



This article is part of the topic “Visual Narrative Research: An Emerging Field in Cognitive Science,” Neil Cohn and Joseph P. Magliano (Topic Editors). For a full listing of topic papers, see [http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1756-8765/earlyview](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1756-8765/earlyview)

Editors’ Introduction and Review: Visual Narrative Research: An Emerging Field in Cognitive Science

Neil Cohn,^a Joseph P. Magliano^b

^a*Department of Communication and Cognition, Tilburg School of Humanities and Digital Sciences, Tilburg Center for Cognition and Communication, Tilburg University*

^b*Department of Learning Sciences at the College of Education & Human Development, Georgia State University*

Received 1 October 2018; received in revised form 29 September 2019; accepted 29 September 2019

Abstract

Drawn sequences of images are among our oldest records of human intelligence, appearing on cave paintings, wall carvings, and ancient pottery, and they pervade across cultures from instruction manuals to comics. They also appear prevalently as stimuli across Cognitive Science, for studies of temporal cognition, event structure, social cognition, discourse, and basic intelligence. Yet, despite this fundamental place in human expression and research on cognition, the study of visual narratives themselves has only recently gained traction in Cognitive Science. This work has suggested that visual narrative comprehension requires cultural exposure across a developmental trajectory and engages with domain-general processing mechanisms shared by visual perception, attention, event cognition, and language, among others. Here, we review the relevance of such research for the broader Cognitive Science community, and make the case for why researchers should join the scholarship of this ubiquitous but understudied aspect of human expression.

Keywords: Cognitive science; Visual narratives; Discourse; Cognition

Correspondence should be sent to Neil Cohn, Department of Communication and Cognition, Tilburg School of Humanities and Digital Sciences, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: neilcohn@visuallanguagelab.com

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. Introduction

Drawn sequences of images are all around us, from comics and picture books, to instruction manuals, storyboards, and infographics. They are among our oldest records of human intelligence, appearing on cave paintings, wall carvings, and ancient pottery, and sequential images extend across human cultures and time periods (Petersen, 2011), making them a fundamental and universal part of human expression. Despite this, the proportion of studies examining visual narratives in Cognitive Science is staggeringly small compared to the wealth of research on language and text-based narratives (Magliano, Higgs, & Clinton, 2019; Magliano, Loschky, Clinton, & Larson, 2013). Here, we examine why Cognitive Science may have overlooked this fundamental aspect of human meaning-making, how visual narratives have typically appeared in studies of cognition, and why researchers should turn toward them with greater focus. We provide a framework that describes the underlying structures involved in drawn sequential narratives. In discussing this framework, we hope to pose interesting questions that engender interest in the study of psychology of visual narratives.

1.1. What are visual narratives?

Visual narratives are sequences of images created with meaningful intent, typically illustrating a continuous event sequence, particularly to tell a story. They may be considered a subset of meaningful sequences of images, which in less narrative forms also include instruction manuals and signage. Drawn visual narratives typically appear in comics, picture stories, and storyboards, and have many historical precedents (Petersen, 2011). Dynamic visual narratives appear as film, which uses actual percepts instead of drawings (animation aside), though they are often preceded in production by drawn narratives in the form of storyboards.

Why might visual narratives be interesting to those who study the mind? Consider first Fig. 1, a wordless comic from *JA!* by Ángela Cuéllar and Jonás Aguilar. In the first panel of this sequence, a man and his pets, a dog and a cat, enter a pet shop. In panel 2, the cat looks at a fish in a bowl, ominously peering into it in panel 3. In the fourth panel, the man then wonders what happened to his dog, only in panel 5 for the dog to be revealed as stuffed into the fish bowl (presumably by the cat).

We will analyze this sequence in more detail below, but let us first point out several important aspects about how a mind might construe the meaning described above. First, a comprehender must recognize that the lines and shapes convey meanings, despite being drawings with a high degree of abstraction unlike natural percepts. Second, within these drawings, some elements do not look like natural percepts, and require conventional knowledge, such as the balloon indicating speech (which contains no text). Third, a comprehender must recognize that there are five distinct sub-images (panels) within this broader spatial array, and that these segmented units must be viewed in a particular order. Fourth, understanding of these sequential images requires linking the meaning across images—knowing that the elements in one image (such as a cat) are the same entities in other images, and that there are changes in

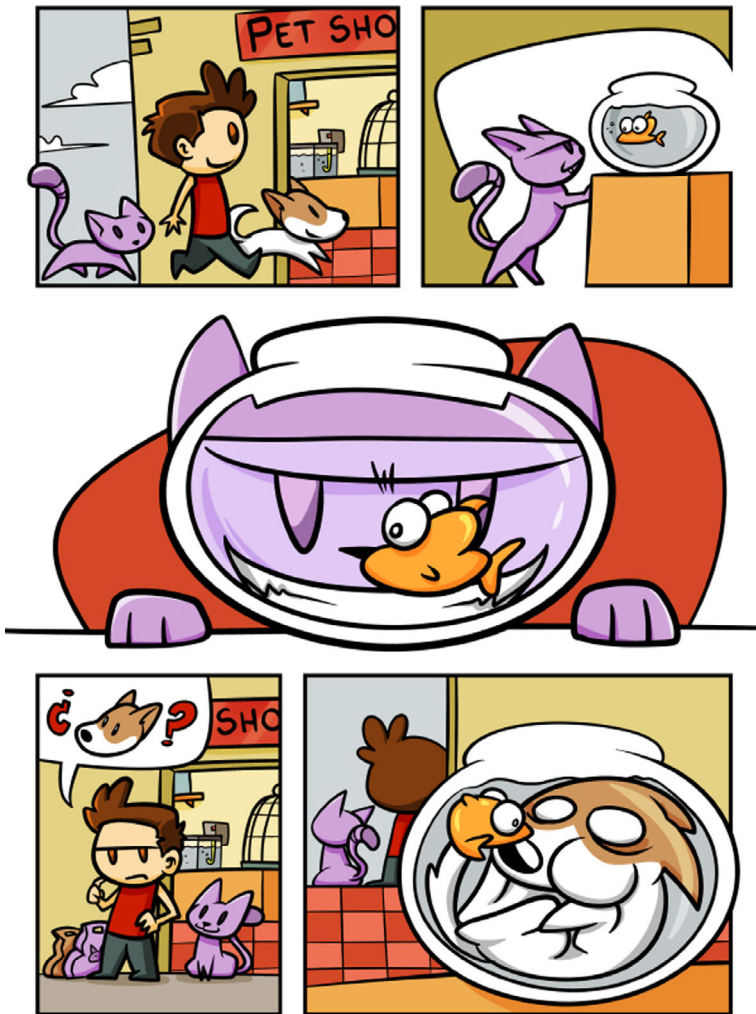


Fig. 1. An example of visual narrative from *JA!* by Ángela Cuéllar and Jonás Aguilar (© 2016; <https://revista-exegesis.com/2016/01/ja/>).

states between them (i.e., “referential continuity”), along with relations in narrative time, space, and causality (e.g., Zwaan & Radvansky, 1998). Fifth, in this sequential understanding, we must recognize that, although we see some events, there are other meanings that we do not see, and that we must infer, such as the cat stuffing the dog into the fishbowl (Graesser, Singer, & Trabasso, 1994). Sixth, we understand that characters have intentions that may motivate their actions (Trabasso, van den Broek, & Suh, 1989), and our expectations might be confounded for narrative effect (such as the possible expectation that the cat wants to eat the fish at panel 3, only to turn out that it wants to stuff the dog into the fishbowl). Seventh, the presentation of this sequential meaning follows particular choices for when we see the things that we do, in what pacing, and with what framing.

None of these aspects of comprehension are trivial, and they relate to many established structures already studied in Cognitive Science: object and scene perception, spatial cognition, event cognition, sequence processing, theory of mind, inferencing, discourse processing, and many other fundamental aspects of cognition. If this example was a block of written language, the importance of studying its cognitive underpinnings would be taken for granted, and that doing so would be fundamental for assessing the mind more generally. Yet, for some reason as a drawn narrative it does not intrinsically carry such priorities. The question is: why?

1.2. *Why aren't visual narratives studied?*

Given the above, why might visual narratives not be studied with the type of seriousness afforded other types of human expression and communication? We see several potential factors involved. First, sequential images are frequently used in tasks and stimuli to investigate *other* aspects of cognition (e.g., Baron-Cohen, Leslie, & Frith, 1986; Boroditsky, Gaby, & Levinson, 2008); this means researchers are certainly aware of visual narratives as potentially effective communicative tools, but they just have not consolidated around studying them in and of themselves.

Second, the use of visual narratives as stimuli often presumes that their understanding is fairly transparent. This may relate to phenomenologically based beliefs that drawings are not internally complex or simply map directly into spatial or event knowledge, despite evidence to the contrary (Cohn, 2012; Willats, 2005; Wilson, 2016). While there may be cross-domain processes that support comprehension across narrative experiences (Magliano et al., 2013), the conventions of the visual modality likely lead to non-trivial dimensions for how these mechanisms operate and the awareness of them (Cohn, 2013b; Magliano et al., 2019).

Third, though attitudes have been changing in recent years, visual narratives have historically been afforded low esteem in culture, in the form of comics or illustrated picture books, for example, compared to language or film. Drawn visual narratives have often been associated with younger readers, again, likely tied to their presumed transparency. Nevertheless, while attitudes have been changing over the past decades in the United States, children have never been the sole audience for visual narratives (Duncan, Smith, & Levitz, 2015), and such stereotypes are not as strong in other countries (Schodt, 1983). If something is perceived as simplistic, with low cultural value, why should cognitive scientists study them seriously? Below, we address these myths of the simplicity of visual narrative comprehension. We then detail why addressing them is important for the use of sequential images in experimentation, and finally make the case for the benefits of studying visual narratives in Cognitive Science.

1.3. *Myths of visual narrative*

We contend that at least three myths about visual narratives have contributed to the lack of a formal field of study on them. The first is the *myth of transparency*. There are many assumptions about the universal transparency of sequential image comprehension.

Sequential images are thought to be understandable by everyone—including young children—with little learning or decoding (e.g., McCloud, 1993). This *assumed transparency* likely arises from beliefs that they require basic perceptual or event cognition alone. The logic follows something like this: Sequential images visually depict objects and events, and basic perceptual processing allows everyone to understand events (Gibson, 2014); thus, everyone should understand sequential images. This thinking has likely motivated the frequent use of sequential images in psychological experiments, IQ tests and clinical assessments, and educational, humanitarian, and/or anthropological research, especially with non-literate and native populations (see below).

Despite these presumptions, sequential image understanding requires exposure and expertise, as suggested by a scattered literature of cross-cultural, developmental, and cognitive research. First, various cross-cultural studies have found that individuals from indigenous communities have difficulty construing even basic aspects of sequential images. These responses often manifest at a basic level, where individuals do not construe that the characters in one image are the same as those in subsequent images—that is, referential continuity. Such findings appear with individuals in Nepal (Fussell & Haaland, 1978), Papua New Guinea (Bishop, 1977; Cook, 1980), and various populations in Africa (Byram & Garforth, 1980; Duncan, Gourlay, & Hudson, 1973; Liddell, 1996), among others. The consistent trend in these studies is that sequential images were not understood *as a sequence*, but rather were interpreted as a series of isolated images each depicting their own scene, with no binding continuity. Higher level deficits (i.e., inference processes) have also been observed by rural populations in Turkey for film (Ildirar & Schwan, 2015). These individuals who do not construe sequential information typically come from rural communities. They also have minimal exposure to visual narratives (comics, illustrated books, films), which often correlates with low literacy or education (Cook, 1980; Le Guen & Pool Balam, 2012). Overall, though, the findings are consistent: *Sequential image understanding requires exposure and practice with visual narratives* (for review, see Cohn, in press).

Second, developmental research has suggested a trajectory for when children begin to comprehend a sequence of images as a sequence. This work has suggested that children at or below the age of 4 have difficulty recognizing the referential continuity of repeated characters as indexing the same entities (Bornens, 1990; Trabasso & Nickels, 1992). Recognition of this continuity begins between 4 and 5, typically with full understanding around age 6 (Bornens, 1990). This age range is also when children begin to become more proficient at picture arrangement tasks, where they organize randomly ordered images into a coherent sequence (Fivush & Mandler, 1985; Friedman, 1990; Weist, Lyytinen, Wysocka, & Atanassova, 1997), and when they are able to describe a coherent sequence of narrative events from sequenced images (Trabasso & Nickels, 1992). Also, around age 5 they begin to recognize and infer omitted content from sequences (Brown & French, 1976; Kunen, Chabaud, & Dean, 1987; Schmidt & Paris, 1978), though both picture arrangement and inference continue developing into later years, and are modulated both by age and by experience with comics (Nakazawa, 2016). Altogether, this research suggests that *children develop an understanding of image sequences across a*

developmental trajectory conditioned by their exposure to visual narratives (for review, see Cohn, in press).

While lack of comprehension due to exposure may seem like an extreme case, effects of expertise also appear between “fluent” comic readers (Lee & Armour, 2016; Zhao & Mahrt, 2018). Metrics of self-assessed comic reading/drawing proficiency have been shown to be consistent predictors of individual differences in studies of visual narratives assessed by reaction times (Cohn, Paczynski, Jackendoff, Holcomb, & Kuperberg, 2012), self-paced viewing times (Cohn & Wittenberg, 2015), segmentation choices (Cohn & Bender, 2017), comprehension ratings (Cohn, Murthy, & Foulsham, 2016; Cohn & Wittenberg, 2015), accuracy judgements (Hagmann & Cohn, 2016), and amplitude differences of brainwaves (Cohn & Kutas, 2015, 2017; Cohn et al., 2012). In addition, processing may not be modulated just by general expertise with visual narratives, but also by the patterns found in *specific* visual narrative systems. Recent analysis of the neurocognitive processes of a particular sequential pattern was found to be modulated by frequency of reading comics that prevalently use that pattern (Cohn & Kutas, 2017). Thus, even among experienced readers of visual narratives, expertise may modulate processing.

Overall, this work suggests that sequential image comprehension requires a “fluency” that is acquired from exposure and practice with visual narratives across a developmental trajectory. These findings go against the *myth of transparency* that sequential image understanding comes “for free” with perception and/or event cognition. It is also noteworthy that the works documenting these findings belong to scattered disciplines, with little integration, much less placed in the context of the broader study of the mind. If we want to truly understand these forms—and their connection to other aspects of cognition like language, spatial cognition, or event understanding—they must find a place within Cognitive Science.

The second myth is that of *universal comprehension processes*. Some have argued that many of the cognitive processes that support meaning-making in text should also support the processing of visual narratives (Baggett, 1979; Gernsbacher, 1990; Magliano et al., 2013). As such, there may not be a perceived need to study the psychology of visual narratives, given a rich history of research on narrative discourse (see McNamara & Magliano, 2009 for an extensive review). However, comprehension proficiency between text and visual narratives is weakly correlated in children (Pezdek, Lehrer, & Simon, 1984). Moreover, only a few studies have directly explored the extent to which processes are common across textual, filmic, or drawn narratives (Baggett, 1979; Coderre et al., 2018; Magliano, Kopp, McNeerney, Radvansky, & Zacks, 2012; Robertson, 2000; West, 1998). Finally, there are reasons to also assume that visual narratives require unique processes (Cohn, 2013b; Magliano et al., 2019).

Consider bridging inferences, of which theories of comprehension universally assume are important for comprehension (McNamara & Magliano, 2009). Bridging inferences establish how two or more narrative events are connected. They involve a range of inferences, from anaphor resolution to causal processing (i.e., establishing that two narrative events are causally connected). A few recent studies have demonstrated that bridging inference is important for the comprehension of visual narratives (Cohn & Kutas, 2015;

Cohn & Wittenberg, 2015; Hutson, Smith, Magliano, & Loschky, 2017; Magliano, Kopp, Higgs, & Rapp, 2017; Magliano, Larson, Higgs, & Loschky, 2015), which is consistent with the assumption of universal comprehension processes. However, bridging inferences in visual narratives involve attentional selection and visual search of images—aspects of scene perception that cannot occur in text (Hutson et al., 2017). While there may be universal cognitive processes (e.g., bridging inferences) across narrative media, modality-specific processes may support that meaning making (Cohn, 2013b; Loughlin, Grossnickle, Dinsmore, & Alexander, 2015). As such, strong adherence to the myth of universal cognitive comprehension processes has potentially prevented important insights into how these processes are achieved across media. The time is ripe for a systematic study of the psychology of visual narratives.

A final myth is that visual narratives are *neutral experimental materials*, which is arguably a consequence of the first two myths. Because of the presumptions of transparency, experiments often rely on drawings and sequential images as experimental stimuli under the assumption that they will require no expertise to decode. Researchers in psychology use sequential images in experimental tasks to study event cognition (Tinaz, Schendan, Schon, & Stern, 2006), temporal cognition (Boroditsky et al., 2008), discourse (Gernsbacher, 1985), theory of mind (Baron-Cohen et al., 1986; Sivaratnam, Cornish, Gray, Howlin, & Rinehart, 2012), and social intelligence (Campbell & McCord, 1996), not to mention as elicitation tools for studying language (Berman & Slobin, 1994; San Roque et al., 2012), among others. They have also become a staple of general intelligence (IQ) tests (WAIS-IQ, WISC), and clinical assessments (Kaufman & Lichtenberger, 2006; Ramos & Die, 1986).

Consider also the widespread use of visual narratives as stimuli in developmental research, as in tasks for Theory of Mind (Baron-Cohen et al., 1986; Sivaratnam et al., 2012). Sequential images are often used to study the developmental trajectory of ToM in children because they are perceived as easily controllable and transparent (and appealing) to children. However, as discussed above, the age of onset for sequential image comprehension appears to be around 4–5 years old. Thus, assessing the development of ToM may be confounded by the *concurrent development* of “visual narrative fluency,” which is not measured. This confound may be a challenge to numerous domains that use visual narratives, including temporal cognition (Ingber & Eden, 2011; Weist, 2009), narrative cognition (Burris & Brown, 2014), and sequential reasoning (Zampini et al., 2017).

Another prevalent use of visual narratives comes in the *Picture Arrangement Task* (PAT), where participants arrange unordered images into a sequence, which is then scored against an expected, target sequence. The PAT has long been a key part of general intelligence (IQ) tests (WAIS-IQ, WISC) and clinical assessments (Kaufman & Lichtenberger, 2006; Ramos & Die, 1986) like brain damage (Breiger, 1956; Huber & Gleber, 1982) and other clinical diagnoses (Beatty, Jovic, & Monson, 1993; Beatty & Monson, 1994). Nevertheless, questions persist about what these tasks index (Lipsitz, Dworkin, & Erlenmeyer-Kimling, 1993; Ramos & Die, 1986; Tulskey & Price, 2003), perhaps because such studies *never include measures of visual narrative reading experience*, despite robust findings that PAT proficiency differs across age and experience with visual narratives

(Fivush & Mandler, 1985; Friedman, 1990; Nakazawa, 2016; Weist et al., 1997). In addition, the PAT would deem any “unexpected” sequence order as incorrect, even though reordering can yield multiple well-formed sequences when accounting for the structural constraints of visual narratives (Cohn, 2014).

Thus, despite the use of visual narratives as stimuli and in tasks across Cognitive Science, rarely do these studies account for aspects of the participants (expertise) and stimuli (stimulus structure). If sequential images are to be used in such experimental contexts, then it behooves us to examine just how they are understood, and how that understanding interacts with language and other cognitive systems.

2. Aspects of visual narratives

A growing number of studies have begun to examine the actual properties of sequential image understanding (Cohn, 2013b; Loschky, Magliano, Larson, & Smith, 2019; Nakazawa, 2016). This work has pointed to a complex structure that affords questions about visual narrative processing as a unique focus of scientific inquiry, and connected to fundamental issues to the broader study of cognition.

2.1. *Sequential image structure is complex*

Despite the stereotype that visual narratives are simple, they actually involve numerous levels of structure which interface with commonly studied aspects of cognition. A model of visual narrative structure is provided in Fig. 1, which divides visual narrative systems across a modality of expression (graphics), organized using combinatorial systems (grammar), which express conceptual information (meaning), and does so across both the unit and sequence levels. While these structures can be described abstractly, here we return to our example from Fig. 1 across its component structures. We will start by discussing aspects of the modality at both the unit (graphic structure) and sequence (layout) levels. Then we will discuss the combinatorial systems that organize form and meaning for both units (morphological structure) and sequences (narrative structure), before addressing aspects of the meaning itself with conceptual/event structures. Finally, we discuss the connection of visual narratives with text in multimodal relationships. Each section concludes with suggestions for possible, future research questions (Fig. 2).

2.1.1. *Graphic structure*

At the most surface level, comprehending a visual sequence requires decoding the **graphic structure**—the system organizing the physical manifestation of the graphics of the representation. Such a level would be analogous to phonological structure in spoken or signed languages, organizing the modality of expression itself. The systematized aspects of graphics—lines, shapes, junctions—map to the meanings expressed in a visual image. To understand Fig. 1, a reader must understand that various shapes correspond to expressed meanings, both referential entities (a man, cat, dog, etc.), locations (inside and outside of a

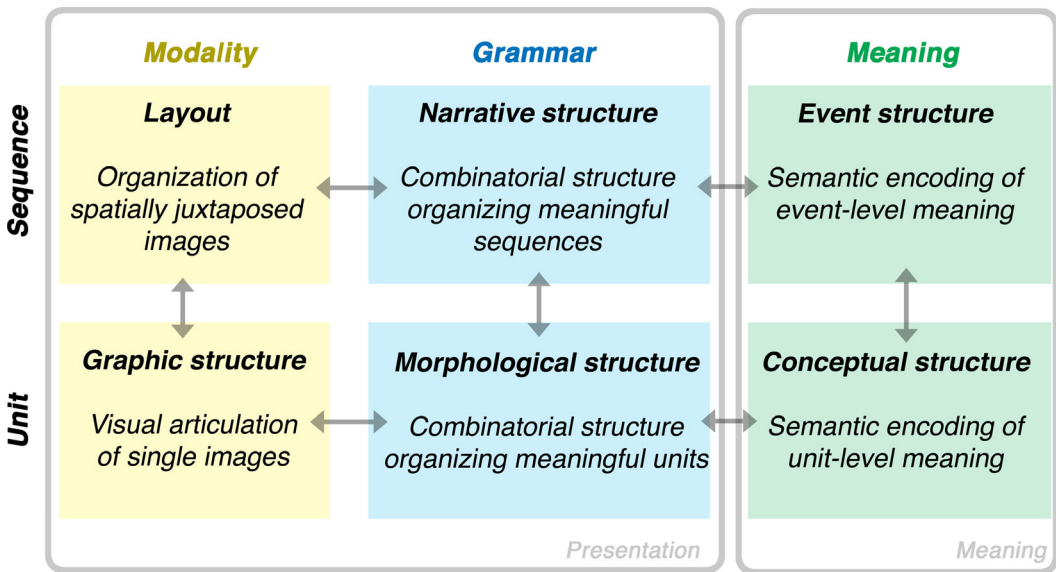


Fig. 2. A model of the structures involved in visual narratives.

pet store), and events (walking, looking, questioning). Processing this information is supported by perceptual and attentional selection (Loschky et al., 2019; Magliano et al., 2013).

Yet, as discussed above, while basic perceptual processes operate in visual narrative understanding, they must interact with specialized knowledge that is unique to the domain, and may be culturally specific (Cohn, 2013b). Graphic schemas entrenched in long-term memory can be considered the “visual vocabulary” of drawn representations, and range in size from the small subparts of images like hands, eyes, and head-shape, to larger composites like whole figures and even scenes (Cohn, 2013b). Learned patterns also exist for abstract relations of “morphology” like speech balloons and thought bubbles, or the motion lines that illustrate the paths of moving objects (discussed further below). All of these patterned representations must be stored in the minds of artists who create them, and, while iconicity may enable broader comprehension, frequent exposure may in turn be encoded preferentially for comprehenders of particular visual lexicons (Nakazawa, 2016).

While understanding these iconic images, in a range of styles (Kendall, Raffaelli, Kingstone, & Todd, 2016), may seem simple to people surrounded by graphic images, understanding and producing graphic representations is fairly complex. First, with regard to production, increasing evidence points to drawing development being guided by the imitation of graphic schemas (Huntsinger, Jose, Krieg, & Luo, 2011; Okada & Ishibashi, 2017; Wilson, 1988). If learners seek to acquire a visual vocabulary of graphic schemas, it would explain why learners in many cultures appear to reach the apex of a critical learning period for drawing around puberty, while no such apex is apparent in cultures

with exposure and practice with rich visual vocabularies (Cohn, 2012; Wilson, 2016). Thus, people who “can’t draw” may not have sufficiently learned a visual vocabulary.

Second, in terms of comprehension, cross-cultural findings have suggested that drawings of different styles are not always comprehended in the ways conforming to Western expectations (Fussell & Haaland, 1978; Wilkins, 1997/2016). For example, Wilkins (1997/2016) recounts showing Australian Aboriginals a drawing of what we would likely construe as a horse at a lateral side-view, running with puffs of smoke kicked up around its hooves. However, many of his Aboriginal respondents reported that the horse was lying down or dead, because their indigenous visual narrative system maintains a fixed aerial viewpoint. With a fixed aerial view, a “lateral view” of a horse was viewed as if looking down on it, and thus it was lying down or dead. Such findings again imply a non-trivial role of exposure and expertise to the understanding of graphic systems.

Finally, recent advances in machine learning have increasingly been able to recognize elements from visual narratives like faces, objects, textures, balloons, and other visual features (Nguyen, Rigaud, & Burie, 2017; Rigaud, Guérin, Karatzas, Burie, & Ogier, 2015; Saito et al., 2015). While these efforts show the growing effectiveness of machine learning for recognizing targeted aspects of visual narratives, such methods often underperform both human testers and computational methods trained on naturalistic photographs (Khetarpal & Jain, 2016; Takayama, Johan, & Nishita, 2012). However, significant advances are being made with such computational methods, which are rapidly becoming more reliable (see Laubrock & Dunst, 2019).

Such results raise various questions for numerous aspects of Cognitive Science. How do people learn to draw? What is the role of exposure to particular graphic systems on comprehension? How do drawn percepts differ from naturalistic percepts? How might models of computer vision adjust to the complexity of visual narrative representations?

2.1.2. *Layout*

The physical properties of images also extend beyond the internal relations in a unit, but also to the *layout* of juxtaposed images, be it spatial or temporal. In a spatial layout, a comprehender must navigate across juxtaposed images in order to comprehend the content. Layouts have a range of complexity. In illustrated picture books, single images are often placed on each page or page-spread. Fig. 1 uses a fairly simple layout, a left-to-right and down “Z-path” that mimics the reading order of Western writing systems (Cohn, 2013a). Other comics use more complex and varying features of layout (Bateman, Veloso, Wildfeuer, Cheung, & Guo, 2016; Cohn, 2013a), involving vertical columns, staggering of panels, inset panels placed inside of other panels, and other decorative features. Negotiating the reading-path across these features involves constraints that go beyond those in writing systems (Bateman, Beckmann, & Varela, 2018; Cohn, 2013a), and are certainly not learned through interactions in the real world.

Some work in the study of temporal cognition has speculated that the layout of narrative images can reveal biases about time-space metaphors (Boroditsky et al., 2008). However, this conflates the understanding of layout and the content of images, as well as disregards that visual narratives have unique constraints guiding layouts, independent of

writing and content (Cohn, 2013a). Though there are relationships between the layout of visual narratives and writing systems, the constraints operating on visual narrative layouts may be among the more modality-specific features of visual narratives.

The study of layout opens a host of interesting questions across fields: What are the constraints on navigating page layouts, and how do they differ across cultures, genres, and so on? What are the principles by which layouts connect to their content? What is the relationship of navigating a visual narrative layout and writing systems? To what extent does navigation of visual narrative layouts rely on experience with visual narratives in general, or specific types (such as culture-specific comic conventions)? How does the structure and navigation of layout in visual narratives relate to that in other media?

2.1.3. *Morphological structure*

In language research, morphemes refer to the smallest units of meaning instantiated in a form, and morphological structure refers to the combinatorial properties governing the units of a language. Similarly, visual narrative systems have been shown to have repeated, combinatorial features similar to the morphological knowledge that supports language (Cohn, 2012, 2013b; Forceville, 2011; Wilson & Wilson, 1977). Such a *visual morphology* constitutes both the schemas involved in basic, iconic drawings, which render images that typically appear like basic percepts and scenes, and more symbolic representations like speech balloons, motion lines, or lightbulbs above heads. These latter visual morphemes depart from perceptual resemblance, using combinatorial structures to attach conventionalized visual signs to other “stems.” For example, in Fig. 1 the fourth panel shows the man “speaking” with a speech balloon (here albeit with imagistic contents). This “carrier” attaches to the stem of the speaker, and forms like these have been likened to bound morphemes—affixes—in verbal languages (Cohn, 2013b; Forceville, 2011). Similar affixation occurs with signs like hearts or stars floating above characters’ heads, motion lines attaching to moving objects, or substitutions of hearts or stars for characters’ eyes (Cohn, 2013b; Forceville, 2011). In many cases, this affixation uses hierarchic embedding: Gears may float above a head to show thinking, but so too might motion lines surround the gears to show they are spinning. Thus, an affix (motion lines) attaches to an affix (gears) which attaches to a stem of a face.

All of these elements use combinatorial principles that are highly constrained and conventionalized, and their understanding is modulated by comic reading expertise (Cohn et al., 2016; Forceville, 2011; Newton, 1985). Often, these conventions are not recognizable without knowing their culture-specific origins, such as nosebleeds depicting lust or a bubble out of the nose as indicating sleep, which come from Japanese manga (Cohn, 2013b), and now subsequently appear in the emoji used in messaging applications. While many visual morphemes use purely conventional meanings, others draw from metaphoric origins (Cohn, 2013b; Forceville, 2011; Szawerna, 2017). For example, gears spinning above a head evoke the idea of the mind as a machine, while steam out of the ears evokes the head as an overflowing pressurized container (Forceville, 2005).

To what extent is visual morphology constrained by spatial and conceptual relationships with other elements (e.g., faces), and to what extent does context modulate their

understanding? Is their comprehension modulated by frequency of exposure, and/or culture-specific familiarity? What are the combinatorial mechanisms that allow us to create composite meanings in a single visual representation, and how quickly is the understanding of new visual signs acquired? To what extent do comprehenders need to compute the metaphors that many of these forms evoke, or are they entrenched in memory? Do cultures use conceptual metaphors in visual morphology that are consistent with metaphors in their languages?

2.1.4. Narrative structure

Though we extract meaning from a visual narrative, as will be discussed below, that meaning is organized by patterns in the sequence itself—its *narrative structure*. For example, Fig. 1 follows a fairly standard narrative arc, with the opening panels establishing the broader context and initiating the primary events, before resolving the sequence at the final image. Interestingly, this sequence lacks a depicted climax, the events of which are instead left inferred (i.e., the dog being stuffed into the fishbowl). Inferences like this are interesting for mental model construction, as discussed below, yet they also arise with intent: The author *chose* to not show the key information of the sequence. This means that such an omission has a structural purpose, not arising merely by happenstance, and such structure can thus be characterized in addition to the meaning-making.

This narrative structure provides constraints on the ordering and organization of information in a visual discourse. The schematic properties of such narrative structures may be modality independent, but growing evidence suggests that drawn narratives may use patterning unique to the medium (Cohn, 2013b, 2013c; Cohn & Kutas, 2017). Because it provides a combinatorial system that mediates the expression of meaning in visual sequences, recent work has argued that the narrative structure of sequential images follows constraints similar to that of a grammatical structure in language (Cohn, 2013c). Functionally, like syntax in language, this narrative structure in visual narratives can be viewed as processing instructions to the cognitive systems responsible for building situation models (Christiansen & Chater, 2016; Givón, 1993). Structurally, also like syntax, narrative imbues units (images) with categorical roles in relation to a sequencing schema, which organizes them into a recursive, hierarchic structure that can be further altered using modifiers (Cohn, 2013b, 2013c). Such a “narrative grammar” expands on earlier discourse theories of “story grammars” (Mandler & Johnson, 1977) which were characterized by older Chomskyan models of syntax (Chomsky, 1965), with ambiguous relations between syntax and semantics (Black & Wilensky, 1979). This newer approach is consistent with contemporary models of construction grammar that motivates sequencing through stored schemas rather than procedural rules (Culicover & Jackendoff, 2005; Goldberg, 1995), with a clear separation between narrative and semantic processing (Cohn, Jackendoff, Holcomb, & Kuperberg, 2014; Cohn et al., 2012). Growing behavioral and neurocognitive research has supported the basic constructs of this narrative grammar, like the presence of narrative categories (Cohn, 2014; Magliano et al., 2017), constituent structure (Cohn & Bender, 2017; Cohn et al., 2014), and entrenched narrative patterns beyond just the canonical arc (Cohn & Kutas, 2017).

Experimental research measuring the electrophysiology of the brain has bolstered the claims of similarities between the narrative structure operating in sequential images and the syntactic structure in sentence processing (for review see Cohn, 2019b). Manipulation of the narrative grammar evokes neural responses consistent with combinatorial structure and/or structural predictions (anterior negativities, often left lateralized) and with processes associated with syntactic revision (P600s) (Cohn et al., 2014; Cohn & Kutas, 2015, 2017). These comparisons suggest similarities in the processing both across domains (visual images, speech/text), and across levels of information structure (sentence level, narrative/discourse level), perhaps pointing toward more abstract aspects of sequence processing in general (Christiansen, Conway, & Onnis, 2011; Patel, 2003).

Recent efforts using computational methods have also begun analyzing visual narrative sequencing (see Laubrock & Dunst, 2019). Some marginal results have been found by convolutional neural networks in recognizing sequencing structure in four-panel comic strips (Ueno & Isahara, 2017; Ueno, Mori, Suenaga, & Isahara, 2016). More extensive analyses have attempted to use deep neural architectures to predict aspects of narrative and inference in a corpus of over a million panels from almost 4,000 classic American superhero comics (Iyyer et al., 2017). In all cases, the computational methods greatly underperformed human assessments, again suggesting that visual narrative structure—even in short four-panel strips—has complexity beyond what computational models currently allow.

This research overall motivates us to ask, what are the representations and structures that guide visual narrative sequencing? To what extent does visual narrative share structural principles and processing mechanisms with those found in other domains (e.g., sentence processing)? Do narrative schemas manifest in similar ways across modalities? How does narrative structure interact with meaning-making and situation model construction?

2.1.5. Conceptual/Event structures

As an expressive and communicative modality, visual narratives convey meanings, both within their units and across sequences. Individual images convey referential information about entities and locations, along with basic information about the events they undertake (i.e., their actions and experiences). These events extend further across sequences, where they are connected into broader event structures. Similarly, narrative text convey events within and across sentences and as such, readers need to compute the meaning of sentences and establish their semantic relationship to the prior discourse (Graesser, Millis, & Zwaan, 1997). Understanding narrative, whether verbal or visual, requires one to build situation models that reflect an underlying structure of the depicted events which unfold in narrative time and space (van Dijk & Kintsch, 1983; Gernsbacher, 1990; Zwaan & Radvansky, 1998).

The literature on narrative text processing has stressed the construction of a situation model of the growing understanding of a narrative episode, that is, characters performing goal-directed actions in a narrative time and space (Zwaan & Radvansky, 1998). Situation models represent the agents involved in the narrative, the events of the narrative, the

spatial temporal locations where these events occur, and how these events are related, in space, time, and causality (Zwaan & Radvansky, 1998). Building situation models requires that understanders monitor continuity in these dimensions, and update their representations when discontinuities are experienced (Gernsbacher, 1990; Zwaan & Radvansky, 1998). As a reader progresses through a visual sequence, changes in dimensions of characters, spatial location, and events trigger the need to update a situation model with new information (Cohn & Kutas, 2015; Loschky et al., 2019; Magliano et al., 2012; Radvansky & Zacks, 2014; Zwaan & Radvansky, 1998).

Certainly, perceptual changes across drawn sequential images play a role in signaling that updating is required (Loschky et al., 2019; Magliano et al., 2012; Radvansky & Zacks, 2014), but remarkably little research has explored how this might be the case (Loschky et al., 2019). Yet, as discussed above, even connecting basic referential information across images requires a “fluency” in sequential image processing (Byram & Garforth, 1980; H. F. Duncan et al., 1973; Ildirar & Schwan, 2015; Liddell, 1996), meaning that perceptual processing alone cannot account for tracking information across images.

Visual narratives also involve meaning-making processes beyond monitoring continuities in situational dimensions (Magliano et al., 2013), such as various types of inferencing (McNamara & Magliano, 2009) and conceptual metaphors (Forceville, 2016). Consider the bridging inference used in Fig. 1 to understand the primary events of the sequence. In panels 2 and 3, the cat peers into a fishbowl, potentially implying that it wants to eat the fish. Yet, in panel 5, we see the dog stuffed into the fishbowl. The primary event is not depicted—the inferred actions of the cat stuffing the dog into the bowl, thereby changing our conception of why the cat looked in the bowl in the first place (with ill intent toward the dog, not the fish). Understanding this comic arguably requires inference for how the dog got into the fishbowl, so that coherence can be established between the final and preceding panels (Cohn & Kutas, 2015; Hutson, Magliano, & Loschky, 2018; Loughlin et al., 2015; Magliano et al., 2017).

As discussed above, inferences like this require that a reader reconcile sequencing that otherwise makes little sense in its surface form. As in the study of discourse (McNamara & Magliano, 2009), inference has been a primary focus in visual narratives since early in their study (McCloud, 1993). Readers are indeed sensitive to omitted information, and such costs are evident in both behavioral and neurocognitive measures (Cohn & Kutas, 2015; Cohn & Wittenberg, 2015; Magliano et al., 2017; Magliano et al., 2015). Such work again suggests connections across modalities, as similar brain responses are implicated to inferences in visual as in verbal narratives (Cohn & Kutas, 2015, 2017), as are similar memory systems (Magliano et al., 2015). Nevertheless, modality-specific processes may also operate while facilitating inference in visual narratives. For example, Hutson et al. (2018) showed that, while constructing bridging inferences in visual narratives, readers engage in attentional selection not afforded by other narrative media, such as text or film. These findings raise questions about the domain-specificity and generality of mental model construction in visual and verbal modalities (Cohn, 2013b, 2019; Magliano et al., 2019; Magliano et al., 2013). To what extent does visual narrative use modality-specific processing, and to what extent are they shared with other modalities?

2.1.6. Multimodality

Finally, in addition to structures of sequential images in isolation, they also unite with other domains in multimodal interactions. While considerable research has examined multimedia expository information (Mayer, 2009), surprisingly little cognitive research has been done on multimodal visual narratives. In film, both the visual and linguistic (and musical) streams of information support situation model construction (Magliano, Dijkstra, & Zwaan, 1996), while static, drawn visual narratives typically integrate text and images. This raises interesting interactions between multiple streams of information that construct a holistic meaning (Cohn, 2016).

On the front end of processing (extracting information from and directing attention to panels/pictures), comprehending such interactions between text and images thus places demands on the attentional and perceptual systems (Magliano et al., 2013; Mayer, 2009), which can affect how readers negotiate between these information sources (Kirtley, Murray, Vaughan, & Tatler, 2018; Laubrock, Hohenstein, & Kümmerer, 2018). On the back end, further cognitive mechanisms must integrate the meaning and narrative structures of (at least) two streams of information (Cohn, 2016; Magliano et al., 2013; Mayer, 2009). This opens up questions about how much and at what level each modality is processed separately or integrated. Prior work has indeed shown that visual sequences can modulate the comprehension of written and spoken language (Manfredi, Cohn, De Araújo Andreoli, & Boggio, 2018; Manfredi, Cohn, & Kutas, 2017), just as sentence contexts can modulate the meaning of images (Ganis, Kutas, & Sereno, 1996; Weissman & Tanner, 2018).

An inherent challenge of studying the multimodality of visual narratives is the complex relationship between linguistic and visual content (Cohn, 2016). Consider the strips in Fig. 3a and b, both from *Journal Comics* by Drew Weing, which use captions above fairly illustrative images. Both strips use a “balanced” combination of text and image where both modalities contribute substantially to the multimodal whole. Yet, when the text is omitted, it should be apparent that they differ in their sequential contributions: Fig. 3a/c has a dedicated sequence to the images alone (a temporal order), while Fig. 3b/d does not (a semantic field about smells). One could thus readily rearrange the images in Fig. 3d, but rearrangement in Fig. 3c would change the overall gist. Thus, in building a situation model from both visual and verbal content, readers must negotiate varying types of combinatorial relations between images as well as between modalities.

Exploring these connections between verbal and visual streams of information opens up interfaces with established studies of multimodality, such as co-speech gesture (Goldin-Meadow, 2003; McNeill, 1992)—does multimodal processing take similar characteristics across different integrated modes? How are these information sources processed and represented to convey meaning? How does the nature of different types of multimodal relationships affect processing, representation, and durability in a mental model? Such questions can well interface with education research, which has already explored issues of multimodality (Mayer, 2009), and where visual narratives have growing advocacy (e.g., Hosler & Boomer, 2011; Wong, Miao, Cheng, & Yip, 2017). Finally, given that corpus research on comics has shown that multimodal interactions can change

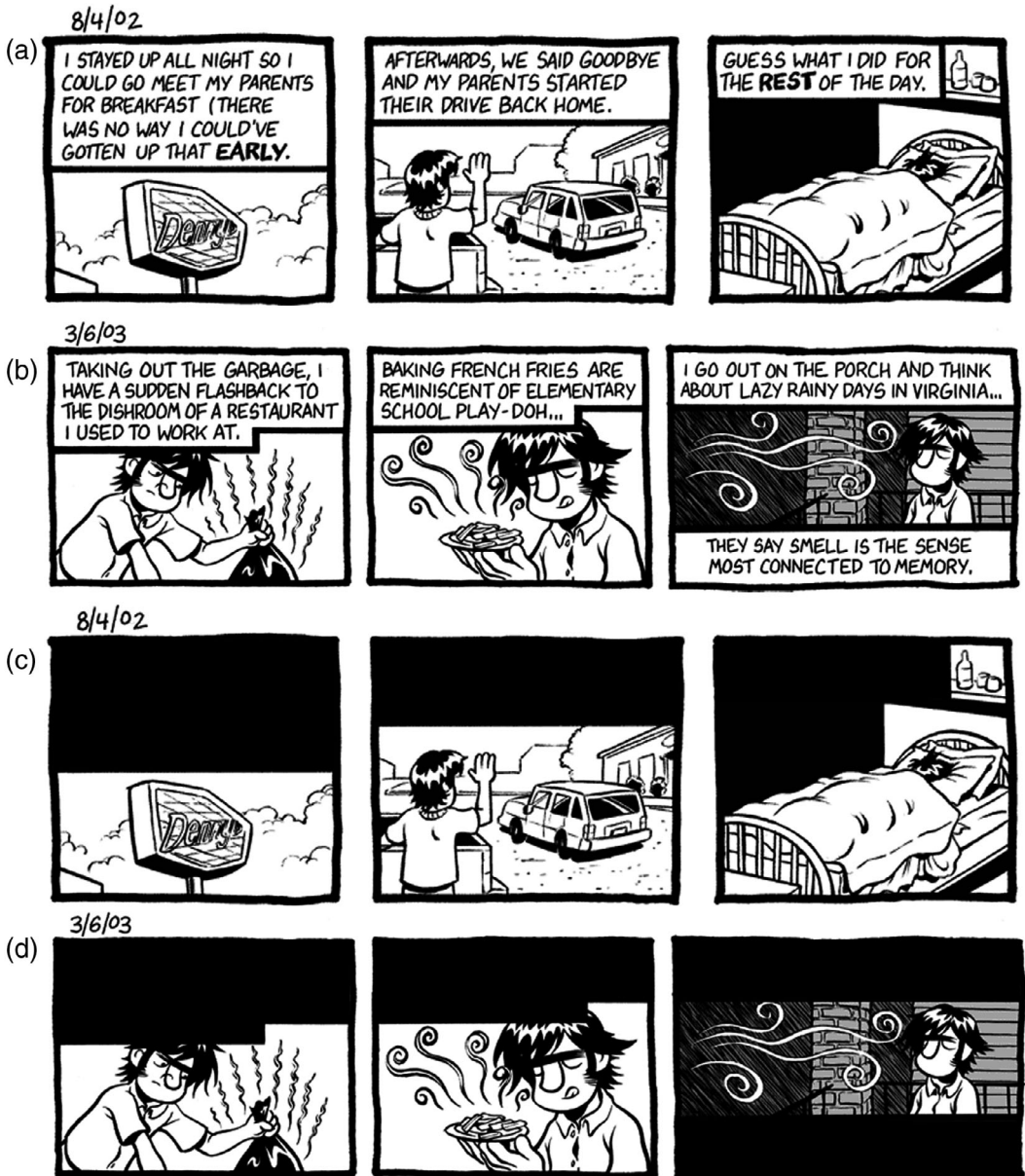


Fig. 3. Comic strips (a and b) from *Journal Comic* (© Drew Weing) which balance the meaning between text and images, but differ in terms of their sequencing: a/c uses a dedicated order (temporal) while b/d uses a semantic field about smells, which could be rearranged.

over time in systematic ways (Cohn, Taylor, & Pederson, 2017), and processing differs based on cross-cultural narrative patterns (Cohn & Kutas, 2017), might readers habituate to processing multimodality in particular ways based on exposure?

2.2. *Sequential image processing is not uniform*

In line with visual narratives' representational levels discussed above, the cognition operating on sequential image processing is multilayered, and not uniform. Studies using event-related brain potentials (ERPs) have directly measured the neural activity of participants comprehending sequential images (for review see Cohn, 2019b). Here, a complex interplay arises between various cognitive mechanisms, such as semantic access indexed by N400 effects (Coderre et al., 2018; Cohn et al., 2012; West & Holcomb, 2002), or integration and updating processes indexed by P600 effects (Cohn et al., 2014; Cohn & Kutas, 2015; Cohn & Maher, 2015). Other effects have been implicated related to combinatorial processing and/or working memory, such as anterior negativities (Cohn et al., 2014; Cohn & Kutas, 2015, 2017), and sensitivity to frequency effects suggested by frontal positivities (Cohn & Maher, 2015; Manfredi et al., 2017). These findings suggest that visual narratives are not simply reducible to one type of processing (ex. updating or inference), but, like most complex cognitive phenomena, evoke multiple interacting mechanisms.

In addition, these neural responses ostensibly appear to overlap with sentence processing and music (Kaan, 2007; Patel, 2003)—outside the context of narratives—thereby suggesting that visual narratives tap into fairly general cognitive mechanisms. Similar overlap in cognitive processing is also implicated for memory systems and segmental processes using behavioral methods (Magliano et al., 2012; Magliano et al., 2015). As noted above, only a few studies have directly compared the cognitive processing of visual narratives (films or drawn narratives) to text-based versions, while attempting to control for narrative content (i.e., attempting to present the same events). These studies have typically explored claims that aspects of comprehension are similar across media, and indeed consistencies do appear in the processing of different versions of a story (Baggett, 1979; Coderre et al., 2018; Magliano, Clinton, O'Brien, & Rapp, 2018; Magliano et al., 2012).

Given that visual narratives have often been used for diagnosis and treatment, additional clues about their processing come from various clinical populations. Damage to the frontal lobe has long been shown to impair visual narrative sequencing (e.g., McFie & Piercy, 1952), as has selective damage to both the left and right hemispheres (Bihrlé, Brownell, Powelson, & Gardner, 1986; Fucetola, Connor, Strube, & Corbetta, 2009; Huber & Gleber, 1982; Marini, Carlomagno, Caltagirone, & Nocentini, 2005; Tinaz et al., 2006). Recent ERP work has shown attenuation for processing semantic incongruities (N400) in both verbal and visual narratives for individuals with Autism Spectrum Disorder (ASD) compared to neurotypical controls (Coderre et al., 2018). Additional deficits in comprehension and sequencing of visual narratives have also been shown by individuals with Developmental Language Disorder, formerly known as Specific Language Impairment (Bishop & Donlan, 2005; Nenadović, Stokić, Vuković, Đoković, & Subotić, 2014), a diagnosis assigned to children with delayed language development, particularly of syntax, who still maintain proficiency at non-verbal IQ tests. Such results further imply connections between the neural architecture of language and visual narratives, and call

into question presumptions that all visual materials may be easier to process than verbal ones for various clinical populations (Coderre, 2019).

Overall, this work suggests that the processing of visual narratives may overlap with that of other media (e.g., Gernsbacher, 1990), and may overlap with different levels or types of expressive systems, *not just narratives* (e.g., sentence processing, music). However, this research by no means can be characterized yet as systematic in exploring the various sources of complexity described above. It does not preclude the possibility that important differences could be attributed to specialized knowledge, processes, or affordances that are unique to the visual modality (Cohn, 2013b; Magliano et al., 2019). Such work thus raises important questions balancing these concerns: What are the brain regions and mechanisms involved in sequential image understanding? How much overlap is there with other, domain-general mechanisms? To what degree does visual narrative processing balance those that are domain-general and those that are domain-specific?

2.3. *Sequential image comprehension is culturally variable*

Contrary to the myth of transparency, visual narrative systems differ in many ways, even when setting aside the modality differences of filmic and drawn narratives. For example, comics and illustrated picture stories vary in many conventions (layout, morphology, narrative structure), despite both conveying meaning across sequential images. Within works labeled as “comics” around the world, systematic variation arises across most all levels of structure. Corpus analyses have suggested cultures’ comics differ across narrative patterns (Cohn, 2019a), page layout (Cohn, Axner, Diercks, Yeh, & Pederson, 2019), motion events (Cohn, Wong, Pederson, & Taylor, 2017), and others. In addition, visual narrative systems have been observed to change over time in dimensions like layout, narrative, and multimodality (Bateman, Veloso, & Lau, 2019; Cohn, Taylor, et al., 2017; Pederson & Cohn, 2016). Other systems vary in ways significantly different from those used in comics, like the visual narratives drawn in the sand by Central Australian Aboriginals (Green, 2014; Wilkins, 1997/2016). Overall, such observations parallel with language: While there may be abstract similarities and typological constructs underlying visual narratives of the world, they manifest in distinct and patterned systems used by a particular population.

While the iconicity of images aids in cross-cultural understanding, fluency is modulated by knowledge of specific visual narrative systems, and knowledge of one visual narrative system does not grant full fluency in others. For example, the aforementioned sand narratives used by Central Australian Aboriginals use highly conventionalized representations that are likely opaque for those unfamiliar with it (Green, 2014; Wilkins, 1997/2016), while some individuals from this community also had trouble understanding the sequencing of Western comics (Wilkins, 1997/2016). In addition, a study among self-described “comic readers” found that neural responses to a pattern found more prevalently in Japanese manga is modulated by readership of those works (Cohn & Kutas, 2017).

Such findings raise questions about the nature of cross-cultural variability of visual narratives and the fluency associated with them: Do visual narratives across cultures and

history share “universal” features and or underlying typological regularities? To what degree is visual narrative comprehension contingent on familiarity with visual narrative systems in general or with specific systems given a particular genre, culture, and/or historical context? To what degree does the typological structure of a visual narrative system overlap with the structures in other expressive capacities, like verbal or signed languages? How does cross-cultural diversity relate to the identity of their producers and the socio-cultural treatment of different types of images?

2.4. Sequential image comprehension requires a fluency, learned across a developmental trajectory

Finally, in discussing myths, we made the case that being a fluent reader of drawn sequential narratives requires exposure. In many cultures, children’s first exposure to reading is through multimodal visual narratives, and various works have suggested that visual narratives might be beneficial for literacy, language learning, and inferences (Kendeou et al., 2019). Yet, as discussed above, understanding of visual narratives themselves does not come “for free” with visual perception and event cognition, and requires a fluency developed with exposure and practice. While foundational research has suggested that understanding sequential images moves through identifiable steps in development, we still know very little about this process in detail: How much exposure and practice are necessary? What are the stages of visual narrative understanding? How does this development balance general cognitive principles and culturally specific patterns?

In addition, there is often an asymmetry between visual narrative comprehension and production—far more people can comprehend sequential images than can create them (Wilson, 2016). Important work has shown that exposure to visual narratives—and especially copying existing systems—is crucial for developing proficiency in creating visual narratives (Stoermer, 2009; Wilson, 1988; Wilson & Wilson, 1977). Yet little is understood about the trajectory of this learning, how it lines up with the development of comprehension, and the effects of such proficiency on other systems, like literacy and intelligence.

All of these issues also raise questions about the development of visual narrative comprehension relative to other domains, like the concurrent development of Theory of Mind and language. To what degree does the development of these domains line up with each other? To what degree is visual narrative understanding (in)dependent of other aspects of cognition?

3. Conclusion

Throughout, we have emphasized that visual narratives reflect fundamental aspects of human meaning-making, with complexity beyond what is often presumed. This complexity deserves a focus of its own within the broader study of human cognition, and should not be stifled by cultural myths of its transparency or universality. At the same time,

though, the complexity of visual narratives affords fundamental questions across Cognitive Science. Scholars in a variety of sub-disciplines of Cognitive Science (event cognition, scene perception, psycholinguistics, clinical psychology, literacy, narrative comprehension, communications, cultural psychology, and more) have begun developing viable programs of research on this topic. We are at a crucial juncture where these efforts could either converge into a new and emerging sub-field of Cognitive Science, or these efforts could simply continue as fractionated research across several sub-disciplines.

How do we prevent such balkanization? One solution is to establish a common understanding of the problem we're trying to understand. We believe the framework presented here provides one approach by giving access points for identifying problems that researchers can address. As discussed above, the framework outlines that visual narratives express meaning through a graphic modality, which is organized using combinatorial structure that operates at both the unit and sequence level. This complex framework manifests in culturally different ways, is multimodal in nature, is learned across a developmental trajectory, and recruits non-uniform processing mechanisms. In the context of this framework, we posed questions about these issues that we find interesting and worth pursuing, but by no means should be considered exhaustive.

Another potential solution is proposing theories that afford intersections between different areas of Cognitive Science. We believe our proposed framework makes the case that this is warranted. Several emerging theories of visual narrative processing range in the extent to which they accomplish this agenda, spanning fields of psycholinguistics and cognitive neuroscience (Cohn, 2013b, 2019b), scene perception and event cognition (Loschky et al., 2019), attention (Smith, 2012), developmental psychology (Nakazawa, 2016), multimodal discourse (Bateman & Wildfeuer, 2014), and computer science (Augereau, Iwata, & Kise, 2018), among others. However, we have intentionally not discussed these theories herein. Rather, we wish to inspire scholars to test these models and/or develop their own, but hope that said theories are sensitive of the need to cut across different areas of research on cognitive processing. In doing so, we believe that we will not only learn about how visual narratives are processed, but can also learn about how the brain and mind are coordinated to make meaning across domains.

References

- Augereau, O., Iwata, M., & Kise, K. (2018). A survey of comics research in computer science. *Journal of Imaging*, 4(87), 1–19.
- Baggett, P. (1979). Structurally equivalent stories in movie and text and the effect of the medium on recall. *Journal of Verbal Learning and Verbal Behavior*, 18(3), 333–356. [https://doi.org/10.1016/S0022-5371\(79\)90191-9](https://doi.org/10.1016/S0022-5371(79)90191-9).
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1986). Mechanical, behavioural and intentional understanding of picture stories in autistic children. *British Journal of Developmental Psychology*, 4(2), 113–125.
- Bateman, J. A., Beckmann, A., & Varela, R. I. (2018). From empirical studies to visual narrative organization: Exploring page composition. In A. Dunst, J. Laubrock, & J. Wildfeuer (Eds.), *Empirical comics research: Digital, multimodal, and cognitive methods* (pp. 127–153). New York: Routledge.

- Bateman, J. A., Veloso, F. O. D., & Lau, Y. L. (2019). On the track of visual style: A diachronic study of page composition in comics and its functional motivation. *Visual Communication*. <https://doi.org/10.1177/1470357219839101>.
- Bateman, J. A., Veloso, F. O. D., Wildfeuer, J., Cheung, F. H., & Guo, N. S. (2016). An open multilevel classification scheme for the visual layout of comics and graphic novels: Motivation and design. *Digital Scholarship in the Humanities*, 32(3), 476–510. <https://doi.org/10.1093/llc/fqw024>.
- Bateman, J. A., & Wildfeuer, J. (2014). A multimodal discourse theory of visual narrative. *Journal of Pragmatics*, 74, 180–208. <https://doi.org/10.1016/j.pragma.2014.10.001>.
- Beatty, W. W., Jovic, Z., & Monson, N. (1993). Picture sequencing by schizophrenic patients. *Bulletin of the Psychonomic Society*, 31(4), 265–267.
- Beatty, W. W., & Monson, N. (1994). Picture and motor sequencing in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, 16(2), 165–172. <https://doi.org/10.1080/01688639408402627>.
- Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum.
- Bihrlé, A. M., Brownell, H. H., Powelson, J. A., & Gardner, H. (1986). Comprehension of humorous and nonhumorous materials by left and right brain-damaged patients. *Brain and Cognition*, 5, 399–411.
- Bishop, A. (1977). Is a picture worth a thousand words? *Mathematics Teaching*, 81, 32–35.
- Bishop, D. V. M., & Donlan, C. (2005). The role of syntax in encoding and recall of pictorial narratives: Evidence from specific language impairment. *British Journal of Developmental Psychology*, 23(1), 25–46. <https://doi.org/10.1348/026151004X20685>.
- Black, J. B., & Wilensky, R. (1979). An evaluation of story grammars. *Cognitive Science*, 3, 213–230.
- Bornens, M.-T. (1990). Problems brought about by “reading” a sequence of pictures. *Journal of Experimental Child Psychology*, 49(2), 189–226. [https://doi.org/10.1016/0022-0965\(90\)90055-D](https://doi.org/10.1016/0022-0965(90)90055-D).
- Boroditsky, L., Gaby, A., & Levinson, S. C. (2008). Time in space. In A. Majid (Ed.), *Field manual*. Vol. 11 (pp. 52–76). Nijmegen: Max Planck Institute for Psycholinguistics.
- Breiger, B. (1956). The use of the W-B picture arrangement subtest as a projective technique. *Journal of Consulting Psychology*, 20(2), 132. Retrieved from <http://europepmc.org/abstract/MED/13306842>.
- Brown, A. L., & French, L. A. (1976). Construction and regeneration of logical sequences using causes or consequences as the point of departure. *Child Development*, 47(4), 930–940. <https://doi.org/10.2307/1128428>.
- Burris, S., & Brown, D. (2014). When all children comprehend: Increasing the external validity of narrative comprehension development research. *Frontiers in Psychology*, 5(168), 1–16. <https://doi.org/10.3389/fpsyg.2014.00168>
- Byram, M. L., & Garforth, C. (1980). Research and testing non-formal education materials: A multi-media extension project in Botswana. *Educational Broadcasting International*, 13(4), 190–194.
- Campbell, J. M., & McCord, D. M. (1996). The WAIS-R comprehension and picture arrangement subtests as measures of social intelligence: Testing traditional interpretations. *Journal of Psychoeducational Assessment*, 14, 240–249.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Christiansen, M. H., & Chater, N. (2016). *Creating language: Integrating evolution, acquisition, and processing*. Cambridge, MA: MIT Press.
- Christiansen, M. H., Conway, C. M., & Onnis, L. (2011). Similar neural correlates for language and sequential learning: Evidence from event-related brain potentials. *Language and Cognitive Processes*, 27(2), 231–256. <https://doi.org/10.1080/01690965.2011.606666>.
- Coderre, E. L. (2019). Dismantling the “Visual Ease assumption”: A review of visual narrative processing in clinical populations. *Topics Cognitive Science*. <https://doi.org/10.1111/tops.12446>
- Coderre, E. L., Cohn, N., Slipher, S. K., Chernenok, M., Ledoux, K., & Gordon, B. (2018). Visual and linguistic narrative comprehension in autism spectrum disorders: Neural evidence for modality-independent impairments. *Brain and Language*, 186, 44–59.
- Cohn, N. (2012). Explaining “I can’t draw”: Parallels between the structure and development of language and drawing. *Human Development*, 55(4), 167–192. <https://doi.org/10.1159/000341842>.

- Cohn, N. (2013a). Navigating comics: An empirical and theoretical approach to strategies of reading comic page layouts. *Frontiers in Psychology*, 4, 1–15. doi:10.3389/fpsyg.2013.00186.
- Cohn, N. (2013b). *The visual language of comics: Introduction to the structure and cognition of sequential images*. London: Bloomsbury.
- Cohn, N. (2013c). Visual narrative structure. *Cognitive Science*, 37(3), 413–452. <https://doi.org/10.1111/cogs.12016>.
- Cohn, N. (2014). You're a good structure, Charlie Brown: The distribution of narrative categories in comic strips. *Cognitive Science*, 38(7), 1317–1359. <https://doi.org/10.1111/cogs.12116>.
- Cohn, N. (2016). A multimodal parallel architecture: A cognitive framework for multimodal interactions. *Cognition*, 146, 304–323. <https://doi.org/10.1016/j.cognition.2015.10.007>.
- Cohn, N. (2019a). Structural complexity in visual narratives: Theory, brains, and cross-cultural diversity. In M. Grishakova & M. Poulaki (Eds.), *Narrative complexity and media: Experiential and cognitive interfaces* (pp. 174–199). Lincoln: University of Nebraska Press.
- Cohn, N. (2019b). Your brain on comics: A cognitive model of visual narrative comprehension. *Topics in Cognitive Science*, <https://doi.org/10.1111/tops.12421>.
- Cohn, N., Axner, J., Diercks, M., Yeh, R., & Pederson, K. (2019). The cultural pages of comics: Cross-cultural variation in page layouts. *Journal of Graphic Novels and Comics*, 10(1), 67–86. <https://doi.org/10.1080/21504857.2017.1413667>.
- Cohn, N., & Bender, P. (2017). Drawing the line between constituent structure and coherence relations in visual narratives. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 43(2), 289–301. <https://doi.org/10.1037/xlm0000290>.
- Cohn, N. (in press). Visual narrative comprehension: universal or not? *Psychonomic Bulletin & Review*.
- Cohn, N., Jackendoff, R., Holcomb, P. J., & Kuperberg, G. R. (2014). The grammar of visual narrative: Neural evidence for constituent structure in sequential image comprehension. *Neuropsychologia*, 64, 63–70. <https://doi.org/10.1016/j.neuropsychologia.2014.09.018>.
- Cohn, N., & Kutas, M. (2015). Getting a cue before getting a clue: Event-related potentials to inference in visual narrative comprehension. *Neuropsychologia*, 77, 267–278. <https://doi.org/10.1016/j.neuropsychologia.2015.08.026>.
- Cohn, N., & Kutas, M. (2017). What's your neural function, visual narrative conjunction? Grammar, meaning, and fluency in sequential image processing. *Cognitive Research: Principles and Implications*, 2 (27), 1–13. <https://doi.org/10.1186/s41235-017-0064-5>.
- Cohn, N., & Magliano, J. P. (2019). *Editors' introduction and review: Visual narrative research: An emerging field in cognitive science*. *Topics in Cognitive Science*. <https://doi.org/10.1111/tops.12473>
- Cohn, N., & Maher, S. (2015). The notion of the motion: The neurocognition of motion lines in visual narratives. *Brain Research*, 1601, 73–84. <https://doi.org/10.1016/j.brainres.2015.01.018>.
- Cohn, N., Murthy, B., & Foulsham, T. (2016). Meaning above the head: Combinatorial constraints on the visual vocabulary of comics. *Journal of Cognitive Psychology*, 28(5), 559–574. <https://doi.org/10.1080/20445911.2016.1179314>.
- Cohn, N., Paczynski, M., Jackendoff, R., Holcomb, P. J., & Kuperberg, G. R. (2012). (Pea)nuts and bolts of visual narrative: Structure and meaning in sequential image comprehension. *Cognitive Psychology*, 65(1), 1–38. <https://doi.org/10.1016/j.cogpsych.2012.01.003>.
- Cohn, N., Taylor, R., & Pederson, K. (2017). A picture is worth more words over time: Multimodality and narrative structure across eight decades of American superhero comics. *Multimodal Communication*, 6(1), 19–37. <https://doi.org/10.1515/mc-2017-0003>.
- Cohn, N., & Wittenberg, E. (2015). Action starring narratives and events: Structure and inference in visual narrative comprehension. *Journal of Cognitive Psychology*, 27(7), 812–828. <https://doi.org/10.1080/20445911.2015.1051535>.
- Cohn, N., Wong, V., Pederson, K., & Taylor, R. (2017). Path salience in motion events from verbal and visual languages. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 1794–1799). Austin, TX: Cognitive Science Society.

- Cook, B. L. (1980). Picture communication in the Papua New Guinea. *Educational Broadcasting International*, 13(2), 78–83.
- Culicover, P. W., & Jackendoff, R. (2005). *Simpler syntax*. Oxford: Oxford University Press.
- Duncan, H. F., Gourlay, N., & Hudson, W. (1973). *A study of pictorial perception among bantu and white school children*. Johannesburg: Witwaterstrand University Press.
- Duncan, R., Smith, M. J., & Levitz, P. (2015). *The power of comics* (2nd ed.). New York: Continuum Books.
- Fivush, R., & Mandler, J. M. (1985). Developmental changes in the understanding of temporal sequence. *Child Development*, 56(6), 1437–1446. <https://doi.org/10.2307/1130463>.
- Forceville, C. (2005). Visual representations of the idealized cognitive model of anger in the Asterix album *La Zizanie*. *Journal of Pragmatics*, 37(1), 69–88.
- Forceville, C. (2011). Pictorial runes in *Tintin and the Picaros*. *Journal of Pragmatics*, 43(3), 875–890.
- Forceville, C. (2016). Conceptual metaphor theory, blending theory, and other cognitivist perspectives on comics. In N. Cohn (Ed.), *The visual narrative reader* (pp. 89–114). London: Bloomsbury.
- Friedman, W. J. (1990). Children's representations of the pattern of daily activities. *Child Development*, 61(5), 1399–1412. <https://doi.org/10.1111/j.1467-8624.1990.tb02870.x>.
- Fucetola, R., Connor, L. T., Strube, M. J., & Corbetta, M. (2009). Unravelling nonverbal cognitive performance in acquired aphasia. *Aphasiology*, 23(12), 1418–1426. <https://doi.org/10.1080/02687030802514938>.
- Fussell, D., & Haaland, A. (1978). Communicating with pictures in nepal: Results of practical study used in visual education. *Educational Broadcasting International*, 11(1), 25–31.
- Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for "common sense": An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience*, 8, 89–106.
- Gernsbacher, M. A. (1985). Surface information loss in comprehension. *Cognitive Psychology*, 17, 324–363.
- Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Lawrence Erlbaum.
- Gibson, J. J. (2014). *The ecological approach to visual perception* (classic ed). New York: Psychology Press.
- Givón, T. (1993). *English grammar: A function-based introduction* (Vol. 2). Amsterdam: John Benjamins.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.
- Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.
- Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse Comprehension. *Annual Review of Psychology*, 48, 163–189.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371–395.
- Green, J. (2014). *Drawn from the ground: Sound, sign and inscription in Central Australian sand stories*. Cambridge, UK: Cambridge University Press.
- Hagmann, C. E., & Cohn, N. (2016). The pieces fit: Constituent structure and global coherence of visual narrative in RSVP. *Acta Psychologica*, 164, 157–164. <https://doi.org/10.1016/j.actpsy.2016.01.011>.
- Hosler, J., & Boomer, K. B. (2011). Are comic books an effective way to engage Nonmajors in learning and appreciating science? *CBE-Life Sciences Education*, 10(3), 309–317. Retrieved from <http://www.lifescied.org/content/10/3/309.abstract>.
- Huber, W., & Gleber, J. (1982). Linguistic and nonlinguistic processing of narratives in aphasia. *Brain and Language*, 16, 1–18.
- Huntsinger, C. S., Jose, P. E., Krieg, D. B., & Luo, Z. (2011). Cultural differences in Chinese American and European American children's drawing skills over time. *Early Childhood Research Quarterly*, 26(1), 134–145. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0885200610000281>.
- Hutson, J. P., Magliano, J., & Loschky, L. C. (2018). Understanding moment-to-moment processing of visual narratives. *Cognitive Science*, 42(8), 2999–3033. <https://doi.org/10.1111/cogs.12699>.
- Hutson, J. P., Smith, T. J., Magliano, J. P., & Loschky, L. C. (2017). What is the role of the film viewer? The effects of narrative comprehension and viewing task on gaze control in film. *Cognitive Research: Principles and Implications*, 2(1), 46. <https://doi.org/10.1186/s41235-017-0080-5>.

- Ildirar, S., & Schwan, S. (2015). First-time viewers' comprehension of films: Bridging shot transitions. *British Journal of Psychology*, 106(1), 133–151. <https://doi.org/10.1111/bjop.12069>.
- Ingber, S., & Eden, S. (2011). Enhancing sequential time perception and storytelling ability of deaf and hard of hearing children. *American Annals of the Deaf*, 156(4), 391–401.
- Iyyer, M., Manjunatha, V., Guha, A., Vyas, Y., Boyd-Graber, J., Daumé, H., III, & Davis, L. (2017). The amazing mysteries of the gutter: Drawing inferences between panels in comic book narratives. *CVPR*, 6478–6487.
- Kaan, E. (2007). Event-related potentials and language processing: A brief overview. *Language and Linguistics Compass*, 1(6), 571–591. <https://doi.org/10.1111/j.1749-818X.2007.00037.x>.
- Kaufman, A. S., & Lichtenberger, E. O. (2006). *Assessing adolescent and adult intelligence* (3rd ed.). Hoboken, NJ: Wiley.
- Kendall, L. N., Raffaelli, Q., Kingstone, A., & Todd, R. M. (2016). Iconic faces are not real faces: enhanced emotion detection and altered neural processing as faces become more iconic. *Cognitive Research: Principles and Implications*, 1(1), 1–14. <https://doi.org/10.1186/s41235-016-0021-8>.
- Kendeou, P., McMaster, K. L., Butterfuss, R., Kim, J., Bresina, B., & Wagner, K. (2019). The inferential language comprehension (iLC) framework: Supporting Children's comprehension of visual narratives. *Topics in Cognitive Science*. <https://doi.org/10.1111/tops.12457>
- Khetarpal, K., & Jain, E. (2016). A preliminary benchmark of four saliency algorithms on comic art. Paper presented at the Multimedia & Expo Workshops (ICMEW), 2016 IEEE International Conference on.
- Kirtley, C., Murray, C., Vaughan, P. B., & Tatler, B. W. (2018). Reading words and images: Factors influencing eye movements in comic reading. In A. Dunst, J. Laubrock, & J. Wildfeuer (Eds.), *Empirical comics research: Digital, multimodal, and cognitive methods* (pp. 264–283). New York: Routledge.
- Kunen, S., Chabaud, S. A., & Dean, A. L. (1987). Figural factors and the development of pictorial inferences. *Journal of Experimental Child Psychology*, 44(2), 157–169. [https://doi.org/10.1016/0022-0965\(87\)90028-2](https://doi.org/10.1016/0022-0965(87)90028-2).
- Laubrock, J., & Dunst, A. (2019). Computational approaches to comics. *Topics in Cognitive Science*. <https://doi.org/10.1111/tops.12476>
- Laubrock, J., Hohenstein, S., & Kümmerer, M. (2018). Attention to comics: Cognitive processing during the reading of graphic literature. In A. Dunst, J. Laubrock, & J. Wildfeuer (Eds.), *Empirical comics research: Digital, multimodal, and cognitive methods* (pp. 239–263). New York: Routledge.
- Le Guen, O., & Pool Balam, L. I. (2012). No metaphorical timeline in gesture and cognition among Yucatec Mayas. *Frontiers in Psychology*, 3, 1–15. <https://doi.org/10.3389/fpsyg.2012.00271>.
- Lee, J. F., & Armour, W. S. (2016). Factors influencing non-native readers' sequencing of Japanese manga panels. In S. Pasfield-Neofitou & C. Sell (Ed.), *Manga vision* (pp. 178–193). Clayton, Australia: Monash University.
- Liddell, C. (1996). Every picture tells a story: South African and British children interpreting pictures. *British Journal of Developmental Psychology*, 14(3), 355–363. <https://doi.org/10.1111/j.2044-835X.1996.tb00711.x>.
- Lipsitz, J. D., Dworkin, R. H., & Erlenmeyer-Kimling, L. (1993). Wechsler comprehension and picture arrangement subtests and social adjustment. *Psychological Assessment*, 5(4), 430–437.
- Loschky, L. C., Magliano, J., Larson, A. M., & Smith, T. J. (2019). The Scene perception & event comprehension theory (SPECT) applied to visual narratives. *Topics in Cognitive Science*. <https://doi.org/10.1111/tops.12455>
- Loughlin, S., Grossnickle, E., Dinsmore, D., & Alexander, P. (2015). “Reading” paintings: Evidence for trans-symbolic and symbol-specific comprehension processes. *Cognition and Instruction*, 33(3), 257–293. <https://doi.org/10.1080/07370008.2015.1076822>.
- Magliano, J. P., Clinton, J. A., O'Brien, E. J., & Rapp, D. N. (2018). Detecting differences between adapted narratives. In A. Dunst, J. Laubrock, & J. Wildfeuer (Eds.), *Empirical comics research: Digital, multimodal, and cognitive methods* (pp. 284–304). New York: Routledge.
- Magliano, J. P., Dijkstra, K., & Zwaan, R. A. (1996). Generating predictive inferences while viewing a movie. *Discourse Processes*, 22, 199–224.

- Magliano, J. P., Higgs, K., & Clinton, J. A. (2019). Sources of complexity in comprehension across modalities of narrative experience. In M. Grishakova & M. Poulaki (Eds.), *Narrative complexity and media: Experiential and cognitive interfaces* (pp. 149–173). Lincoln: University of Nebraska Press.
- Magliano, J. P., Kopp, K., Higgs, K., & Rapp, D. N. (2017). Filling in the gaps: Memory implications for inferring missing content in graphic narratives. *Discourse Processes*, 54(8), 569–582. <https://doi.org/10.1080/0163853X.2015.1136870>.
- Magliano, J. P., Kopp, K., McNeerney, M. W., Radvansky, G. A., & Zacks, J. M. (2012). Aging and perceived event structure as a function of modality. *Aging, Neuropsychology, and Cognition*, 19(1–2), 264–282. <https://doi.org/10.1080/13825585.2011.633159>.
- Magliano, J. P., Larson, A. M., Higgs, K., & Loschky, L. C. (2015). The relative roles of visuospatial and linguistic working memory systems in generating inferences during visual narrative comprehension. *Memory & Cognition*, 44(2), 207–219. <https://doi.org/10.3758/s13421-015-0558-7>.
- Magliano, J. P., Loschky, L. C., Clinton, J. A., & Larson, A. M. (2013). Is reading the same as viewing? an exploration of the similarities and differences between processing text- and visually based narratives. In B. Miller, L. Cutting, & P. McCardle (Eds.), *Unraveling the behavioral, neurobiological, and genetic components of reading comprehension* (pp. 78–90). Baltimore, MD: Brookes.
- Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 9, 111–151.
- Manfredi, M., Cohn, N., De Araújo Andreoli, M., & Boggio, P. S. (2018). Listening beyond seeing: Event-related potentials to audiovisual processing in visual narrative. *Brain and Language*, 185, 1–8. <https://doi.org/10.1016/j.bandl.2018.06.008>.
- Manfredi, M., Cohn, N., & Kutas, M. (2017). When a hit sounds like a kiss: an electrophysiological exploration of semantic processing in visual narrative. *Brain and Language*, 169, 28–38. <https://doi.org/10.1016/j.bandl.2017.02.001>.
- Marini, A., Carlomagno, S., Caltagirone, C., & Nocentini, U. (2005). The role played by the right hemisphere in the organization of complex textual structures. *Brain and Language*, 93(1), 46–54. <https://doi.org/10.1016/j.bandl.2004.08.002>.
- Mayer, R. E. (2009). *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge, UK: Cambridge University Press.
- McCloud, S. (1993). *Understanding comics: The invisible art*. New York: Harper Collins.
- McFie, J., & Piercy, M. F. (1952). Intellectual impairment with localized cerebral lesions. *Brain: A Journal of Neurology*, 75, 292–311.
- McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. *Psychology of Learning and Motivation*, 51, 297–384.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.
- Nakazawa, J. (2016). Manga literacy and manga comprehension in Japanese children. In N. Cohn (Ed.), *The visual narrative reader* (pp. 157–184). London: Bloomsbury.
- Nenadović, V., Stokić, M., Vuković, M., Đoković, S., & Subotić, M. (2014). Cognitive and electrophysiological characteristics of children with specific language impairment and subclinical epileptiform electroencephalogram. *Journal of Clinical and Experimental Neuropsychology*, 36(9), 981–991. <https://doi.org/10.1080/13803395.2014.958438>.
- Newton, D. P. (1985). Children's perception of pictorial metaphor. *Educational Psychology*, 5(2), 179–185. <https://doi.org/10.1080/0144341850050207>.
- Nguyen, N.-V., Rigaud, C., & Burie, J.-C. (2017). Comic characters detection using deep learning. Paper presented at the Document Analysis and Recognition (ICDAR), 2017 14th IAPR International Conference.
- Okada, T., & Ishibashi, K. (2017). Imitation, inspiration, and creation: Cognitive process of creative drawing by copying Others' artworks. *Cognitive Science*, 41(7), 1804–1837. <https://doi.org/10.1111/cogs.12442>.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6(7), 674–681. <https://doi.org/10.1038/nn1082>.

- Pederson, K., & Cohn, N. (2016). The changing pages of comics: Page layouts across eight decades of American superhero comics. *Studies in Comics*, 7(1), 7–28. https://doi.org/10.1386/stic.7.1.7_1.
- Petersen, R. S. (2011). *Comics, manga, and graphic novels: A history of graphic narratives*. Santa Barbara, CA: ABC-CLIO.
- Pezdek, K., Lehrer, A., & Simon, S. (1984). The relationship between reading and cognitive processing of television and radio. *Child Development*, 55(6), 2072–2082. <https://doi.org/10.2307/1129780>.
- Radvansky, G. A., & Zacks, J. (2014). *Event Cognition*. Oxford, UK: Oxford University Press.
- Ramos, M. C., & Die, A. H. (1986). The Wais-R picture arrangement subtest: what do scores indicate? *The Journal of General Psychology*, 113(3), 251–261. <https://doi.org/10.1080/00221309.1986.9711036>.
- Rigaud, C., Guérin, C., Karatzas, D., Burie, J.-C., & Ogier, J.-M. (2015). Knowledge-driven understanding of images in comic books. *International Journal on Document Analysis and Recognition (IJ DAR)*, 18(3), 199–221.
- Robertson, D. A. (2000). *Functional neuroanatomy of narrative comprehension*. (Doctoral dissertation). Madison: University of Wisconsin.
- Saito, Y., Hirai, K., & Horiuchi, T. (2015). Construction of manga materials database for analyzing perception of materials in line drawings. In *Proceedings of the 2015 Color and Imaging Conference (Vol. 2015, pp. 201–206)*: Society for Imaging Science and Technology.
- San Roque, L., Gawne, L., Hoenigman, D., Miller, J. C., Rumsey, A., Spronck, S., ... Evans, N. (2012). Getting the story straight: Language fieldwork using a narrative problem-solving task. *Language Documentation and Conservation*, 6, 135–174.
- Schmidt, C. R., & Paris, S. G. (1978). Operativity and Reversibility in Children's Understanding of Pictorial Sequences. *Child Development*, 49(4), 1219–1222. <https://doi.org/10.2307/1128764>.
- Schodt, F. L. (1983). *Manga! Manga! The world of Japanese comics*. New York: Kodansha America.
- Sivaratnam, C. S., Cornish, K., Gray, K. M., Howlin, P., & Rinehart, N. J. (2012). Brief report: Assessment of the social-emotional profile in children with autism spectrum disorders using a novel comic strip task. *Journal of Autism and Developmental Disorders*, 42(11), 2505–2512. <https://doi.org/10.1007/s10803-012-1498-8>.
- Smith, T. J. (2012). The attentional theory of cinematic continuity. *Projections*, 6(1), 1–27. <https://doi.org/10.3167/proj.2012.060102>.
- Stoermer, M. (2009). Teaching between the frames: Making comics with seven and eight year old children, a search for craft and pedagogy. (Doctoral Dissertation). Indiana University, Indiana.
- Szawerna, M. (2017). *Metaphoricity of Conventionalized Diabetic Images in Comics: A Study in Multimodal Cognitive Linguistics*. Frankfurt am Main: Peter Lang.
- Takayama, K., Johan, H., & Nishita, T. (2012). Face detection and face recognition of cartoon characters using feature extraction. Kuching, Malaysia. Paper presented at the Image, Electronics and Visual Computing Workshop.
- Tinaz, S., Schendan, H. E., Schon, K., & Stern, C. E. (2006). Evidence for the importance of basal ganglia output nuclei in semantic event sequencing: An fMRI study. *Brain Research*, 1067(1), 239–249. <https://doi.org/10.1016/j.brainres.2005.10.057>.
- Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes*, 15, 249–275.
- Trabasso, T., van den Broek, P., & Suh, S. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes*, 12, 1–25.
- Tulsky, D. S., & Price, L. R. (2003). The joint WAIS-III and WMS-III factor structure: development and cross-validation of a six-factor model of cognitive functioning. *Psychological Assessment*, 15(2), 149–162.
- Ueno, M., & Isahara, H. (2017). Story Pattern Analysis Based on Scene Order Information in Four-Scene Comics. Recognition (ICDAR). Paper presented at the 2017 14th IAPR International Conference on Document Analysis.
- Ueno, M., Mori, N., Suenaga, T., & Isahara, H. (2016). Estimation of structure of four-scene comics by convolutional neural networks. Paper presented at the Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding.

- van Dijk, T., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Weissman, B., & Tanner, D. (2018). A strong wink between verbal and emoji-based irony: how the brain processes ironic emojis during language comprehension. *PLoS ONE*, *13*(8), e0201727.
- Weist, R. M. (2009). Children think and talk about time and space. In P. Łobacz, P. Nowak & W. Zabrocki (Eds.), *Language, Science, and Culture*. Wydawnictwo Naukowe UAM: Poznań.
- Weist, R. M., Lyytinen, P., Wysocka, J., & Atanassova, M. (1997). The interaction of language and thought in children's language acquisition: A crosslinguistic study. *Journal of Child Language*, *24*(1), 81–121.
- West, W. C. (1998). Common versus multiple semantic systems: An electrophysiological examination of the comprehension of verbal and picture stories. (Doctoral Dissertation). Medford, MA: Tufts University.
- West, W. C., & Holcomb, P. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research*, *13*, 363–375.
- Wilkins, D. P. (1997/2016). Alternative representations of space: Arrernte narratives in sand. In N. Cohn (Ed.), *The visual narrative reader* (pp. 252–281). London: Bloomsbury.
- Willats, J. (2005). *Making sense of children's drawings*. Mahwah, NJ: Lawrence Erlbaum.
- Wilson, B. (1988). The artistic Tower of Babel: Inextricable links between culture and graphic development. In G. W. Hardiman, & T. Zernich (Eds.), *Discerning art: Concepts and issues* (pp. 488–506). Champaign, IL: Stipes.
- Wilson, B. (2016). What happened and what happened next: Kids' visual narratives across cultures. In N. Cohn (Ed.), *The visual narrative reader* (pp. 185–227). London: Bloomsbury.
- Wilson, B., & Wilson, M. (1977). An iconoclastic view of the imagery sources in the drawings of young people. *Art Education*, *30*(1), 4–12.
- Wong, S. W. L., Miao, H., Cheng, R. W.-Y., & Yip, M. C. W. (2017). Graphic novel comprehension among learners with differential cognitive styles and reading abilities. *Reading & Writing Quarterly*, *33*(5), 412–427. <https://doi.org/10.1080/10573569.2016.1216343>.
- Zampini, L., Zanchi, P., Suttora, C., Spinelli, M., Fasolo, M., & Salerni, N. (2017). Assessing sequential reasoning skills in typically developing children. *BPA-Applied Psychology Bulletin (Bollettino di Psicologia Applicata)*, *65*(279), 44–50.
- Zhao, F., & Mahrt, N. (2018). Influences of comics expertise and comics types in comics reading. *International Journal of Innovation and Research in Educational Sciences*, *5*(2), 218–224.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185.

Papers in this topic

- Cohn, N., & Magliano, J. P. (2019). Editors' Introduction and Review: Visual Narrative Research: An Emerging Field in Cognitive Science. *Topics in Cognitive Science*, *12*(1), 197–223.
- Cohn, N. (2019). Your brain on comics: A cognitive model of visual narrative comprehension. *Topics in Cognitive Science*, *12*(1), 352–386.
- Coderre, E. (2019). Dismantling the "Visual Ease Assumption:" A Review of Visual Narrative Processing in Clinical Populations. *Topics in Cognitive Science*, *12*(1), 224–255.
- Kendeou, P., McMaster, K. L., Butterfuss, R., Kim, J., Bresina, B., & Wagner, K. (2019). The Inferential Language Comprehension (iLC) Framework: Supporting Children's Comprehension of Visual Narratives. *Topics in Cognitive Science*, *12*(1), 256–273.
- Laubrock, J., & Dunst, A. (2019). Computational approaches to comics. *Topics in Cognitive Science*, *12*(1), 274–310.
- Loschky, Lester C., Magliano, Joseph, Larson, Adam M., & Smith, Tim J. (2019). The scene perception & event comprehension theory (SPECT) applied to visual narratives. *Topics in Cognitive Science*, *12*(1), 311–351.