting studies assessing memory and learning.

Criteria	Inclusion	Exclusion
1. Reading, writing, or listening to narrative texts	Texts are the central task of the learning activity	Texts are secondary
2. Definition of narrative	Narrative/narrative features, story, or novel	Only other terms are used (e.g., primary vs. secondary literature; history of science)
3. Participants	Students, from preschoolers to pre-service teachers	Others (e.g., teachers)
4. Educational outcomes	Memory or learning is one of the outcomes	Only other outcomes (e.g., engagement; understanding the nature of science)

features, e.g., temporal organization), even if other terms are used simultaneously (e.g., "hybrid"); authors' definitions were not called into question. As we wanted to examine comprehension outcomes at different educational levels, (3) participants should be students, that is, completing a formal educational degree (or in preschool, which in some countries is considered formal education) and (4) outcomes must examine memory and/or learning of contents from the scientific body of knowledge (i.e., explanations intrinsic to science, Norris et al., 2005); other outcomes could be analyzed (e.g., affective aspects; understanding the nature of science), insofar as the outcomes of interest were directly examined. To include different materials, we searched for papers in the following databases: Web of Science (very encompassing), ERIC (directed at education) and OATD (theses and dissertations). It is worth noting that the main addressed concepts were quite encompassing (e.g., science; learning; narrative) and the outcomes of interest could be described using various terms (e.g., comprehension, learning, recall, memory). As such, broader searches would produce many papers that did not fit our aims, but narrowing the search terms too much would likely result in missing relevant papers. In an attempt to achieve a balance, we decided to combine the following search terms: "science", "scientific", "narrative text", "learning", "comprehension", "memory", "retention", "recall", and "teaching". However, because effective Boolean queries including various terms can be difficult to achieve (e.g., Karimi et al., 2010; Scells & Zuccon, 2018; Wang et al., 2022), we opted for performing various searches that flexibly combined the relevant search terms, guided by the goal of examining the overall (in)consistency of previous evidence on narrative learning benefits. For these reasons, it proved challenging to apply a stricter search methodology that would have allowed to perform an exhaustive search of the retrieved results. The search terms were preferentially looked for in the abstracts. The search covered works produced between January 1990 and December 2019, as many of the seminal theoretical and empirical works on the benefits of narrative texts for memory and comprehension dated from the 1980's. Studies could be written in English, European Portuguese, Brazilian Portuguese, Italian or French, and we also did some searches using the search terms translated in these languages.

We began by checking the titles and abstracts against the inclusion criteria. When papers were related to our question, we inspected them fully for potentially relevant references. We started by inspecting the methods and results sections, to check whether memory and learning outcomes had been directly addressed. These outcomes should have been to some extent learned through a text-based activity of the study. We did not consider papers in which the learning activity was only a means of applying knowledge acquired before/elsewhere. We were interested in both quantitative and qualitative measures. Quantitative measures had to be supported by some numerical data (e.g., means; percentages; test statistics). Qualitative measures had to include some sort of analysis or explanation of the observed comprehension outcomes, connecting them with least one illustrative example. We did not include papers which only presented general statements on learning gains or on the quality/relevance of narrative tools or interventions. When papers fitted our purposes, we also checked the papers that cited them.

We analyzed the findings of 36 papers, whose main characteristics are summarized in Table 2 (see also Tables 3 and 4). Among them, 20 compared narrative materials to some control material (narrative vs. control studies; henceforth, N vs. C). In the other 16 studies, narratives were examined as stand-alone tools (narrative-only studies; henceforth, N-O). As using science narrative texts is often presented as an alternative to conventional teaching methods, directly comparing narratives with a control material provides more straightforward evidence on a potential narrative effect in science learning, while also minimizing confounding effects. However, many educators are not interested in comparing different teaching methods, but rather in exploring in more depth the narrative tool.

These different paradigms often used different methods and tasks to examine the outcomes of interest. A similar number of N versus C and N-O studies (3 and 4, respectively) used mixed methods to assess memory and learning. Yet, whereas N versus C studies used predominantly quantitative measures, the same number of N-O studies (i.e., 6) used either quantitative or qualitative measures. A similar number of N versus C and N-O studies (i.e., between 5 and 7), examined memory and/or learning using interviews, discussions or observations, and recall, retell or re-rewrite. All but four N versus C studies used written tests of questions, while N-O studies used more diverse tasks, such as story writing or journaling and drawings or hands-on-activities. Despite these differences, the general design is often similar. In about half of the studies of each paradigm, participants' knowledge was assessed before and after reading, writing, or listening to the texts (i.e., pre-post design). The remaining studies either only applied posttests or analyzed several measures in the context of case studies. Regardless of paradigm and design, some studies also included delayed assessments.

In the case of quantitative comparisons, we verified whether there was evidence that the interventions produced gains in memory and/or learning (e.g., pre vs. post measures) and whether these gains were stronger in specific conditions (e.g., with N vs. C; with different narrative texts; in delayed vs. immediate assessments). When effect sizes (ES) were not reported and sufficient information was available, we calculated them for between or within effects (Lakens, 2013). For the sake of clarity, when studies included several tasks, conditions (e.g., groups) or factors irrelevant to our purpose, we either made a selection or aggregated the results when they were of similar direction and significance (see notes of Tables 3 and 4). As regards qualitative analyses, we checked which kind of memory and/or learning outcomes were observed, and how they were patent in the illustrative examples presented by the authors.

As we will discuss further in the Results, the C-I model differentiates memory from learning. Yet, memory and learning are intertwined at both the cognitive and design level, with some measures tapping both processes at once (e.g., tests; interviews). When tests were reapplied after a delay, we considered them as assessing both learning and memory. In addition, some studies explicitly stated they included measures tapping different levels of representation but did not present results separately; in this case, we interpreted them as pertaining to both levels.

TABLE 2 Measures, design and tasks of the 36 studies selected to examine the "If" question.

	Quantitative measures	Qualitative measures	Mixed-methods	Pre-post design	Delayed assessment	Written tests or questions	Interview, discussion or observation	Recall, retell or rewrite	Writing stories or journaling	Drawings or hands-on-activities
NN versus C	17	0	3	9	7	16	4	7	1	0
NN-O	6	6	4	8	4	7	5	5	6	5
Total	24	6	6	17	11	24	9	12	7	5

Note: The first three columns refer to the outcomes we wished to examine, not necessarily to the methodology of the full study. Studies using delayed assessment could have applied a pre-post design or other (post-test or case study). A same study could have examined memory and/or learning using more than one type of task. Recall and retell could have been done orally (e.g., interview) or through writing.

TABLE 3 Characterization of the main variables (Educational level, memory outcomes, learning outcomes) of the N versus C studies.

No.	Study	Educational level	Main memory and/or learning findings	Effect sizes (when available)
1.	Maria and Johnson (1989)	Gr 5 and 7	Gr 7 scored higher than Gr 5 in Im Misc Post and Im App Post but there we no differences between N and E; Gr 7 scored higher than Gr 5 in Del App Post and N scores were higher than E scores	Im Misc Post Gr 5: $d = 0.47$ $r = 0.23$; Gr 7: $d = 0.24$, $r = 0.12$; Im App Post Gr 5: $d = 0.45$ $r = 0.22$, Gr 7: $d = 0.47$ $r = 0.23$; Del App Post Gr 5: $d = 0.76$ $r = 0.35$; Gr 7: $d = 0.24$ $r = 0.12$
2.	Maria and Junge (1993)	Gr 5	No N versus C differences in Im Rec and Del test; N Rec were longer; Rec had few Sci ideas	Im Rec (length): $d = 1.07$, $r = 0.47$; Im Rec (Sci ideas): $d = 0.20$, $r = 0.10$; Del test: $d = 0.19$, $r = 0.1$
3.	Jetton (1994)	Gr 2	No N versus C differences in Im free response or Del Rec; Rec included more on N ideas	Im story ideas: $d = 0.57$ $r = 0.27$; Im Sci ideas: $d = 0.17$ $r = 0.08$; Del Rec: $d = 0.13$ $r = 0.06$
4.	Alvermann et al. (1995)	Gr 9	E advantage in Rec & App Post	Rec: $d = -0.88$ $r = -0.40$; App $d = -0.90$ $r = -0.41$
5.	Hellstrand and Ott (1995)	Gr12	N LRN advantage in Post	$d = 0.57$, $r = 0.27$
6.	Lamartino (1995)	Gr 3	No N versus C LRN differences in Post	$d = 0.50$, $r = 0.24$
7.	Wolfe and Mienko (2007)	UndGr	No significant N versus C differences in LRN and Rec	LRN: $d = -0.13$ $r = -0.06$; Rec: $d = 0.27$ $r = 0.14$
8.	McQuiggan et al. (2008)	Gr 8	Largest Pre-Post gains for E text, followed by min N and lastly full N	N (full) versus E: $d = -0.99$ $r = -0.44$; N (full) versus N (min): $d = -0.32$ $r = -16$; N (min) versus E: $d = -0.62$ $r = -0.30$
9.	Cervetti et al. (2009)	Gr 3 and 4	General E advantage in Post LRN and Retell, but only significant for one of the Sci topics	General LRN: $d = -0.47$ $r = -0.23$; General Retell: $d = -0.46$ $r = -0.22$
10.	Negrete and Lartigue (2010)	UndGr	E advantage in Im Post MEM/LRN; N advantage in Del MEM/LRN Post	
11.	Wolfe and Woodwyk (2010)	UndGr	N MEM advantage for total text elements, but E MEM advantage for common text elements; Textbase stronger for N and decreased significantly in Del; situation model marginally stronger for E and no decrease in Del	MEM (total): $d = 0.57$ $r = 0.28$; MEM (common): $d = -0.77$ $r = -0.36$; Im textbase: $d = 0.44$, $r = 0.22$; Del textbase: $d = 0.19$, $r = 0.10$; Im situation model: $d = 0.41$, $r = 0.20$; Del situation model: $d = 0.22$, $r = 0.11$

TABLE 3 (Continued)

No.	Study	Educational level	Main memory and/or learning findings	Effect sizes (when available)
12.	Rosa (2010)*	Gr 11	N Pre/Post LRN and CCPT elaboration advantage as measured by tests, DISC and INTVW	
13.	Ritchie et al. (2011)	Gr 6	N Pre/Post LRN advantage as measured by story WRT and INTVW	PreWRT to Story Part A: $d = 1.59$, $r = 0.69$; PreWRT to Story Part B: $d = 1.16$
14.	Arya and Maul (2012)	Gr 7 and 8	Im N LRN advantage for one Sci topic and Del N LRN advantage for both Sci topics; Gr 8 students did not benefit from Radioactivity N	Im (Radioactivity) $d = 0.17$ $r = 0.08$; Im (Galilean telescope) $d = 0.43$ $r = 0.20$; Del (both topics) $d = 0.95$ $r = 0.43$
15.	Hadzigeorgiou et al. (2012)	Gr 9	N advantage in Im and Del LRN Post	Im Post: $d = 1.31$ $r = 0.55$; Del Post: $d = 1.72$ $r = 0.65$
16.	Hong and Lin-Siegler (2012)	Gr 10	No N (struggles) versus N (achievements) versus C differences in Im LRN or Rec. N (struggles) advantage (vs. other two texts) in Del LRN and Rec	Im Rec: $d = -0.35$ $r = -0.17$; Del Rec: $d = 0.67$ $r = 0.32$; Im LRN: $d = 0.07$ $r = 0.04$, Del LRN: $d = 0.90$ $r = 0.41$
17.	Reuer (2012)*	Gr 12	N LRN advantage in Post and INTVW	Chapter test: $d = 0.37$, $r = 0.18$; Exam: $d = 1.21$, $r = 0.51$
18.	Browning and Hohenstein (2015)	Gr 1, 2 and 3	N LRN advantage in all Gr, but especially in Gr 3	Total Gr: $d = 1.06$ $r = 0.47$; Gr 1: $d = 0.83$ $r = 0.38$; Gr 2 $d = 0.73$ $r = 0.34$; Gr 3 $d = 1.72$, $r = 0.65$
19.	Akarsu et al. (2015)	Gr 7	N Post LRN advantage	Pre: $d = -0.15$, $r = -0.08$; Post: $d = 1.42$, $r = 0.58$
20.	Dinsmore et al. (2017)	UndGr	Highest increase in LRN complexity with N, followed by E and decrease with persuasive text	$p\eta^2 = 0.14$

Note: White cells represent better memory and/or learning outcomes for the narrative (N) text, light gray cells represent the absence of difference between N and control (C) texts or mixed results (e.g., N advantage only in delayed measures), and dark gray cells represent better memory and/or learning outcomes for the expository (E) text. The letter *d* stands for Cohen's effect size and *r* for Pearson's correlation coefficient. Selection of conditions/tasks: in study 1, we only present the comparisons between the N text and one of the applied E texts (considerate E text) because they had equivalent length and their scores only significantly differed in Del measures. In study 4, we only present part of the applied tasks, one representative of each relevant outcome, and the Control condition was not examined. Aggregation of results: the following results were collapsed: in study 2, the results from good readers and bad readers; in study 4, the results from the Discussion web and Question/answer conditions; in study 7, the results from the two applied E texts (Topical E text and Sequential E text); in study 14, the results from both Sci topics in the Del assessmen. In study 5, effect sizes were calculated by approximation (approximately 25 students per class). Abbreviations: App, application; CCPT, concept/conceptual; Del, delayed; DISC, discussion(s); Gr, grade(s); Im, immediate; INTVW, interview(s); LRN, learning; MEM, memory; Misc, misconception; Pre, pretest; Post, post-test; Rec, recall; Sci, science/scientific; Ss, students; UndGr, undergraduates; WRT, writing/wrote; yr-o, year-old(s).
*All studies were published articles except for 2 MSc theses.

TABLE 4 Characterization of the main variables (Educational level, memory outcomes, learning outcomes) of the N-O studies.

No.	Study	Educational level	Main memory and/or learning findings	Effect sizes (when available)
21.	Banister and Ryan (2001)	Gr 4	Ss showed CCPT change from Pre/Post questions and reWRT of N; in Del INTVW Ss used more abstract ideas in retell of N than in description of Sci CCPT; some imperfections in CCPT development	
22.	Wilcken (2008)*	High school	Better Post Im and Del LRN for N with concrete details (vs. abstract N)	Im: Cohen's *dav* = 0.45; Del: Cohen's *dav* = 0.79
23.	Ritchie et al. (2008)	Gr 4	LRN was evaluated as functional and fluent use of Sci CCPT through N WRT, INTVW and field observation	
24.	Corni et al. (2010)	Gr 3	Ss drawings and WRT texts showed LRN (from descriptions to interpretation and formulation of hypotheses)	
25.	Frisch (2010)	UndGr (preservice teachers)	Average-scoring Ss (exam) used N they WRT to understand Sci CCPT more than below- or above-average-scoring Ss; teacher guidance helped Ss integrate and LRN SCI CCPT in N they WRT	
26.	Kokkotas et al. (2010)	Gr 6	Ss showed LRN in comprehension questions, classroom DISC and by implementing experiments	
27.	Tomas et al. (2011)	Gr 9	Ss were able to transform Sci knowledge to WRT accurate Sci N but this knowledge was better explained and elaborated in INTVW (which can explain decreases)	Improvements: N PreWRT to Part B: $d = 1.25$; N Part A to N Part B: $d = 0.85$; Decreases: N PreWRT to Part C: $d = 0.55$; N Part A to N Part C: $d = 0.89$; N Part B to N Part C: $d = 2.19$
28.	Kalogiannakis and Violintzi (2012)	Preschool	LRN improved from Pre to Post as assessed by INTVW and drawings	
29.	Legare et al. (2013)	5 to 12 yr-o	Specific N (desire-based, need-based, natural selection) promoted corresponding explanations of evolution in Im LRN and Rec (INTVW)	Older versus younger desire-based explanations $d = -2.29$; $r = -0.75$; Older versus younger need-based explanations $d = 3.18$, $r = 0.85$; Older versus

TABLE 4 (Continued)

No.	Study	Educational level	Main memory and/or learning findings	Effect sizes (when available)
			Older Ss recalled more content, used more need-based and evolution-based explanations, and used more evolution CCPT than younger Ss; the later used more desire-based explanations	younger evolution-based explanations $d = 6.08$ $r = 0.95$; Older versus younger evolution CCPT $d = 6.55$ $r = 0.96$
30.	Morais (2015)	Primary school (8–10 yr-o)	Ss drawings showed LRN of Sci CCPT	
31.	Lin-Siegler et al. (2016)	Gr 9 and 10	Ss who read intellectual struggle N or life struggle N had better Post LRN (tests) than Ss who read achievement N	Achievement N versus intellectual struggle N: $d = 0.04$, $r = 0.02$; achievement N versus life struggle N: $d = 0.09$, $r = 0.05$; intellectual struggle N versus life struggle N: $d = 0.06$, $r = 0.03$
32.	Mutonyi (2016)	Gr 9–11	Ss journals, focus-groups DISC and INTVW demonstrated CCPT LRN; in many cases it was retained after months/a year	
33.	Prins et al. (2017)	Gr 10 and 11	Both Gr scored high in Im and Del LRN and MEM Post (except Retell); Gr 11 had higher Im Retell; worse Del retell (less Sci information) in both Gr	Retelling between Gr: $d = 2.02$ $r = 0.71$; Retelling between sessions: Cohen's $dz = 0.74$
34.	Flynn and Hardman (2019)	Gr 12	Ss improved LRN (from first 15 Post questions to last 15 Post questions)	$d = 0.71$
35.	Morais and Araújo (2019)	Gr 8	Ss showed LRN of Sci CCPT in N WRT (explanation of Sci ideas) and creation of hands-on-activities (connecting and App of Sci ideas)	
36.	Walan and Enochsson (2019)	Preschool (4–6 yr-o) and primary school (7–8 yr-o)	Ss demonstrated different levels of SCI LRN and Rec in Del INTVW (from no identified LRN to LRN connected to reality); Im drawings supported Rec)	

Note: The letter *d* stands for Cohen's effect size and *r* for Pearson's correlation coefficient. Aggregation of results: in study 22, the results from the factors Structure and Gender were collapsed with the results of the factor Concreteness.
Abbreviations: App, application; CCPT, concept/conceptual; DISC, discussion(s); E, expository; Del, delayed; Gr, grade(s); Im, immediate; INTVW, interview(s); LRN, learning; MEM, memory; N, narrative; Pre, pretest; Post, post-test; Rec, recall; Retell, retelling; Sci, science/scientific; Ss, students; UndGr, undergraduates; WRT, writing/wrote; yr-o, year-old(s).
*All studies were published articles except for 1 PhD dissertation.

4.2 | The "How" question

The aim of the "How" question was to provide an overview on aspects that, based on the outlined theoretical frameworks, should be relevant to characterize the process of learning science through narrative texts. We did this by gathering and linking ideas, hints, and evidence, establishing connections with a broader literature in a more exploratory mode.

The 36 studies examined in the "If" question were also used as the main basis for the analysis, whenever they contributed to characterize the process leading to the observed learning outcomes. Many of the variables not included in the analysis of the "If" question were relevant and thus were considered here, such as prior knowledge, interest, or readers' social and cultural identities. However, we also drew on and established connections with a broader literature with that same purpose of characterizing the underlying process in mind. We included information from theoretical and empirical works on science narratives and on conventional narratives more generally. These were mostly retrieved during the previously described searches, especially by inspecting the full papers for relevant references. When relevant, we also made connections to more general research on learning and science learning, and to literature on specific features from the examined conditions and processes. These literatures were again not exhaustively searched or described, as our goal was to sketch a theoretical overview. The presented evidence is also not exhaustive, but instead illustrative. We included different kinds of evidence: direct, that is, the feature has been investigated in direct connection to learning outcomes stemming from science narrative texts; indirect, that is, the feature has been investigated in a learning intervention that used science narrative texts but was not directly connected to learning outcomes; or more tentative, namely, the feature was investigated using nonscience narrative texts.

We organized the analysis of this question under two main sub-questions, each addressing specific features. Namely, drawing on the outlined theoretical frameworks, we mapped a set of conditions (texts, activities, and populations, as well as their interactions with the wider context) and underlying processes (prior knowledge, affective dispositions, and cognitive abilities) that are relevant to characterize the process of learning through science narrative texts.

5 | RESULTS

5.1 | If: Can science narrative texts improve memory and/or learning?

The presentation of results will be organized according to the memory versus learning distinction provided by the C-I model, and the tasks usually used to assess them (e.g., Ferstl, 2001; Kintsch, 2012).According to the C-I model (Kintsch, 1988; Kintsch, 1994), memory and learning correspond to different levels of representation. Memory is related to the reproduction and paraphrasing of information, being strongly associated to the textbase level. Learning is related to changes in knowledge and requires integration of new content within prior knowledge, being more closely associated to the situation model level. As outlined, authors from N versus C and N-O studies used different tasks to measure these outcomes. Table 3 lists all the examined N versus C studies, along with information on the main variables (educational level and memory and/or learning outcomes) and, whenever possible, the corresponding ES. Table 4 presents the same information for the N-O studies. As can be seen in Tables 3 and 4, both types of studies included a range of educational levels, but whereas N-O studies mainly included preschoolers and only seldom undergraduates, N versus C studies focused more on the latter and none included preschoolers.

Memory and learning are preferentially measured by specific tasks (e.g., Ferstl, 2001; Kintsch, 2012). Memory is often tapped by free recall (e.g., 7[1]), recognition questions (e.g., multiple-choice; 10), comprehension questions about explicitly mentioned information (e.g., fill-in questions, 10; interviews, 21), questions that probe automatic inferences based on explicitly mentioned information (e.g., 14), and questions that evaluate retention, such as comparing immediate and delayed measures (e.g., 16), or repeating learning assessment sometime later (e.g., test, 15).

N versus C studies reported mixed results, using mostly written tasks and some interviews (see Table 3). Among undergraduates, there was no overall difference between text types in immediate recall (7) but depending on the recalled items (the total items of the text versus the items common to both texts) and level of representation (textbase vs. situation model), either a narrative or expository advantage was found. Delayed recall among younger students (primary: 3; first middle school year: 2) did not benefit from one text type specifically. Expository gains in immediate recall were reported among primary and middle school students (e.g., 4, 9, respectively; the latter just for one topic). Narrative gains, on the other hand, were seldomly observed in immediate recall (14, but only for one topic) but were clear cut in delayed assessments of students from several educational levels (middle school: 1,14, 15; high school: 16; university: 10; medium to large ES). It is interesting to note that when only one of the presented topic yielded a text type advantage, this topic had been deemed as less interesting (14) or more difficult (9) by students. Many ES were medium to large, with some exceptions (1, 2, 3, 7, 9, 11, 14, and 16).

N-O studies reported encouraging retention outcomes (see Table 4). In interviews, primary and preschool students showed good recall of information after 3–5 months (21, 36). In the latter case (36), drawings made by the students supported recall and there were different levels of performance. In middle-to-high-school students, the retention interval could amount to 1 year (32). Among high schoolers, after a one-week delay, one study found that the narrative texts with concrete details promoted better retention than their abstract counterparts (22; large ES), and another found that students rewrote the narrative they had read using less factual information (33; large ES). Intriguingly, when comparing primary students' retelling of the narrative with their explicit description of the scientific model it contained (both after 3 months), more abstract science ideas were included in the former. When directly compared, older students recalled more content than younger ones in immediate measures (29, 33).

Learning can be assessed by problem-solving tasks that demand the transfer or the application of information (e.g., complex problem-solving, 16; implementing an experiment, 25; applying classroom-acquired knowledge in other contexts, 22), by inference questions that cannot be answered with explicitly mentioned information (e.g., 14), and by questions directed at determining knowledge change or improvement, such as pre- versus posttests (e.g., 7) and delayed reassessments (e.g., 1).

The pattern of results from N versus C studies mirrors the previously presented one for memory but contains more findings (see Table 3). In some studies, there were no significant differences between text types in immediate measures using written questions or tests (2, 3, 6, 7; primary school, middle school, and university). Other studies reported an expository advantage in immediate measures of mostly the same kind (study 9 used interviews) and among the same educational levels (9, 4, 8, 11). However, more studies reported a narrative advantage. This learning advantage was scarcely observed on immediate measures (primary school: 18; middle school: 14; interview and written test, respectively). However, it was often reported in delayed assessments (1, 10, 14, 16; primary school, middle school, and high school) and in assessments in which the "moment" of learning is harder to categorize (5, 12, 13, 15, 17, 18, 19; primary school, high school, and middle school), as they were developed through the course of several weeks (e.g., school term). Two of these studies also assessed learning through discussions and interviews (e.g., 12, 13, and 17) and story writing (13). When directly compared, older students tended to score higher (1, 18), and in one study older students did not benefit from the narrative text in one of the science topics (14). ES were mostly medium to large (exceptions: 1, 2, 3, 9, 14, 17, and 20).

Positive outcomes relating to learning were also reported in N-O studies, at different educational levels and using a more varied set of tasks (see Table 4). Among preschool, primary school, and middle school students, conceptual appropriation and development was observed with written questions (21, 26, and 31), interviews (21, 23, 27, 29, and 36), drawings (24, 28, and 30), written stories (23, 27, and 35), created hands-on-activities (26, 35) and field observations (23). However, in many studies, knowledge demonstration was not without imperfections, as evaluated by incorrections or naive concepts in drawings and interviews (21, 28, and 36) and even by a decrease in the demonstration of learning in the last part of the task of story writing (27). Moreover, in one study conducted with participants from 5 to 12 years-old, younger participants endorsed more naive and anthropomorphic explanations of evolution (desire-based), whereas older students endorsed more natural selection explanations and used more evolution concepts (although they also endorsed more need-based explanations as well). Many of the older students (including undergraduates, 25) demonstrated conceptual learning and refinement not only in written tests (25, 31, 33, 34), but also through journaling (32), discussions and interviews (32), as well as story writing (25). ES were medium to large, with one exception (31).

Overall, although there is encouraging evidence for the use of narrative texts in science learning, it is difficult to build a clear pattern from these results. This can be partly explained by the variability in features, which, beyond educational level, can be found in these studies and impact memory and learning outcomes. We provided a few hints of such features (e.g., narrative text elements, students' interest, configuration of the activity) that can help characterize this learning process. We will delve further into this question in the next section.

5.2 | How: Characterizing the process of learning from science narrative texts

5.2.1 | Under which conditions

Text features

Regardless of the narrative educational materials used, it is useful to consider how their features can specifically impact memory and learning, so that the conditions in which they are more effective for comprehension can be ascertained (e.g., Norris et al., 2005). The way narration is structured (i.e., temporally organized events) is proposed to aid memorization and learning (e.g., Prins et al., 2017; Strube, 1994). This textual feature resembles the way the human mind organizes experiences (Bruner, 1991; Fisher, 1987; Kintsch, 1998), and there is evidence that people build temporally organized representations of texts even when that structure is absent (e.g., Claus & Kelter, 2006). Two studies that used science narratives report results in line with this idea. One of them reports that, compared to the narrative text, the expository text more frequently caused chronological confusions that interfered with comprehension (Browning & Hohenstein, 2015). In another, the expository text presenting events in a temporal order promoted greater knowledge integration than another presenting events by topics and, interestingly, than the narrative text (Wolfe & Mienko, 2007). Yet, although the temporal expository and narrative texts had a similar structure, they differed greatly in another text feature, namely its content, a pragmatic feature.

Narration contents tend to focus on personal and social events (e.g., Arya & Maul, 2012; Corni et al., 2010; Klein, 2006), and science narratives have portrayed these contents differently. They can include fictional elements (e.g., Wolfe & Mienko, 2007), which are strongly associated to sociocultural conceptions and practices of narratives, or stick to factual information, which is more characteristic of science discourse. Examples of the fiction elements used in science narratives are anthropomorphism (e.g., Banister & Ryan, 2002; Cervetti et al., 2009), myths (e.g., Kalogiannakis & Violintzi, 2012), and fantasy and science fiction (e.g., Akarsu et al., 2015; Wolfe & Mienko, 2007). Some authors argue that, because fiction suspends disbelief, it creates unrestricted hypothetical worlds that are useful to illustrate complex (e.g., Negrete, 2005) or counterintuitive (Browning & Hohenstein, 2015) science concepts. At the same time, there is the concern that fantastic and anthropomorphic elements may make it difficult to separate fact from fiction, promoting inaccuracies and misconceptions (e.g., Broemmel & Rearden, 2006; Gomez-Zweip & Straits, 2006), or animistic or teleological explanations of science (Klein, 2006).

Many studies that used fictional elements in science narratives report positive memory and/or learning outcomes (e.g., Akarsu et al., 2015; Corni et al., 2010; Kalogiannakis & Violintzi, 2012), and one suggests that anthropomorphic elements enhanced students' recall of ideas (Banister & Ryan, 2001). Yet, others report that fantastic and anthropomorphic elements interfered with recall, with students exhibiting more misconceptions (Cervetti et al., 2009), less scientifically accurate interpretations (Legare et al., 2013), and difficulties in separating facts from fiction (Prins et al., 2017) or in integrating science contents into stories (Frisch, 2010; Tomas et al., 2011). Students also recalled more story than science ideas from the texts (Jetton, 1994; Maria & Junge, 1993; Wolfe & Woodwyk, 2010), and some expository materials benefited knowledge integration further (Wolfe & Mienko, 2007; Wolfe & Woodwyk, 2010). Importantly, the interference of interesting yet irrelevant elements (e.g., Garner et al., 1989), such as fiction, seems to depend on how coherent and intertwined with the text's topic these elements are (Glaser et al., 2009; Lehman et al., 2007; Negrete, 2005). In at least one study, the authors acknowledged that this was not the case in their materials (Wolfe & Mienko, 2007).

Despite their close pragmatic association, narrative materials do not necessarily contain fictional contents. Examples of factual or feasible fictional information used in studies with science narratives are depictions of daily or contemporary events (Dinsmore et al., 2017; Reuer, 2012; Rosa, 2010) and historical/discovery accounts (e.g., Arya & Maul, 2012; Lin-Siegler et al., 2016). There are different proposals on how scientists should be portrayed in the latter. Some educators argue that a romanticized view of scientists, that brings out heroic and wonder-like qualities, can facilitate learning (Egan, 1997; Hadzigeorgiou et al., 2012). Other educators argue that scientists should be portrayed in a realistic and accessible way that highlights their struggles and challenges (e.g., Arya & Maul, 2012; Hong & Lin-Siegler, 2012) and avoids stereotypical images of innate ingenuity and monumentality (e.g., Allchin, 2003; Solomon, 2002). In one study that followed the former approach, students learned better with the narrative than with the expository text, and more than half associated romantic qualities to science knowledge in their journals (Hadzigeorgiou et al., 2012). In other studies, however, narratives focused on scientists' achievements and innate intelligence did not produce the same learning gains as narratives focused on scientists' struggles did (Hong & Lin-Siegler, 2012; Lin-Siegler et al., 2016). Narratives portraying scientists' discoveries also yielded better learning than their expository counterparts (Arya & Maul, 2012).

The use of historical-based accounts is claimed to serve another purpose, which is to connect scientific knowledge with the social and cultural context in which it was discovered or developed (Arya & Maul, 2021; Klassen, 2007). This provides students with a better grasp of what science is and how it works (i.e., the nature of science), averting misconceptions that can be harmful to learning (e.g., Allchin, 2003; Clough, 2011). Some authors further contend that the scientific process is fairly narrative in itself (e.g., Bruner, 1996; Hadzigeorgiou, 2016; Larison, 2018).

Science narratives have been reported to encourage students to challenge their perceptions (e.g., Arya & Maul, 2021; Dinsmore et al., 2017; Erten et al., 2013) and hold a more accurate image of science, such as viewing it as a process (Evangelista & Zimmermann, 2008; Leipzig, 2018). One study reported that the narrative text made students use more evidence in their responses and display a more complex learning of the science contents (Dinsmore et al., 2017).

Activity features

Besides the text's content, it is important to consider situational aspects or circumstances of learning, pertaining to the activity and its interactions with the wider context (Snow, 2002). One feature important to consider is the goal(s) of the activity. Educators and researchers may communicate goals to students, but students generate their own goals, which can be influenced by existing schemata. People often draw on schemata when interpreting texts (e.g., Adam, 2011; van Dijk & Kintsch, 1983), which include genre-specific processing strategies, activated according to the knowledge of what is usually expected from texts with specific features (e.g., Hidi et al., 1982; Rastier, 2001).

Instructional/study goals can favor the activation of expository-processing strategies, as learning is associated with this kind of materials (e.g., Kloser, 2013; Wang, 2009). A text with features associated to the narrative textual genre (e.g., temporal organization; fictional information) may activate an entertaining aim, stemming from socioculturally-based expectations. The overtly communicated goal (e.g., to learn) may thus conflict with activated pragmatic knowledge (e.g., to entertain), and interfere with comprehension (Snow, 2002). Indeed, students exposed to science narratives have expressed that narratives were not adequate to learn science (Prins et al., 2017), or were surprised to have learned from them (Murmann & Avraamidou, 2014). Additionally, students may activate a story or an informational mode depending on the activity they are doing (e.g., short stories and drawings vs. instructed group work, respectively, Murmann & Avraamidou, 2014).

Studies using science narratives have communicated different goals to students. For instance, some goals were related students' own evaluation (understanding how students make sense of difficult information; Alvermann et al., 1995; solving a mystery; McQuiggan et al., 2008; studying; Wolfe & Mienko, 2007; Wolfe & Woodwyk, 2010), and students tended to benefit more from expository texts in these cases. In one study where the goal was not related to students' own evaluation, but instead to the evaluation of the texts, they benefited more from the narrative texts (Arya & Maul, 2012). Some goals were related with entertainment, such as solving a mystery (McQuiggan et al., 2008), where students benefited more from the expository text, or co-writing an ecological mystery (Ritchie et al., 2008, a N-O study), in which case students demonstrated written and spoken fluency of the scientific concepts. One study specifically manipulated the activity's goal (hearing the same text as a "story" or as a "book") but, regardless of the instruction, all students focused more on story elements (Jetton, 1994). As 2nd graders, these students must have been familiar to narrative texts via social practices, yet have little contact with expository materials, leading them to activate story-processing strategies under both conditions.

Another relevant feature of the activity is its duration. As commented in the analysis of the "If" question, studies ran over the course of weeks tended to yield more positive results for narrative than for expository texts. Authors from one study reasoned that the narrative storyline may have overloaded students' cognition, which was assessed on the same day (McQuiggan et al., 2008). Duration can also provide the opportunity to integrate more and varied activities. Examples of activities included in longer studies with science narratives, but not shorter ones, are fieldtrips (e.g., Ritchie et al., 2008), journaling (e.g., Hadzigeorgiou et al., 2012; Mutonyi, 2016), and preparing hands-on activities and experiments (e.g., Kokkotas et al., 2010; Morais & Araújo, 2019. Reasons for the learning gains afforded by longer and more varied activities may include knowledge consolidation (e.g., Squire et al., 2015) and increased meaning making (e.g., Bruner, 1990). Finally, the duration and variety of learning activities are likely to vary depending on economic and cultural factors of the learning setting (Snow, 2002), such as socioeconomic status (SES) and cultural features of the neighborhood.

Reader features

The reader is at the center of the learning process. In addition of differing by their educational level, readers may have varied social and cultural identities. At the same time, developing text materials and learning from them takes place in a sociocultural context that reflects the interpretations of specific cultural groups (e.g., Adam, 1997; Snow, 2002), particularly the dominant ones (e.g., Arya & Maul, 2021; Phillips Galloway et al., 2020). The underrepresentation of historically/currently marginalized social and cultural groups in mainstream science poses challenges to learning (e.g., Jackson et al., 2016; Harper & Kayumova, 2022; Visintainer, 2020). The need of integrating the thinking and learning dynamics of these marginalized groups has been stressed by some authors (e.g., Harper & Kayumova, 2022; Lee & Grapin, 2022; Mutonyi, 2016), who claimed that cultural background impacts text interpretation and knowledge construction (e.g., Greenfield, 1997, as cited in Arya & Maul, 2021; van Dijk, 2001, as cited in Arya & Maul, 2012), and that learners feel the need to see people like them doing science (e.g., Arya & Maul, 2021; Bowman et al., 2022; Gilbert et al., 2005).

Being culturally relevant mental models about the world (Bruner, 1986; Kintsch, 1988), narratives are proposed to help readers connect to science by bringing them closer to familiar and relevant contexts (e.g., Avraamidou & Osborne, 2009; Graesser et al., 2002). This issue has been tackled in some studies with science narratives. In one study, preschool Greek children learned about volcanoes through a Greek myth (Kalogiannakis & Violintzi, 2012) and in another, Ugandan students made use of cultural tools like proverbs and stories to learn more about HIV, a very socially relevant issue in their country (Mutonyi, 2016). Both studies report engagement and learning gains. Other studies found through interviews that narratives detailing episodes of discovery were effective in reaching different genders and cultural backgrounds. Because they felt they could also be scientists, students found science more relatable and interesting, which may have boosted their comprehension of the contents (Arya & Maul, 2021; Lin-Siegler et al., 2016). This sense of relatedness may be dependent on the level of match between the text's social and cultural elements and the student's own background (Lin-Siegler et al., 2016).

Additionally, SES has well established effects on brain and cognition (for a review, see e.g., Farah, 2017), including in science learning (e.g., Lee & Luykx, 2007; Yang, 2003). There is some evidence that science narratives work well (e.g., Lin-Siegler et al., 2016; Mutonyi, 2016) or better than expository texts (Arya & Maul, 2012; Hong & Lin-Siegler, 2012) among middle and high-school participants from low and middle-income backgrounds. Importantly, SES is often confounded with ethnicity (e.g., Cheng & Goodman, 2015), and in some studies the students who benefited more from narratives were both from low SES and predominantly Latinx and Black (Arya & Maul, 2012; Lin-Siegler et al., 2016). Yet, another study with mostly Black students found worse results in the narrative condition (Alvermann et al., 1995).

5.2.2 | Through which underlying processes

Ascertaining the extent to which reader's features are engaged during the process of learning from science narrative texts can help understand how these texts generate memory and learning outcomes (e.g., Norris et al., 2005). As such, in this section these features will be framed as underlying processes.

Integration with prior knowledge

The notion than learning from text requires linking and integrating new information with prior knowledge is central to the C-I model (e.g., Kintsch, 1998). These processes have been vastly investigated using science expository texts, which are more dependent of the integration processes than nonscience narrative texts (e.g., Best et al., 2008; McNamara et al., 2011).

Narrative texts are proposed to provide meaningful organizing structures (e.g., Negrete & Lartigue, 2010; Strube, 1994) that help activate prior knowledge (e.g., Leipzig, 2018; Maria & Johnson, 1989) and integrate information (Negrete, 2005; Prins et al., 2017). For these reasons, they can be particularly useful as scaffolding tools for beginner or struggling learners (e.g., Gilbert et al., 2005; Klassen, 2007; Mutonyi, 2016).

Narratives have been successfully used to derive science teaching methodologies for preschool (e.g., Kalogiannakis & Violintzi, 2012; Morais, 2015) and primary school (e.g., Corni et al., 2014) students. Additionally, it has been shown that high-school students with very little prior knowledge were able to develop adequate scientific understanding through science narratives (Prins et al., 2017), or were the only ones demonstrating learning gains (Flynn & Hardman, 2019). Yet, significant correlations between prior knowledge and learning through science narratives were not always found (Wilcken, 2008). When compared to expository texts, results are mixed. Even though text type did not have an overall impact in the learning and memory of undergraduates, it interacted with their prior knowledge: students with minimal prior knowledge learned better with the narrative, and students with higher knowledge learned better with the expository text (Wolfe & Mienko, 2007). Prior knowledge did not correlate with narrative text recall, but it correlated with expository text recall (Wolfe & Mienko, 2007; Wolfe & Woodwyk, 2010). Middle-school students who had previous contact with the topic also benefited less from the narrative texts (Arya & Maul, 2012; but prior knowledge was not directly measured). However, younger middle and primary school students all benefited more from the narrative text, regardless of age and prior knowledge (Browning & Hohenstein, 2015; Maria & Johnson, 1989).

There is also some evidence that narratives may be used as scaffolding tools for struggling learners. In Reuer (2012), narratives were on average more effective for learning than textbooks, but especially so for average and low achievers, a pattern matched by the students' own perceptions. Narrative texts also improved the learning of below-average and average achieving students (Hadzigeorgiou et al., 2012; Hong & Lin-Siegler, 2012; Lin-Siegler et al., 2016). Yet, in one study it did not particularly benefit struggling readers (Maria & Junge, 1993).

Affective dispositions

Even though affective processes have received less attention than cognitive ones in science learning research, their impact has been documented (e.g., engagement, Fredricks et al., 2018; self-efficacy, Britner & Pajares, 2006; emotions, Sinatra et al., 2014).

Narratives have been consistently suggested to have a positive effect on students' affective dispositions (e.g., Avraamidou & Osborne, 2009; Bruner, 1986; Norris et al., 2005). Studies using science narratives report a range of results in support of this idea. Students seem to react with interest to the reading (e.g., Arya & Maul, 2021; Hadzigeorgiou et al., 2012) and writing (e.g., Evangelista & Zimmermann, 2008; Tomas et al., 2011) of science narratives. Students' engagement was also manifest by the expressiveness of their drawings about the narrative-based intervention (Morais, 2015). Moreover, students described the experience of reading (e.g., Prins et al., 2017; Reuer, 2012) or writing (Ritchie et al., 2008; Ritchie et al., 2011; Tomas et al., 2011) science narratives as enjoyable and engaging, writing more (Hadzigeorgiou et al., 2012) and more positive (Akarsu et al., 2015) journal entries than students in the expository text condition. High levels of immersion during narrative-based science learning activities were also reported (McQuiggan et al., 2008; Murmann & Avraamidou, 2014). Finally, students with low self-reported levels of interest in science increased this interest by reading narratives about scientists' struggles (Hong & Lin-Siegler, 2012), and one study found a significant positive correlation between students' interest in the narrative texts and their performance on an science exam (Reuer, 2012).

Interventions with science narratives have also reported behaviors suggestive of active learning, an intrinsically motivated type of learning (Deci & Ryan, 1982) expressed through autonomy, initiative, and responsibility for one's learning (Kane, 2004). Students were curious (e.g., Akarsu et al., 2015; Morais & Araújo, 2019), participated actively (e.g., Evangelista, 2008; Kokkotas et al., 2010), engaged in the preparation and execution of tasks (e.g., Klassen, 2007; Kokkotas et al., 2010; Vrasidas et al., 2015) and spontaneously wrote stories (Akarsu et al., 2015) and planning notes (Klassen, 2007). Students also showed interest in learning more about the science topic (Evangelista, 2008; Rosa, 2010) and proactively made additional research on it (Evangelista, 2008; Hadzigeorgiou et al., 2012).

Some studies found evidence that learning with science narratives impacted the willingness and belief in the capacity to achieve by positively affecting the ratings of self-efficacy (McQuiggan et al., 2008; Tomas et al., 2011) and self-confidence (Flynn & Hardman, 2019).

Finally, many authors contend that the science learning gains prompted by narratives are in part due to its ability to involve readers emotionally, particularly with the thoughts, feelings, and actions of characters (e.g., Banister & Ryan, 2001; Hadzigeorgiou et al., 2012; Lin-Siegler et al., 2016; Murmann & Avraamidou, 2014). Emotions are a fundamental part of the narrative experience (e.g., Bruner, 1986; Egan, 2005; Oatley, 2016) and it is known that understanding nonscience narratives recruits a wide brain network (e.g., Mar, 2004; Mason & Just, 2009; Xu et al., 2005) that includes areas related to emotional processing, perspective-taking, and theory of mind.

Studies with science narratives offer some concordant evidence. A range of emotional responses has been observed: enthusiasm and excitement (e.g., Hadzigeorgiou et al., 2012; Vrasidas et al., 2015); laughter (Banister & Ryan, 2001; Klassen, 2007); comments of how enjoyable and fun the intervention was (Murmann & Avraamidou, 2014; Tomas et al., 2011); and other emotionally-charged appraisals (Mutonyi, 2016). Some studies offer evidence of students' emotional involvement with the text's characters, whether they were scientists (e.g., connecting with the scientist's life and work, Hadzigeorgiou et al., 2012; Lin-Siegler et al., 2016) or fictional characters (adopting the character's point of view, Jetton, 1994; Murmann & Avraamidou, 2014).

Cognitive abilities

Studies with science narrative texts also mention cognitive abilities that have an established role in the comprehension of narratives (e.g., attention: van den Broek et al., 1999; mental imagery; Sadoski et al., 1990) and expository science texts (e.g., attention: van den Broek, 2010; working memory: Linderholm & van den Broek, 2002).

Narratives are thought to capture attention because they center on human action (e.g., Banister & Ryan, 2001; Bruner, 1986; Corni et al., 2010). Studies that used science narratives reported that students were more focused and attentive than in regular classes (Hadzigeorgiou et al., 2012; Morais & Araújo, 2019; Murmann & Avraamidou, 2014).

One study comparing science narrative and expository texts directly examined working memory (Wolfe & Mienko, 2007), namely, the processes involved in the maintenance and manipulation of information during cognitive tasks (Baddeley & Hitch, 1974). Contrary to previous studies cited by Wolfe and Mienko (2007) on both expository texts (e.g., Linderholm & van den Broek, 2002) and narrative texts (e.g., Hambrick & Engle, 2002), working memory did not predict learning or recall from either text.

Some studies with science narrative texts offer evidence on the engagement of mental imagery/visualization abilities. Students claimed to have been able to visualize invisible concepts that were difficult to understand (Akarsu et al., 2015) and that traditional school texts did not promote visualization the same way narrative texts did, making them less memorable (Prins et al., 2017). Students also described elaborate mental images of the stories, seemingly using them as a basis for understanding the learning activity (Murmann & Avraamidou, 2014; Vrasidas et al., 2015). One study (Wilcken, 2008) found that narrative texts with concrete details were more easily understood and remembered than more abstract narratives, possibly due to an enhancement of mental imagery (e.g., Driscoll, 2000, as cited in Wilcken, 2008).

Imagination is proposed to encourage processes that can aid the learning of complex and counterintuitive science contents, such as divergent thinking (Browning & Hohenstein, 2015), suspension of disbelief (Alvermann et al., 1995; Browning & Hohenstein, 2015), and the envisioning of different realities (Bruner, 1986; Gilbert et al., 2005), including scientists' own reality (e.g., Arya & Maul, 2012). In some studies, using science narrative texts, teachers (e.g., Klassen, 2007; Vrasidas et al., 2015) and students (Tomas et al., 2011) claimed that these materials allowed students to exercise imagination to a greater extent than traditional activities.

The use of imagination is thought to be closely connected to the interpretative nature of narrative texts, as readers must draw on their imagination to fill parts that are ambiguous or left unanswered (e.g., Bruner, 1986; Klassen, 2007; Negrete & Lartigue, 2004). This interpretative effort is proposed to trigger high-level abilities such as abstraction, thinking and reflection (e.g., Bruner, 1991; Klassen, 2007; Rosa, 2010). In some studies, specific kinds of narrative (personal, Skydsgaard et al., 2016; historical, Evangelista, 2008; science-fiction, Vrasidas et al., 2015) helped students reflect and develop critical thinking skills.

Finally, specific features from these texts should engage abilities not directly discussed in studies using science narratives, such as inference generation and executive functioning. Namely, filling ambiguous or unanswered parts should involve inference generation, processing multiple perspectives (Bruner, 1996) can engage cognitive flexibility, and event sequentiality (e.g., Negrete & Lartigue, 2004; Norris et al., 2005; Reuer, 2012) might trigger planning abilities. These abilities are deemed crucial to learn from text by the C-I model (Kintsch, 1998). They have been often compared in non-scientific narrative and expository texts, but have been reported to be more important for the comprehension of the latter (e.g., Eason et al., 2012; Wu et al., 2020).

6 | DISCUSSION

The current theoretical review had a twofold aim: to examine if science narrative texts have consistently benefited learning and/or memory outcomes (the "If" question) at different educational levels, and to provide an overview of aspects that characterize the learning process leading to such outcomes (the "How" question). These aims were grounded on three theoretical frameworks based on concepts from pedagogy, text linguistics, and cognitive psychology.

The "If" question revealed encouraging results for the use of science narrative texts. Students from different educational levels benefited from narrative texts in memory and learning outcomes. However, this advantage was particularly marked in delayed assessments and in longer-lasting interventions. Despite the strong pragmatic association between narrative texts and younger students, these results suggest that narrative texts can be appropriate science education tools to students of diverse educational levels. However, there is a need for more studies with higher level students, such as undergraduates.

Despite this overall pattern, it is not always clear which representation level benefited most from narrative texts, partly because they were not always differentiated in the studies. Importantly, narrative texts did not always provide an advantage to expository texts. This lack of consistency contradicts the idea of a single narrative effect. We further argue that such a narrative effect fails to consider the multifaceted nature of the learning process, which is further highlighted by the unavoidable variability of conditions between different studies.

In the "How" question, we provided an overview of conditions and processes that are part of the learning process and may impact its outcomes, attempting to connect them with evidence on learning outcomes whenever possible. As regards conditions, we discussed how different text features can differ in science narrative texts, how aims and duration can differ in learning activities, and how readers' social and cultural identities can vary. As for processes, we discussed how integration with prior knowledge, affective dispositions and cognitive abilities can be engaged by science narrative texts. This mapping stems from the three theoretical frameworks.

Based on our analysis, we propose that education and research should focus on specific narrative *effects*, that specify with what (texts), with who (learners), when and where (activities and wider context) these effects occur.

We should, however, keep in mind that these conclusions derive from a theoretical review, in which we mainly attempted to understand the narrative effect in relation to science learning. In the "If" question, we made a restricted analysis, focusing on the (in)consistency of previous results, and not a comprehensive analysis of the existing literature. This choice reflects a set of difficulties. First, as already commented on in the Method, developing a unified search strategy proved difficult. In addition, we wanted to include various types of data, stemming from quantitative, qualitative, and mixed methods studies, as well as from masters' theses and PhD dissertations. Satisfactorily integrating quantitative and qualitative data in systematic reviews can prove challenging, and including data from different research levels likely introduces varying levels of rigor. Combined, these issues may arguably increase a false sense of precision. However, despite these challenging aspects, we are not in a place to say that we ensured that all relevant previous evidence was included. Future studies may use our theoretical proposal and qualitative suggestions to perform mixed methods syntheses (e.g., Heyvaert et al., 2013), providing systematic reviews that combine qualitative and quantitative evidence and research elements, and hence add systematicity and comprehensiveness to our approach.

In addition, although we considered theoretical concepts from three different domains, we did not provide a fully unified and integrated theoretical frame, which, although was not the purpose of the current work, could undoubtedly be useful. Our interpretations and the reach of our findings are necessarily limited by these choices.

Nevertheless, in addition to supporting the interpretation of previous results, the theoretically-grounded mapping we provided can also contribute to design future education practices and interventions, as well as research. Embracing these different features of the process of learning science through narrative texts can also help bring together multidisciplinary educators and researchers interested in these educational tools. We provide some recommendations below.

7 | RECOMMENDATIONS FOR RESEARCH

According to RAND, it is important to distinguish what readers take from the activity (i.e., which outcomes) from what they bring to the activity (i.e., which underlying processes). Whereas there are several evidence on the former, which is particularly interesting for educators, the latter is less understood, and researchers can help shed light on it. Examining how features from texts, learners, and activities impact learning from science narrative texts can thus provide valuable insight as to in which conditions and *why* these materials work.

It is clear from the hints we gathered and linked, that ideas and proposals about why narrative texts can improve science learning have not been tightly connected to evidence. For example, is increased mental imagery what makes concrete details improve memory and learning? Is visualization more important when learning from some science topics? And how do these elaborate images specifically promote content learning? Another example is the claim that science narrative texts allow students to exercise imagination to a greater extent than traditional activities. How does the use of imagination promote science learning? Is it by activating divergent thinking, as it has been suggested? At the same time, many abilities that have been extensively investigated with nonscience narratives have not been addressed, or even mentioned, in studies using science narrative (e.g., perspective-taking; inference generation; executive functioning). More directed research can help fill in the gaps.

It is also noteworthy that many justifications for the use of science narrative texts as learning tools stem from theoretical or empirical works using more conventional, nonscience, narratives. The latter body of literature can certainly provide interesting sources for future research. Yet, science and nonscience narrative texts differ in important ways, namely in terms of their connections to wider sociocultural practices such as learners' conceptions of textual genres. Because of this, they can engage abilities very differently. In a related vein, researchers should also consider the difference between sequences and textual genres. It is perfectly fine to develop and investigate highly prototypical narrative materials, but any conclusions drawn should pertain to the level of sequences, and not to be generalized to a textual genre as a whole. Instead, more experimental research can benefit from focusing on specific and well-defined text features.

As science narrative texts contradict socially-based expectations, they may activate different processing strategies. It would be interesting to check whether resulting conflicts can be minimized, or even activated in a complementary manner. Another activity-related aspect worth examining is why narrative learning gains are less evident in shorter interventions.

In addition, the effectiveness of any feature will likely vary according to variability in other features, so these interactions are important to keep in mind. It would be important to ascertain what narrative features are more likely to engage specific processes, among which readers and using which activities. For instance, examining which features promote deep thinking and reflection would inform how to build or select science narrative texts that are appropriate for more complex learning and knowledgeable learners. It would also be very relevant to examine which kind of processes are engaged by science narrative texts that bring out social and cultural elements of science, and how they can improve content learning. This would both shed light on how the learning of students from diverse backgrounds benefits from science narratives, and on how understanding the nature of science improves learning.

Finally, even though the C-I model has privileged the investigation of expository texts, the interest in using narrative texts as science education tools should spark more research on these texts based on the C-I model.

8 | RECOMMENDATIONS FOR EDUCATIONAL PRACTICES

Educators have frequently pointed out a set of challenges to science learning that can easily be connected to text, reader, activity and sociocultural aspects of learning. The language of texts is seen as dense and technical (e.g., Plavén-Sigray et al., 2017; Snow, 2010); readers find many science ideas unfamiliar or even counter-intuitive (e.g., Browning & Hohenstein, 2015; Gilbert et al., 2005); the framing discourse of learning activities is authoritative or even dogmatic (Kloser, 2013; Negrete & Lartigue, 2004); and education is generally decontextualized from human and cultural aspects (e.g., Harper & Kayumova, 2022; Sánchez Tapia et al., 2018; Solomon, 2002). The features tackled in the "How" question can therefore help address these challenges.

When selecting or creating narrative texts for science education it is important to consider the ways in which they connect readers and activities to the wider context. For instance, fiction can be used to address complex or counterintuitive ideas and may engage students as well as prompt positive emotions; however, these elements can also distract students or induce misconceptions. Educators should ensure that fictional elements are well weaved with science contents, in a coherent and contextualized way.

Historical and discovery narratives can help educators contextualize science within a more human context. These insights can help reduce damaging misconceptions about what science is and who gets to do it, building more inclusive science education practices. For example, realistic and accessible depictions of scientists can help readers from varied, often marginalized, social and cultural identities connect to science. Narrative materials can also be a means for educators to connect science learning activities with students' immediate (e.g., neighborhood) and wider (e.g., current world affairs) contexts.

Narrative texts may, however, activate conflicting goals in learners. To tackle this, educators can: make explicit the connections between a text's structure, content, and function; explain that these features are flexible; and develop study strategies using narrative materials from early-on. This, in turn, will enable students to consciously adjust their expectations and adopt a more flexible approach to comprehension (Snow, 2002). Entertainment and instructing goals should be easier to reconcile in longer interventions, which provide more opportunities to alternate between the two and to integrate information. Educators may also reinforce the connections between formal and nonformal settings, taking advantage of the lower restrictiveness of the latter to use narrative-based activities (e.g., Littrell et al., 2022; Murmann & Avraamidou, 2014).

As the latter are not always viable, when planning activities, educators should at least consider the impact that certain narrative materials might have on student's cognition. This impact can be manifold and interact with education materials and activities in important ways. To name a few examples, educators can use narrative texts as scaffolding tools with less knowledgeable and/or struggling learners, but also to promote critical thinking, reflection, and autonomy. Narrative materials can also be used to capture learners' attention, interest and emotions, and to trigger mental imagery and imagination. Yet, the extent to which and the conditions in which (e.g., what text features) such processes are engaged has still to be further determined by research, which may then further inform the design and adaptation of educational practices.

Finally, when building assessment measures, educators can benefit from distinguishing between different levels of comprehension, and the C-I model can be a useful referent.

9 | CONCLUSION

One way to address science learning challenges is by tightening science contents and literacy processes. Tackling such challenges and building more tailored practices can benefit from recognizing and embracing the multifaceted nature of learning. Narrative texts are a flexible educational tool that can help achieve such goals, as they connect the learners, the texts, the activities and the wider context in several and important ways. They have also been shown to improve the memory and learning outcomes of students at various educational levels, albeit not consistently. Together, these results suggest that learning from science narrative texts should be approached as a multitude of specific narrative effects that capture the complex interactions between the different elements of the learning process, instead of a single, overarching, narrative effect. Under this view, education and research should focus on what (texts), with whom (learners), when and where (activities and broader context) narrative materials can be used as effective science learning tools. A multidisciplinary theoretical framework combining complementary fields can thus be pivotal when developing practices and research based on this educational tool.

ACKNOWLEDGMENTS

We are very grateful to Ana Moreira for her invaluable help. This research was funded by Fundação para a Ciência e a Tecnologia (PD/BD/135345/2017 : COVID/BD/135345/2017). R. Kolinsky is Research Director of the Fonds de la Recherche Scientifique–FNRS (FRS–FNRS), Belgium.

ORCID

Sara Soares https://orcid.org/0000-0001-7668-5411
Matilde Gonçalves https://orcid.org/0000-0003-0039-4401

Rita Jerónimo https://orcid.org/0000-0003-4571-8086
Régine Kolinsky https://orcid.org/0000-0002-2492-2144

ENDNOTE

[1] In the Results section, the numbers in parentheses refer to the numbering of the studies as presented in Tables 3 and 4.

REFERENCES

Adam, J. M. (1997). Genres, textes, discours: Pour une reconception linguistique du concept de genre. *Revue Belge de Philologie et d'histoire*, *75*, 665–681.

Adam, J. M. (2011, May 20–21). *The narrative sequence: History of a concept and a research area* [Paper presentation]. *1er colloque International organisé par le Réseau Romand de Narratologie (RRN)*, Fribourg, Switzerland.

Akarsu, B., Kariper, A., & Coşkun, H. (2015). The effect of using scientific stories on teaching science and on the academic achievement of the students. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, *11*(2), 349–365. https://doi.org/10.17860/efd.045.04596

Allchin, D. (2003). Scientific myth-conceptions. *Science Education*, *87*, 329–351.

Alvermann, D. E., Hynd, C. E., & Qian, G. (1995). Effects of interactive discussion and text type on learning counterintuitive science concepts. *The Journal of Educational Research*, *88*, 146–154. https://doi.org/10.1080/00220671.1995.9941293

Arya, D. J., & Maul, A. (2012). The role of the scientific discovery narrative in middle school science education: An experimental study. *Journal of Educational Psychology*, *104*, 1022–1032. https://doi.org/10.1037/a0028108

Arya, D. J., & Maul, A. (2021). Why sociocultural context matters in the science of reading and the reading of science: Revisiting the science discovery narrative. *Reading Research Quarterly*, *56*, S273–S286. https://doi.org/10.1002/rrq.393

Avraamidou, L., & Osborne, J. (2009). The role of narrative in communicating science. *International Journal of Science Education*, *31*, 1683–1707. https://doi.org/10.1080/09500690802380695

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89. https://doi.org/10.1016/S0079-7421(08)60452-1

Bakhtin, M. M. (1984). *Esthétique de la création verbale*. Gallimard.

Banister, F., & Ryan, C. (2001). Developing science concepts through story-telling. *School Science Review*, *82*, 75–84.

Baram-Tsabari, A., & Yarden, A. (2005). Text genre as a factor in the formation of scientific literacy. *Journal of Research in Science Teaching*, *42*(4), 403–428. https://doi.org/10.1002/tea.20063

Best, R. M., Floyd, R. G., & McNamara, D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts. *Reading Psychology*, *29*(2), 137–164. https://doi.org/10.1080/02702710801963951

Bowman, N. A., Logel, C., LaCosse, J., Jarratt, L., Canning, E. A., Emerson, K. T. U., & Murphy, M. C. (2022). Gender representation and academic achievement among STEM-interested students in college STEM courses. *Journal of Research in Science Teaching*, *59*(10), 1876–1900. Portico. https://doi.org/10.1002/tea.21778

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching*, *43*(5), 485–499. https://doi.org/10.1002/tea.20131

Broemmel, A. D., & Rearden, K. T. (2006). Should teachers use the teachers' choices books in science classes? *The Reading Teacher*, *60*, 254–265. https://doi.org/10.1598/RT.60.3.5

Bronckart, J.-P. (1997). *Activité langagière, textes et discours*. Lausanne.

Browning, E., & Hohenstein, J. (2015). The use of narrative to promote primary school children's understanding of evolution. *Education*, *43*, 530–547. https://doi.org/10.1080/03004279.2013.837943

Bruner, J. (1986). *Actual minds, possible worlds*. Harvard University Press.

Bruner, J. (1990). *Acts of meaning*. Harvard University Press.

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, *18*(1), 1–21.

Bruner, J. (1996). *The culture of education*. Harvard University Press.

Cervetti, G. N., Bravo, M. A., Hiebert, E. H., Pearson, P. D., & Jaynes, C. A. (2009). Text genre and science content: Ease of reading, comprehension, and reader preference. *Reading Psychology*, *30*, 487–511. https://doi.org/10.1080/02702710902733550

Cheng, T. L., & Goodman, E. (2015). Race, ethnicity, and socioeconomic status in research on child health. *Pediatrics*, *135*(1), e225–e237. https://doi.org/10.1542/peds.2014-3109

Claus, B., & Kelter, S. (2006). Comprehending narratives containing flashbacks: Evidence for temporally organized representations. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *32*, 1031–1044. https://doi.org/10.1037/0278-7393.32.5.1031

Clough, M. P. (2011). The story behind the science: Bringing science and scientists to life in post-secondary science education. *Science & Education*, *20*, 701–717. https://doi.org/10.1007/s11191-010-9310-7

Corni, F., Giliberti, E., & Fuchs, H. U. (2014). Student teachers writing science stories: A case study. In C. P. Constantinou, N. Papadouris, & A. Hadjigeorgiou (Eds.), *E-book proceedings of the ESERA 2013 conference: Science education research for evidence-based teaching and coherence in learning* (pp. 2494–2505). European Science Education Research Association.

Corni, F., Giliberti, E., & Mariani, C. (2010, August 22–27). *A Story as innovative medium for science education in primary school* [Paper Presentation]. GIREP-ICPE-MPTL International Conference, Reims, France.

Deci, E. L., & Ryan, R. M. (1982). Curiosity and self-directed learning: The role of motivation in education. In L. Katz (Ed.), *Current topics in early childhood education* (pp. 71–85). Albex.

Dinsmore, D. L., Zoellner, B. P., Parkinson, M. M., Rossi, A. M., Monk, M. J., & Vinnachi, J. (2017). The effects of different types of text and individual differences on view complexity about genetically modified organisms. *International Journal of Science Education*, *39*(7), 791–813. https://doi.org/10.1080/09500693.2017.1298871

Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader–text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology*, *104*, 515–528. https://doi.org/10.1037/a0027182

Egan, K. (1997). *The educated mind: How cognitive tools shape our understanding*. The University of Chicago Press.

Egan, K. (2005). *An imaginative approach to teaching*. Jossey-Bass.

Erten, S., Kiray, A., & Sen-Gumus, B. (2013). Influence of scientific stories on students ideas about science and scientists. *International Journal of Education in Mathematics, Science and Technology*, *1*, 122–137.

Evangelista, P., & Zimmermann, E. (2008, October 21). *Produção de histórias cientificas infantis na formação de pedagogos para o ensino de fisica* [Paper presentation]. *XI Encontro de Pesquisa em Ensino de Física*, Curitiba, Brasil.

Evangelista, P. C. Q. (2008). *Produção de histórias infantis científicas no curso de pedagogia* [Master's thesis, Universidade de Brasília]. Open Access Theses And Dissertations.

Farah, M. J. (2017). The neuroscience of socioeconomic status: Correlates, causes, and consequences. *Neuron*, *96*(1), 56–71. https://doi.org/10.1016/j.neuron.2017.08.034

Ferstl, E. C. (2001). Learning from text. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 8605–8609). Elsevier. https://doi.org/10.1016/B0-08-043076-7/01478-9

Fisher, W. (1987). *Human communication as narration: Towards a philosophy of reason, value and action*. University of South Carolina Press.

Flynn, S., & Hardman, M. (2019). The use of interactive fiction to promote conceptual change in science. *Science & Education*, *28*, 127–152. https://doi.org/10.1007/s11191-019-00032-6

Fredricks, J. A., Hofkens, T., Wang, M. T., Mortenson, E., & Scott, P. (2018). Supporting girls' and boys' engagement in math and science learning: A mixed methods study. *Journal of Research in Science Teaching*, *55*(2), 271–298. https://doi.org/10.1002/tea.21419

Frisch, J. K. (2010). The stories they'd tell: Pre-service elementary teachers writing stories to demonstrate physical science concepts. *Journal of Science Teacher Education*, *21*, 703–722. https://doi.org/10.1007/s10972-010-9210-3

Garner, R., Gillingham, M. G., & White, C. S. (1989). Effects of "seductive details" on macroprocessing and microprocessing in adults and children. *Cognition and Instruction*, *6*, 41–57. https://doi.org/10.1207/s1532690xci0601_2

Gilbert, J., Hipkins, R., & Cooper, G. (2005, May 30–June 1). *Faction or fiction: Using narrative pedagogy in school science education* [Paper Presentation]. Redesigning pedagogy: *Research, policy, practice conference*, Nanyang, Singapore.

Glaser, M., Garsoffky, B., & Schwan, S. (2009). Narrative-based learning: Possible benefits and problems. *Communications*, *34*, 429–447. https://doi.org/10.1515/COMM.2009.026

Gomez-Zweip, S., & Straits, W. (2006). Analyzing anthropomorphisms. *Science and Children*, *44*(3), 26–29.

Gonçalves, M. (2019). Quelques réflexions autour des activités de langage, des genres textuels et de l′ implicite. In I. Seara & I. S. Marques (Eds.) *L'implicite Dans les Langues Romanes* (pp. 66–71). Universidade Aberta.

Graesser, A., Olde, B., & Klettke, B. (2002). How does the mind construct and represent stories? In M. Green, J. Strange, & T. Brock (Eds.), *Narrative impact* (pp. 229–262). Lawrence Erlbaum Associates.

Graesser, A. C., Hauft-Smith, K., Cohen, A. D., & Pyles, L. D. (1980). Advanced outlines, familiarity, and text genre on retention of prose. *The Journal of Experimental Education*, *48*, 281–290. https://doi.org/10.1080/00220973.1980.11011745

Hadzigeorgiou, Y. (2016). *Imaginative science education*. The Central Role of Imagination in Science Education.

Hadzigeorgiou, Y., Klassen, S., & Klassen, C. F. (2012). Encouraging a 'romantic understanding' of science: The effect of the Nikola tesla story. *Science & Education*, *21*, 1111–1138. https://doi.org/10.1007/s11191-011-9417-5

Harper, A., & Kayumova, S. (2022). Invisible multilingual Black and Brown girls: Raciolinguistic narratives of identity in science education. *Journal of Research in Science Teaching*. Portico. https://doi.org/10.1002/tea.21826

Hellstrand, A., & Ott, A. (1995). The utilization of fiction when teaching the theory of relativity. *Physics Education*, *30*(5), 284–286.

Heyvaert, M., Maes, B., & Onghena, P. (2013). Mixed methods research synthesis: Definition, framework, and potential. *Quality & Quantity*, *47*(2), 659–676. https://doi.org/10.1007/s11135-011-9538-6

Hidi, S., Baird, W., & Hildyard, A. (1982). That's important but is it interesting? Two factors in text processing. In A. Flammer & W. Kintsch (Eds.), *Advances in psychology: Discourse processing (Vol. 8, pp. 63–75)*. North-Holland. https://doi.org/10.1016/S0166-4115(08)62681-3

Hong, H.-Y., & Lin-Siegler, X. (2012). How learning about scientists' struggles influences students' interest and learning in physics. *Journal of Educational Psychology*, *104*, 469–484. https://doi.org/10.1037/a0026224

Jackson, M. C., Galvez, G., Landa, I., Buonora, P., & Thoman, D. B. (2016). Science that matters: The importance of a cultural connection in underrepresented students' science pursuit. CBE—Life sciences. *Education*, *15*(3), ar42. https://doi.org/10.1187/cbe.16-01-0067

Jetton, T. L. (1994). Information-driven versus story-driven: What children remember when they are reading informational stories. *Reading Psychology*, *15*, 109–130. https://doi.org/10.1080/0270271940150203

Kalogiannakis, M., & Violintzi, A. (2012). Intervention strategies for changing pre-school children's understandings about volcanoes. *Journal of Emergent Science*, *4*, 12–19.

Kane, L. (2004). Educators, learners and active learning methodologies. *International Journal of Lifelong Learning*, *23*, 275–286. https://doi.org/10.1080/0260/37042000229237

Karimi, S., Pohl, S., Scholer, F., Cavedon, L., & Zobel, J. (2010). Boolean versus ranked querying for biomedical systematic reviews. *BMC Medical Informatics and Decision Making*, *10*, 58. https://doi.org/10.1186/1472-6947-10-58

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*, 294–303. https://doi.org/10.1037//0003-066x.49.4.294

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Kintsch, W. (2012). Psychological models of reading comprehension and their implications for assessment. In J. P. Sabatini, E. R. Albro, & T. O'Reilly (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 21–38). Rowman & Littlefield Education.

Kintsch, W., & Rawson, K. A. (2005). Comprehension. In M. J. Snowling & C. Hulme (Eds.), *Blackwell handbooks of developmental psychology. The science of reading: A handbook* (pp. 209–226). Blackwell Publishing 9780470757642.ch12.

Kintsch, W., & van Dijk, T. A. (1975). Comment on se rappelle et on résume des histoires. *Langages*, *40*, 98–116.

Kintsch, W., & Young, S. R. (1984). Selective recall of decision-relevant information from texts. *Memory & Cognition*, *12*, 112–117. https://doi.org/10.3758/BF03198424

Klassen, S. (2006). A theoretical framework for contextual science teaching. *Interchange*, *37*(1), 31–62. https://doi.org/10.1007/s10780-006-8399-8

Klassen, S. (2007). The application of historical narrative in science learning: The Atlantic cable story. *Science & Education*, *16*, 335–352. https://doi.org/10.1007/s11191-006-9026-x

Klein, P. D. (2006). The challenges of scientific literacy: From the viewpoint of second-generation cognitive science. *International Journal of Science Education*, *28*, 143–178. https://doi.org/10.1080/09500690500336627

Kloser, M. (2013). Exploring high school biology students' engagement with more and less epistemologically considerate texts. *Journal of Research in Science Teaching*, *50*, 1232–1257. https://doi.org/10.1002/tea.21109

Kokkotas, P., Rizaki, A., & Malamitsa, K. (2010). Storytelling as a strategy for understanding concepts of electricity and electromagnetism. *Interchange*, *41*, 379–405. https://doi.org/10.1007/s10780-010-9137-9

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lamartino, A. (1995). *Science and reading* [Master's thesis, Kean University]. Open Access Theses and Dissertations.

Larison, K. D. (2018). Taking the scientist's perspective. *Science & Education*, *27*, 133–157. https://doi.org/10.1007/s11191-018-9957-z

Lee, O., & Grapin, S. E. (2022). The role of phenomena and problems in science and STEM education: Traditional, contemporary, and future approaches. *Journal of Research in Science Teaching*, *59*(7), 1301–1309. Portico. https://doi.org/10.1002/tea.21776

Lee, O., & Luykx, A. (2007). Science education and student diversity: Race/ethnicity, language, culture, and socioeconomic status. *Handbook of Research on Science Education*, *1*, 171–197.

Legare, C. H., Lane, J. D., & Evans, E. M. (2013). Anthropomorphizing science: How does it affect the development of evolutionary concepts? *Merrill-Palmer Quarterly*, *59*(2), 168–197. https://doi.org/10.13110/merrpalmquar1982.59.2.0168

Lehman, S., Schraw, G., McCrudden, M. T., & Hartley, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology*, *32*, 569–587. https://doi.org/10.1016/j.cedpsych.2006.07.002

Leipzig, P. (2018). *What's the story?: The effects of narratives in science classrooms* [Doctoral dissertation, Colorado State University]. Open Access Theses and Dissertations.

Linderholm, T., & van den Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology*, *94*(4), 778–784. https://doi.org/10.1037/0022-0663.94.4.778

Lin-Siegler, X., Ahn, J. N., Chen, J., Fang, F. F. A., & Luna-Lucero, M. (2016). Even Einstein struggled: Effects of learning about great scientists' struggles on high school students' motivation to learn science. *Journal of Educational Psychology*, *108*(3), 314–328. https://doi.org/10.1037/edu0000092

Littrell, M. K., Gold, A. U., Koskey, K. L. K., May, T. A., Leckey, E., & Okochi, C. (2022). Transformative experience in an informal science learning program about climate change. *Journal of Research in Science Teaching*, *59*(6), 1010–1034. Portico. https://doi.org/10.1002/tea.21750

Mar, R. A. (2004). The neuropsychology of narrative: Story comprehension, story production and their interrelation. *Neuropsychologia*, *42*, 1414–1434. https://doi.org/10.1016/j.neuropsychologia.2003.12.016

Maria, K., & Johnson, J. M. (1989, November 28 – December 2). *Correcting misconceptions : Effects of type of text* [Paper Presentation]. *Annual meeting of the national reading conference*, Austin, TX, United States of America.

Maria, K., & Junge, K. (1993, December 1– 4). *A comparison of fifth graders' comprehension and retention of scientific information using a science textbook and an informational storybook* [Paper Presentation]. *Annual meeting of the national reading conference*, Charleston, SC, United States of America.

Mason, R. A., & Just, M. A. (2009). The role of the theory-of-mind cortical network in the comprehension of narratives. *Language and Linguistics Compass*, *3*(1), 157–174. https://doi.org/10.1111/j.1749-818X.2008.00122.x

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. *Psychology of Learning and Motivation*, *51*, 297–384. https://doi.org/10.1016/S0079-7421(09)51009-2

McNamara, D. S., Ozuru, Y., & Floyd, R. G. (2011). Comprehension challenges in the fourth grade: The roles of text cohesion, text genre, and readers' prior knowledge. *International Electronic Journal of Elementary Education*, *4*(1), 229–257.

McQuiggan, S. W., Rowe, J. P., Lee, S., & Lester, J. C. (2008). Story-based learning: The impact of narrative on learning experiences and outcomes. In B. P. Woolf, E. Aïmeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent tutoring systems. ITS 2008. Lecture notes in computer science* (Vol. 5091, pp. 530–539). Springer. https://doi.org/10.1007/978-3-540-69132-7_56

Morais, C. (2015). Storytelling with chemistry and related hands-on activities: Informal learning experiences to prevent 'chemophobia' and promote young children's scientific literacy. *Journal of Chemical Education*, *92*, 58–65. https://doi.org/10.1021/ed5002416

Morais, C., & Araújo, J. L. (2019). Awakening to chemistry through storytelling and practical activities: Middle school students interacting with pre-school children. *Chemistry Education Research and Practice*, *20*(1), 302–315. https://doi.org/10.1039/C8RP00096D

Morais, J., & Kolinsky, R. (2016). Literacia científica: Leitura e produção de textos científicos. *Educar Em Revista*, *62*, 143–162.

Morais, J., & Kolinsky, R. (2020). Seeing thought: A cultural cognitive tool. *Journal of Cultural Cognitive Science*, *5*(2), 181–228. https://doi.org/10.1007/s41809-020-00059-0

Murmann, M., & Avraamidou, L. (2014). Animals, emperors, senses: Exploring a story-based learning design in a museum setting. *International Journal of Science Education, Part B*, *4*, 66–91. https://doi.org/10.1080/21548455.2012.760857

Mutonyi, H. (2016). Stories, proverbs, and anecdotes as scaffolds for learning science concepts. *Journal of Research in Science Teaching*, *53*, 943–971. https://doi.org/10.1002/tea.21255

Negrete, A. (2005). Facts via fiction stories that communicate science. In N. Sannit (Ed.), *Motivating science. Science communication from a philosophical, educational and cultural perspective* (pp. 95–102). Pantaneto Press.

Negrete, A., & Lartigue, C. (2004). Learning from education to communicate science as a good story. *Endeavour*, *28*(3), 120–124. https://doi.org/10.1016/j.endeavour.2004.07.003

Negrete, A., & Lartigue, C. (2010). The science of telling stories: Evaluating science communication via narratives (RIRC method). *Journal Media and Communication Studies*, *2*, 98–110. https://doi.org/10.5897/JMCS.9000080

Norris, S. P., Guilbert, S. M., Smith, M. L., Hakimelahi, S., & Phillips, L. M. (2005). A theoretical framework for narrative explanation in science. *Science Education*, *89*, 535–563. https://doi.org/10.1002/sce.20063

Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, *87*, 224–240. https://doi.org/10.1002/sce.10066

Oatley, K. (2016). Fiction: Simulation of social worlds. *Trends in Cognitive Sciences*, *20*, 618–628. https://doi.org/10.1016/j.tics.2016.06.002

Olson, R. (2015). *Houston, we have a narrative: Why science needs story*. University of Chicago Press.

Pearson, P. D., & Cervetti, G. N. (2015). Fifty years of reading comprehension theory and practice In P. D. Pearson & E. H. Hiebert (Eds.), *Research-Based Practices for Teaching Common Core Literacy* (pp. 1–24). Teachers College Press.

Phillips Galloway, E., McClain, J. B., & Uccelli, P. (2020). Broadening the lens on the science of reading: A multifaceted perspective on the role of academic language in text understanding. *Reading Research Quarterly*, *55*(S1), S331–S345. https://doi.org/10.1002/rrq.359

Phillips, L. M., & Norris, S. P. (2009). Bridging the gap between the language of science and the language of school science through the use of adapted primary literature. *Research in Science Education*, *39*(3), 313–319. https://doi.org/10.1007/s11165-008-9111-z

Plavén-Sigray, P., Matheson, G. J., Schiffler, B. C., & Thompson, W. H. (2017). The readability of scientific texts is decreasing over time. *eLife*, *6*, e27725. https://doi.org/10.7554/eLife.27725

Prins, R., Avraamidou, L., & Goedhart, M. (2017). Tell me a story: The use of narrative as a learning tool for natural selection. *Educational Media International*, *54*(1), 20–33. https://doi.org/10.1080/09523987.2017.1324361

Rastier, F. (2001). *Arts et sciences du texte*. Presses universitaires de France.

Reuer, M. D. (2012). *Backroads to learning: The use of narratives in high school biology* [Master's thesis, Montana State University]. Open Access Theses and Dissertations.

Ritchie, S., Rigano, D., & Duane, A. (2008). Writing an ecological mystery in class: Merging genres and learning science. *International Journal of Science Education*, *30*, 143–166. https://doi.org/10.1080/09500690601161783

Ritchie, S. M., Tomas, L., & Tones, M. (2011). Writing stories to enhance scientific literacy. International. *Journal of Science Education*, *33*, 685–707. https://doi.org/10.1080/09500691003728039

Rosa, V. (2010). *Ciência em forma de História para a compreensão dos conceitos científicos* [Doctoral dissertation, Universidade Regional de Blumenau]. Open Access Theses and Dissertations.

Rukeyser, M. (1968). *The speed of darkness*. Random House.

Sadoski, M., Goetz, E. T., Olivarez, A., Jr., Lee, S., & Roberts, N. M. (1990). Imagination in story reading: The role of imagery, verbal recall, story analysis, and processing levels. *Journal of Reading Behavior*, *22*(1), 55–70. https://doi.org/10.1080/10862969009547694

Sánchez Tapia, I., Krajcik, J., & Reiser, B. (2018). "We do not know what is the real story anymore": Curricular contextualization principles that support indigenous students in understanding natural selection. *Journal of Research in Science Teaching*, *55*(3), 348–376. https://doi.org/10.1002/tea.21422

Scells, H., & Zuccon, G. (2018). Generating better queries for systematic reviews. In Proceedings of *the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval* (pp. 475–484). ACM. https://doi.org/10.1145/3209978.3210020

Shanahan, M. C. (2012). Reading for evidence in hybrid adapted primary literature. In S. P. Norris (Ed.), *Reading for evidence and interpreting visualizations in mathematics and science education* (pp. 41–63). Sense Publishers.

Sinatra, G. M., Broughton, S. H., & Lombardi, D. O. U. G. (2014). Emotions in science education. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 415–436). Routlege.

Skydsgaard, M. A., Møller Andersen, H., & King, H. (2016). Designing museum exhibits that facilitate visitor reflection and discussion. *Museum Management and Curatorship*, *31*, 48–68. https://doi.org/10.1080/09647775.2015.1117237

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Rand Corporation.

Snow, C. (2010). Academic language and the challenge of Reading for learning about science. *Science*, *328*, 450–452. https://doi.org/10.1126/science.1182597

Solomon, J. (2002). Science stories and science texts: What can they do for our students? *Studies in Science Education*, *37*, 85–105.

Solomon, J., Duveen, J., Scott, L., & McCarthy, S. (1992). Teaching about the nature of science through history: Action research in the classroom. *Journal of Research in Science Teaching*, *29*, 409–421. https://doi.org/10.1002/tea.3660290408

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a021766. https://doi.org/10.1101/cshperspect.a021766

Strube, P. (1994). Narrative in the science curriculum. *Research in Science Education*, *24*(1), 313–321. https://doi.org/10.1007/BF02356358

Sweet, A. P., & Snow, C. E. (Eds.). (2003). *Rethinking reading comprehension*. Guilford Press.

Tomas, L., Ritchie, S. M., & Tones, M. (2011). Attitudinal impact of hybridized writing about a socioscientific issue. *Journal of Research in Science Teaching*, *48*, 878–900. https://doi.org/10.1002/tea.20431

van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science*, *328*(5977), 453–456. https://doi.org/10.1126/science.1182594

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the online construction of memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71–98). Lawrence Erlbaum Associates.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.

Visintainer, T. (2020). "I think at first glance people would not expect me to be interested in science": Exploring the racialized science experiences of high school students of color. *Journal of Research in Science Teaching*, *57*, 393–422. https://doi.org/10.1002/tea.21597

Vrasidas, C., Avraamidou, L., Theodoridou, K., Themistokleous, S., & Panaou, P. (2015). Science fiction in education: Case studies from classroom implementations. *Educational Media International*, *52*, 201–215. https://doi.org/10.1080/09523987.2015.1075102

Wang, D. (2009). Factors affecting the comprehension of global and local main idea. *Journal of College Reading and Learning*, *39*(2), 34–52. https://doi.org/10.1080/10790195.2009.10850317

Wang, S., Scells, H., Koopman, B., & Zuccon, G. (2022). Automated MeSH term suggestion for effective query formulation in systematic reviews literature search. *Intelligent Systems with Applications*, *16*, Article 200141. https://doi.org/10.1016/j.iswa.2022.200141

Walan, S., & Enochsson, A.-B. (2019). The potential of using a combination of storytelling and drama, when teaching young children science. *European Early Childhood Education Research Journal*, *27*(6), 821–836. https://doi.org/10.1080/1350293x.2019.1678923

Webb, P. (2010). Science education and literacy: Imperatives for the developed and developing world. *Science*, *328*(5977), 448–450. https://doi.org/10.1126/science.1182596 PMID: 20413487.

Wharton, C., & Kintsch, W. (1991). An overview of construction-integration model: A theory of comprehension as a foundation for a new cognitive architecture. *ACM SIGART Bulletin*, *2*(4), 169–173. https://doi.org/10.1145/122344.122379

Wilcken, W. M. (2008). *Comprehension and retention: The effect of concrete details and causal structure in scientific narrative* [Master's thesis, Brigham Young University]. Open Access Theses and Dissertations.

Wolfe, M. B., & Mienko, J. A. (2007). Learning and memory of factual context from narrative and expository text. *British Journal of Educational Psychology*, *77*, 541–564. https://doi.org/10.1348/000709906X143902

Wolfe, M. B., Schreiner, M. E., Rehder, B., Laham, D., Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). Learning from text: Matching readers and texts by latent semantic analysis. *Discourse Processes*, *25*, 309–336. https://doi.org/10.1080/01638539809545030

Wolfe, M. B., & Woodwyk, J. M. (2010). Processing and memory of information presented in narrative or expository texts. *British Journal of Educational Psychology*, *80*, 341–362. https://doi.org/10.1348/000709910X485700

Wu, Y., Barquero, L. A., Pickren, S. E., Barber, A. T., & Cutting, L. E. (2020). The relationship between cognitive skills and reading comprehension of narrative and expository texts: A longitudinal study from grade 1 to grade 4. *Learning and Individual Differences*, *80*, 101848. https://doi.org/10.1016/j.lindif.2020.101848

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *NeuroImage*, *25*, 1002–1015. https://doi.org/10.1016/j.neuroimage.2004.12.013

Yang, Y. (2003). Dimensions of socio-economic status and their relationship to mathematics and science achievement at individual and collective levels. *Scandinavian Journal of Educational Research*, *47*(1), 21–41. https://doi.org/10.1080/00313830308609

Zabrucky, K. M., & Moore, D. (1999). Influence of text genre on adults' monitoring of understanding and recall. *Educational Gerontology*, *25*, 691–710. https://doi.org/10.1080/036012799267440

Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 920–933. https://doi.org/10.1037/0278-7393.20.4.920

How to cite this article: Soares, S., Gonçalves, M., Jerónimo, R., & Kolinsky, R. (2023). Narrating science: Can it benefit science learning, and how? A theoretical review. *Journal of Research in Science Teaching*, *60*(9), 2042–2075. https://doi.org/10.1002/tea.21851