



Understanding Comics. A Comparison between Children and Adults through a Coherence/Incoherence Paradigm in an Eye-tracking Study

Lorena A. Martín-Arnal^a, José A. León^a, Paul van den Broek^b, and Ricardo Olmos^a

^aUniversidad Autónoma de Madrid, Spain; ^bUniversity of Leiden, The Netherlands

ARTICLE INFO

Article history:

Received 11 February 2019

Accepted 15 April 2019

Keywords:

Reading comprehension
Visual narrative grammar
Comic comprehension
Eye movements
Eye tracking
Coherence/incoherence
paradigm
Children and adults

ABSTRACT

Theories about visual narrative understanding accentuate the difference between patterns of reading comprehension in children and adults when they read text and images. This study was conducted to explore the differences in eye movement patterns when children and adults read different comic stories using a coherence/incoherence paradigm. A total of 63 participants, 31 children (10-12 years old) and 32 undergraduate university students from the Universidad Autónoma de Madrid, read 20 comic stories, each of them with both coherent and incoherent versions, for the two ending frames. Fixation durations, number of fixations, and number of regressions were recorded by an eye-tracker, Tobii x-120. A crossed random effects model was applied. Results showed that even though children reach a similar level of understanding than adults they spend more time and have longer fixations than adults, showing more effort to reach the whole comprehension of the stories. Besides, results do not detect significant differences between eye movements' patterns in peak and release for the two groups studied, and therefore both components of the visual narrative grammar are considered equally relevant in the understanding of comics.

Comprendiendo cómics. Una comparación entre niños y adultos a través de un paradigma de coherencia/incoherencia en un estudio de movimientos oculares

RESUMEN

Algunas teorías sobre la comprensión narrativa visual acentúan la diferencia entre los patrones de comprensión lectora en niños y adultos ante la lectura de textos e imágenes. Este estudio se realizó para explorar las diferencias en los patrones de movimientos oculares cuando los niños y adultos leen diferentes historias de cómics utilizando un paradigma de coherencia/incoherencia. Un total de 63 participantes, 31 niños (de 10 a 12 años de edad) y 32 estudiantes universitarios de la Universidad Autónoma de Madrid, leyeron 20 historias de cómics, cada una de ellas con dos finales, coherente e incoherente. La duración de las fijaciones, el número de fijaciones y el número de regresiones fueron registrados en un Tobii x-120. Se aplicó un modelo de efectos aleatorios cruzados. Los resultados mostraron que los niños, a pesar de que obtienen un nivel de comprensión similar al de los adultos, muestran un mayor esfuerzo para alcanzar la comprensión completa de las historias, como lo indican el mayor número y tiempo de las fijaciones. Por otra parte, los resultados tampoco detectan diferencias significativas entre el patrón de movimientos oculares del *peak* y del *release* entre los dos grupos estudiados, por lo que ambos componentes de la gramática narrativa visual se consideran igualmente relevantes en la comprensión de cómics.

Reading comprehension is one of the most complex human activities. It is the product of many processes, strategies, and previous knowledge involved during reading in order to elaborate a complete mental representation (León & Escudero, 2017). This complexity is described by several theoretical models aimed at understanding the cognitive and linguistic processes involved during reading comprehension, such as the Construction-Integration Model (Kintsch, 1988), the Landscape Model (van den Broek, Rapp, & Kendeou, 2005),

the Structure-Building Framework Model (Gernsbacher, 1996), or the Event-Indexing Model (Zwaan & Radvansky, 1998). In these models, cognitive processes are described as mental activities that engage to maintain both, local and global textual coherence, applied to pronouns, sentences in the text, and across larger text distances, but often applied to texts. However, there is little research on image or comics' comprehension, focused on analyzing whether visual comprehension of comics involve cognitive processes, mental

Cite this article as: Martín-Arnal, L. A., León, J. A., van den Broek, P., & Olmos, R. (2019). Understanding comics. A comparison between children and adults through a coherence/incoherence paradigm in an eye-tracking study. *Psicología Educativa*, 25, 127-137. <https://doi.org/10.5093/psed2019a7>

Funding: This study was supported by Grant PSI2013-47219-P from the Ministry of Economic and Competitiveness (MINECO) of Spain, and European Union.

Correspondence: lorena.martin@uam.es; loleleimajere@gmail.com (L. A. Martín-Arnal).

ISSN: 1135-755X/© 2019 Colegio Oficial de Psicólogos de Madrid. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

representations, strategies, or previous knowledge similar to text reading comprehension. Research in the field is also interested in building some models that propose explanations about how visual comprehension of comics and images occur (e.g., Cohn, 2013b; Cohn, Jackendoff, Holcomb, & Kuperberg, 2014).

Some researchers have proposed models that understand cognitive processes on visual narrative comprehension as an organization with a hierarchic constituent structure called Visual Narrative Grammar (VNG) (e.g., Cohn, 2013a; Cohn et al., 2014). For example, Cohn and Bender (2017) claim that “VNG draws an analogy between the structure of sequential images and the structure of sentences; those panels take on functional “grammatical” roles that can be organized into hierarchic constituents” (Cohn & Bender, 2017, p. 2). This VNG resembles other previous models called “story grammars” for verbal narratives stories (see León, 1986, for a complete revision; Mandler & Johnson, 1977; Rumelhart, 1975; Stein & Glenn, 1979; Thorndyke, 1977), which grouped sentences into constituents based on characters’ goal-directed events, as well as temporary and causal relationships among events. VNG establishes an analogy between the sequential frames of a comic and the organization of a written sentence (Cohn, 2013a). More specifically, Cohn (2013b) consider that visual narratives have argued sequences of images—like those in comics—are organized by a narrative “grammar” using constituent structures that go beyond linear coherence relationships between individual images (e.g., Cohn, 2013b; Cohn et al., 2014). Then, VNG posits a structure that follows the subsequent narrative schema: a narratively coherent comic would have four main parts, similar to written stories. These four main parts would be: establisher—beginning of a narrative structure that sets up the context—, initial—beginning of the action, it starts the interaction—, peak—the action itself, the highpoint of the narrative sequence—, and release—conclusion, the consequences of the action that dissolve the narrative tension (Cohn, 2013a). This can be better understood with the help of the next example (Figure 1).

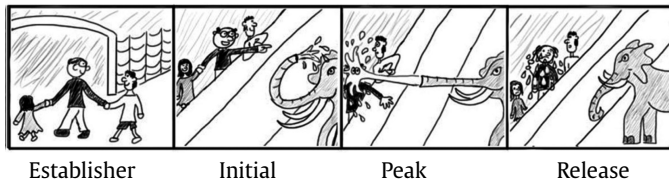


Figure 1. Example of the Structure of a Visual Sequence with four Narrative Constituents.

Another important thing about VNG is that it postulates a merging between narrative structure and meaning that cannot be split up (Cohn et al., 2014; Cohn & Kutas, 2015; Cohn, Paczynski, Jackendoff, Holcomb, & Kuperberg, 2012). Every one of these four parts has its importance. The “release”, for example, gives the final clue to the comprehension of the whole story, as forward connections of events—connections with the consequences of events—tend to have more recall than backwards connections in films (or comics) and in written narratives (León et al., 2015; León & Pérez, 2001; van den Broek & Lorch, 1993; van den Broek, Lorch, & Thurlow, 1996). The “peak” is seen as one of the most important parts, because the action itself happens within it (Cohn et al., 2014). The “establishers” and “initials” are also important, as Gernsbacher (1996) postulated with the Structure Building Framework theory, where the first step to construct comprehension of units that are larger than a word or image is laying a foundation, hence, readers/viewers have to spend more time and number of fixations on these first parts of stories.

Given the importance of these theories for visual understanding, our study will deal with this VNG considering these four components (establisher, initial, peak, and release), where we will

apply a paradigm of coherence vs. incoherence for the two most important components (peak and release).

Eye Movements in Children and Adults when Reading Text and Comics

In the last decade several experiments aimed to describe cognitive processing through eye tracking (e.g., König et al., 2016), which have proven to be a good measure of reading (Rayner, 2009). Eye movements during reading are under direct cognitive control (e.g., Rayner, 1978, 1998). This direct cognitive control is revealed when a reader fixates—which words they skip or regress to and where they land in words—and by analyzing how long they fixate on words. Both provides a great deal of useful information about moment-to-moment processing during reading. As such, our understanding about our reading skill and about how this process is developed has been learnt greatly by monitoring readers’ eye movements and by attempting to infer the basic cognitive processes underlying reading. Moreover, although there are several different paradigms when searching for images studies (Rayner, 2009), very little research analyse deep comprehension with images, particularly if they lean on eye movements control. Thus, the aim of this study dives deep into the relation between age and coherence when adults and children read comic strips without any word.

There is an extensive literature about differences in reading processing when comparing beginners and skilled readers who read sentences or texts. For example, opposing beginners to skilled readers, children who are at the beginning of the reading process have longer fixations, shorter saccades, more re-fixations, and make many regressions (e.g., Huestegge, Radach, Corbic, & Huestegge, 2009; Vorstius, Radach, & Lonigan, 2014). Patterns of eye movements revealed that children (from grades 2 to 6) showed some difficulty decoding words while reading (see Blythe, 2014; Leinenger & Rayner, 2017 for reviews). There is evidence suggesting that developing readers are less efficient at identifying printed words (e.g., Aghababian & Nazir, 2000). However, results of studies using the “disappearing text paradigm” have revealed that children as young as seven years old can capture visual information similarly to skilled adult readers (Blythe, Liversedge, Joseph, White, & Rayner, 2009). Other studies have demonstrated that, even for beginner readers, more exposure to a given word speeds identification and eye movements are under direct cognitive and linguistic control (e.g., Hyönä & Olson, 1995; Joseph, Nation, & Liversedge, 2013). Visual extraction, then, occurs at a normal rate, but beginner readers simply encode less visual information than skilled adult readers on a given fixation. That is to say, they take in visual information at the same rate, but from a reduced region (Leinenger & Rayner, 2017). Results of studies that used the “moving window paradigm” to directly investigate developing readers’ perceptual spans support the idea that developing readers take in less visual information per fixation. These and other authors also say that reading skill influences the perceptual span, since the difficulty of encoding the fixated word could lead to a reduction in the span for beginning readers (Häikiö, Bertram, Hyönä, & Niemi, 2009; Rayner 2009; Rayner & Duffy, 1986). These studies confirm that developing readers have smaller perceptual spans, which reach adult size around fourth grade (Häikiö et al., 2009; Rayner, 1986). A possible explanation for this reduced perceptual span could be that it changes as a function of text difficulty, instead of being a consequence of different eye movements’ patterns between children and adults. Thus, the size of the perceptual span can be reduced when more attention and resources are used to process the foveal word. In fact, skilled readers could seem beginner readers when they are presented texts that are too difficult for them (Rayner, 1986). This suggests that beginner readers have

smaller perceptual spans because of the difficulty they experience processing foveal words (Leinenger & Rayner, 2017).

Regarding reading processes in scene or images perception there are controversial findings due to the different paradigms and the conditions and design controls of the experiments. Across different cultures, readers need congruence when they are trying to understand any story, no matter if it is text, comic, or any other format. There is a need of semantic and structural/syntactic coherence. If one of them is missing, both, text and image reading, can be affected (Cohn, 2014a; Rayner, 2009; Yang, Wang, Chen, & Rayner, 2009).

Some results could be taken from the little research on sequence images or comics. Cohn (2014b), for instance, found a correlation between the ability to order images correctly and the age and experience of the participants when reading comics. Some recent research points out that eyes are guided by the saliency of the scene. That is, they tend to fixate unusual parts of scenes first. As far as researchers have known, beginning readers use to have longer fixations, shorter saccades, and more regressions than skilled readers (Rayner, 2009). There are also differences in patterns of reading comprehension for sentences and images in children and skilled readers (Leinenger & Rayner, 2017). Thus, given that there are not previous research on differences between children and adults, this will be one of the main objectives of this study.

Objective

The aim of this study was to analyse how two frames of comic stories—peak and release—affect online selective processing of story comprehension—e.g., number of fixations, fixation duration, and regressions—through a coherent/incoherent paradigm of short four frames comic strips, and how these patterns relate to strategies children and adults use for reading comprehension. To do so, we will present the participants some four frames comic strips, changing their sense of coherence by varying the relative positions of the two main characters in the two last frames—peak and release. Therefore, our main hypotheses is that reading comprehension—and its eye-tracking patterns—will be affected by both, the age of participants and the coherence of stimuli. Resulting from this idea, there are some main research questions and some derived—also interesting—research questions that we would like to answer:

- A) Are these different patterns of eye movements for comics similar for children than for adults? That is to say, are they universal patterns or specific patterns that differ and change with age?
 1. Will children's eye movement patterns show more effort—more, longer fixations, more regressions—to reach a whole comprehension of our stories than adults?
 2. Will there be a difference between “peak” and “release” patterns due to age differences?
- B) What do we expect from the presentation of these materials with three conditions: “coherent”, “incoherent peak”, and “incoherent release”? Will eye movements' patterns differentiate between conditions? Will they remain independently from age group?
 1. Will incoherent frames be more complex—receive more and longer fixations and more regressions—than coherent frames?
 2. What will happen with coherent frames from incoherent versions? Will they receive the same amount and length of fixations and regressions as those same frames from the “coherent” version? Alternatively, will the incoherence be interacting with them in any way?
 3. Will children be more affected than adults by incoherence—longer and more fixations and regressions for incoherent versions?

Method

Participants

We recruited for this study a total of 63 participants, 32 undergraduate university students from the Universidad Autónoma de Madrid with an age range from 18 to 24 years old (5 male, 15.60%; 27 female, 84.40%), who received a reward for their participation, and 31 children from 10 to 12 years old (12 females, 38.71%; and 19 males, 61.29%), from the Hipatia Fuhem school in Madrid. All of them volunteered for the experiment and had good natural or corrected vision. Their native language was Spanish. All participants gave a written consent according to Autónoma University's Human Subjects Review, and an authorization of the parents and the educational centre-only for the children.

Design

There are three independent variables in our design: coherence—coherent vs. incoherent—, frame—peak vs. release—, which are both within—subjects variables, and age group—adults vs. children—, which is a between-subjects variable. The construct that we are studying, reading comprehension, was operationalized through several DVs: on one hand, four eye-tracking measures as an on-line procedure and, on the other hand, an off-line procedure consisting of a success/error test to check if participants understood stories properly, as described below.

The on-line eye-tracking measures were: mean fixation duration—mean duration (milliseconds, ms) of the fixations on each of the two frames—; number of fixations—the sum of all fixations that participants make on AOI (areas of interest), defined as each of the frames, peak and release—; total fixation time—the total duration of all fixations on a frame, which encompasses both, durations and frequency of fixations—; and number of regressions in—the sum of times that one of the frames was revisited after the first fixation on each one of them.

Second pass fixation time proved to be an important measure of deeper processing—higher order processing—, cognitive and intentional, during reading (Jian, Wu, & Su, 2014); it is important for comprehension (Rayner, 1998) and for integration of information (Mason, Tornatora, & Pluchino, 2013). Moreover, total fixation time on pre-specified areas of interest—like the ones in this experiment—is a useful measure to reveal how readers look at the whole AOI during the entire trial. It shows how much time is spent fixating it (Hyönä, 2010); thus, it was decided to use it to ensure the whole reading/viewing process is englobed in this study about comprehension, that is why total fixation time is included in the analysis. However, first pass fixations are also important for initial processing of features of the images and texts (Jian et al., 2014), that is why mean fixation duration is also included, despite the fact that it could sound redundant together with total fixation time.

The off-line measure was operationalized with a question about the detection of any anomaly in the comic, where it was, and which that anomaly was. Participants were asked to tell if they saw something weird or the whole comic strip made sense.

Materials and Apparatus

An eye-tracker Tobii-x120, which runs at 120 Hz, was used to register eye movements in the present study. Besides, stimuli were presented on a 15" computer screen. There was no need of a chin, because Tobii-x120 allows to move within a 30 x 22 x 30 cm cube—width, height, and depth—, being at 70 cm of distance from the screen.

Stimuli used in this study were ten comic strips created and drawn by us—and three more trial strips—. Each stimulus tells a different

story. The structure of each stimulus follows Cohn's (2013b) Visual Narrative Grammar (VNG) pattern: establisher, initial, peak, and release. All four of them appeared at the same time on the screen, in order, as a comic strip. Coherence was manipulated by changing the position of the characters either on the peak or on the release—thus, creating three coherence conditions: “coherent”, where both frames were correct, “incoherent peak”, where the peak had the change of the character's position, and “incoherent release”, where the release had the change of the character's position. Final stimuli appeared as shown in the examples of Figure 2—the purple arrow did not appear on the presentation of the stimuli.

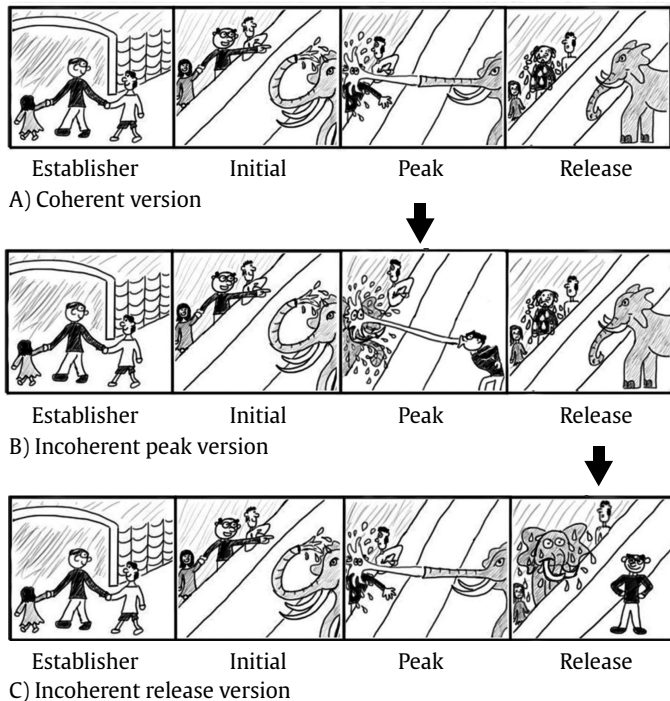


Figure 2. Structure of a Visual Sequence with Narrative Constituents in the Three Versions Analysed in this Study: Coherent Version, Incoherent Peak Version, and Incoherent Release Version.

Procedure

The experiment was administered to participants individually. Each participant, after reading the details of it and signing the informed consent, was seated in front of the visualizing screen, where the corresponding stimuli would be presented. The eye-tracker was situated under this screen, without any obstruction of the participant's line of sight towards it.

When a participant was seated and comfortable, instructions were presented on the screen, with all the time he/she needed to read them, and any doubt about them was solved. They were warned about the three trial stories before the real test. Besides, they were informed that their oral answers would be recorded starting to record at that moment, in order to transcribe them afterwards.

Once there were no doubts, participants were shown their eyes on the screen and they were briefly explained how the machine worked, captured and followed their pupils, allowing them to slightly move, move their eyes, blink, and wink an eye to check it. After that, we proceeded with a nine-dot calibration, which we repeated as many times as needed until it reached the most optimal adjust. When the calibration succeeded, participants were aware of not moving beyond the imaginary window explained before, so the eye-tracker did not lose their gaze.

The next step was beginning the test. The presentation sequence of the stimuli was randomized, as well as the presentation of the coherence version of each stimulus. Therefore, all subjects viewed every stimulus in one of its three versions; besides, subjects viewed all coherence versions, but only one of them for each stimulus. Stimuli appeared on screen as follows:

- The first stimulus appeared for 10 seconds. After several tests performed before the experiment, a mean of 10 seconds was calculated for each stimulus. This time was enough for both children and adults to visualize all the frames, but not enough to let their eyes wander around the screen after finishing the view of the stimuli, messing the data.
- Screen changed into generic questions—“Did you see any anomaly in the story? Where? What was it?”—, which they had to answer aloud. They had all the time they wanted here and, once they had finished their response, participants were asked to press the space bar to make their way to the next story.

Evaluation and Statistical Analysis

Once all the oral answers were transcribed for the off-line measure—correct or incorrect answers depending on the correct/incorrect detection of an anomaly—, they were categorized into: a) “correct”, when they detected an anomaly based on an incoherent frame; b) “incorrect”, when there had been no detection of an anomaly that was there, due to a misunderstanding of the story, or when there had been a detection of an anomaly that was not there, due to other weird or illogical explanations; and c) “ambiguous”, when the anomaly had been detected, but the answer had been illogic or when there had been a detection of an anomaly that was not there, due to a new, logical explanation. In this last case, answers of this kind were less than a 2.50% of the total sample, so the final decision was to treat them as “incorrect”, giving as a result a final categorization of correct/incorrect answers.

Both, these and the data from the four on-line eye movements' measures—namely, mean fixation duration, number of fixations, total fixation time, and number of regressions in—were statistically analysed with the SPSS v23 software, offering results that are shown next.

As we have two random effects, a crossed random mixed model (Hoffman & Rovine, 2007) was conducted to study our coherence, age group, and frame factors. The two crossed random effects were items—comics—and subjects. A total of 15 parameters were estimated. Twelve fixed effects were analysed—the intercept, the age group effect, the frame effect, the coherence effect (two parameters), and the interactions between age group and coherence (two parameters), age group and frame, frame and coherence (two parameters), and age group, frame, and coherence (two parameters) effects. Also, three random effects were estimated: two random intercept variances for items and subjects respectively and level-1 residual variance.

Concerning the distribution of the measures, mean was 3.87—by image— for number of fixations, 347 ms for mean fixation duration, 1,345 ms for total fixation time, and 1.02 for number of regressions in. Standard deviation was 2.14 for number of fixations, 100 ms for mean fixation duration, 809 ms for total fixation time, and .902 for number of regressions in. The asymmetry index was 1.04 for number of fixations, 1.889 for mean fixation duration, 1.05 for total fixation time and .717 for number of regressions in. Hence, there was some indication of positive skew in the distribution of all the measures (i.e., violation of normality), but the amount of skewness was similar for all conditions, and given that it is reasonable to think that it will not constitute an important problem in the estimations of model parameters (Pardo & San Martín, 2015). There were also occasional outliers but not enough to affect the analyses. We conducted analyses

Table 1. Fixed and Random Effects for Age Group, Coherence and Frame for Mean Fixation Duration, Number of Fixations, and Total Fixation

Fixed effect	Mean fixation duration				Number of fixations				Total fixation time				Number of regressions "in"			
	df ₁	df ₂	F	p	df ₁	df ₂	F	p	df ₁	df ₂	F	p	df ₁	df ₂	F	p
Intercept	1	10.10	850.45	.001	1	27.90	1,048.02	.001	1	17.50	539.84	.001	1	16.30	226.19	.001
Age group	1	61.00	23.42	.001	1	61.10	1.13	.292	1	61.10	8.32	.005	1	1,242.30	2.06	.152
Coherence	2	1,182.60	3.77	.023	2	1,170.30	6.03	.002	2	1,168.60	11.86	.001	2	1,238.70	0.72	.489
Frame	1	1,150.30	19.37	.001	1	1,149.20	7.05	.008	1	1,149.20	0.07	.796	1	1,228.80	263.10	.001
Age group * Coherence	2	1,182.40	1.71	.182	2	1,169.60	0.43	.649	2	1,168.30	0.55	.575	2	1,244.00	0.17	.841
Group * Frame	1	1,150.30	0.21	.645	1	1,149.20	0.00	.954	1	1,149.20	0.03	.864	1	1,228.80	2.18	.140
Coherence * Frame	2	1,149.50	19.17	.001	2	1,148.70	92.83	.001	2	1,148.70	118.52	.001	2	1,228.80	7.34	.001
Age group * Coherence * Frame	2	1,149.50	2.00	.135	2	1,148.70	0.46	.634	2	1,148.70	2.14	.119	2	1,228.80	0.42	.658
Random effect	Est.	SE	Wald Z	p	Est.	SE	Wald Z	p	Est.	SE	Wald Z	p	Est.	SE	Wald Z	p
Residual variance	7,666.05	319.95	23.96	.001	3.62	0.15	23.95	.001	496,727.62	20,733.27	23.96	.001	.61	.02	24.79	.001
Intercept subject variance	514.65	165.12	3.12	.002	0.57	0.14	4.17	.001	85,047.46	20,032.93	4.24	.001	.03	.02	1.81	.071
Comic intercept variance	1,349.92	665.46	2.03	.043	0.06	0.04	1.45	.147	27,697.87	14,957.43	1.85	.064	.04	.01	2.90	.004

on both the log-transformed data and the untransformed data with identical patterns of results. Therefore, we have reported the results of analyses in the untransformed data.

Correlations between the four eye tracking variables were also measured for all the frames and conditions, for all the frames dividing the sample into the three different conditions, and for peak and release only, with the total sample and dividing it into the three conditions too. Results always showed high correlations ($p < .001$) between number of fixations and total fixation time, and between mean fixation duration and total fixation time, but not between mean fixation duration and number of fixations. These results were expected, because mean fixation duration (first pass measure) is, in fact, part of total fixation time, and this and the number of fixation are both global measures (first pass and second pass measures).

Results

Success/Error

Beginning with the off-line measure, it starts from very acceptable percentages of success in the answers:

On one hand, adults had a mean of success in the answers of the causal question of 88.43%, their median was 90%, and ranged from 70% to 100%, except for one participant, who succeeded in 60%. This indicates that all but one adult that underscoring beneath 65%—number of successes considered acceptable in previous literature (e.g., León, 2003; León y Escudero, 2003; León y Pérez, 2001; León, Solari, Olmos y Escudero, 2011)—had a high comprehension of the stories.

On the other hand, children obtained a mean of successes of 71.94%, a median of 80%, and ranged from 70% to 100%. In this case, there were seven exceptions, with 30%, 30%, 40%, 40%, 50%, 50% and 60% of success, beneath the acceptable 65%. This data lead us to think that comprehension of the stories was also high for children, except for these seven cases.

Therefore, given that the number of participants with success under an acceptable percentage is very low—eight participants that represent less than a 13% of the total sample—we decided not to drop these participants, in order to have a more realistic sample. Notwithstanding, most of the analyses were replicated, suppressing these eight participants to be sure that they were not adding noise to the data, and no substantial change was found. Hence, it is worth saying that the registered eye-tracking data are attributable to deep comprehension processes in a way that will be explained in further sections.

Eye Movements Results

First of all, the analysis was divided into two groups: a) only the coherent condition was analyzed to see whether there were differences due to the age group independent variable—i.e., between children and adults—and due to the frame independent variable—i.e., between the peak and release means, along with establisher and initial means. Random comic and subjects intercept variances were included in the mixed model, too. The same approach was taken with the analyses of all the measures: mean fixation duration, number of fixations, total fixation time, and number of regressions in—all of them defined in the previous section—, and it seized a baseline with which to b) compare the results obtained from our experimental paradigm with the three set variables. Hence, a second mixed model was conducted for each of the three variables. In this model, three independent variables were analyzed: age group—children and adults—, frame—peak and release—, and coherence—coherent, incoherent peak, and incoherent release. Again, random effects for subjects and comics intercepts were included. All the fixed effects and random effects are shown in Table 1.

Coherent version

Mean fixation duration. Only the age group fixed effect was significantly different from zero, $F(1, 59) = 7.42, p = .008$. The average fixation duration for coherent comics for children was 353 ms ($SE = 9.70$) while for the adults it was 331 ms ($SE = 9.80$). Besides, only marginally significant differences in fixation duration means were found between establisher ($M = 343, SE = 10.40$), initial ($M = 331, SE = 10.30$), peak ($M = 340, SE = 10.30$) and release ($M = 354, SE = 10.30$) frames, $F(3, 727) = 2.52, p = .057$. There was no significant interaction found between age group and frame effects, $F(3, 727) = 1.28, p = .283$. The random effects showed that subjects intercept variance was significant ($\sigma^2_{\text{subject}} = 409.880, \text{Wald } Z = 2.280, p = .023$) and comic intercept variance was marginally significant ($\sigma^2_{\text{comic}} = 625.146, \text{Wald } Z = 1.796, p = .072$). The estimated residual variance was 7,282.223 ($\text{Wald } Z = 19.027, p < .001$). Mean fixation durations are shown in Figure 3.

Number of fixations. Results only showed a significant main effect of the frame condition, $F(3, 735) = 22.16, p < .001$. The release ($M = 4.33, SE = 0.1$) received on average more fixations than the peak ($M = 3.95, SE = 0.1$). There was only a marginally significant effect for the age group variable, $F(1, 63) = 3.45, p = .068$. There was no significant effect for the interaction between age group and frame conditions, $F(3, 735) = .35, p = .790$. The random effects showed that subjects intercept variance was significant ($\sigma^2_{\text{subject}} = .196, \text{Wald } Z = 2.316, p = .021$) and comic intercept variance was not significant ($\sigma^2_{\text{comic}} = 0, \text{Wald } Z = 0, p = 1$). The estimated residual variance was 3.494 ($\text{Wald } Z = 19.133, p < .001$). Mean fixation durations are shown in Figure 4.

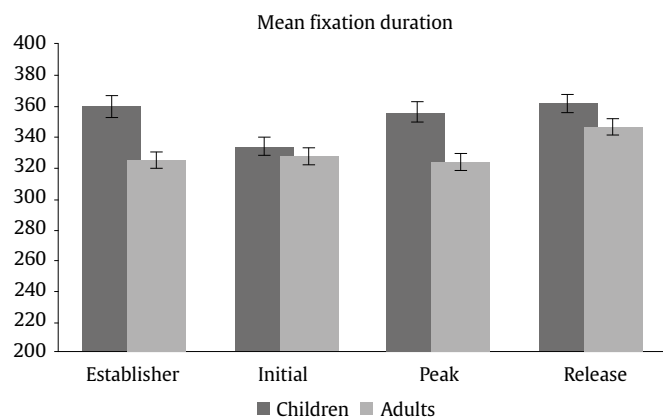


Figure 3. Mean Fixation Durations for Children and Adults and All the Four Frames for Coherent Stories: 1 (establisher), 2 (initial), 3 (peak), and 4 (release).

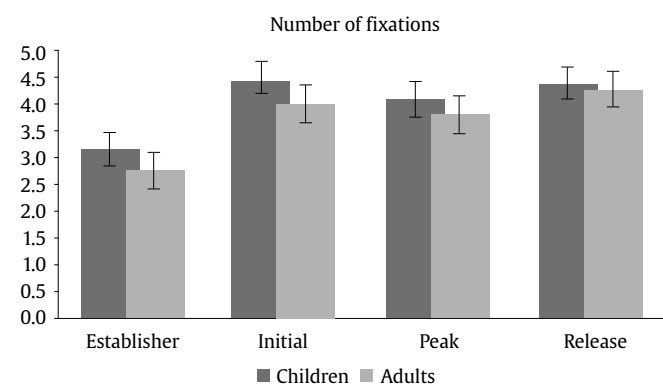


Figure 4. Number of Fixations for Children and Adults and all the Four Frames for the Coherent Condition: 1 (establisher), 2 (initial), 3 (peak), and 4 (release).

Total fixation time. Results showed significant main effects of the frame variable, $F(3, 727) = 18.63, p < .001$, and the age group variable, $F(1, 60) = 10.17, p = .002$. For the frame, the release ($M = 1,521, SE = 53.10$) showed more total fixation time (ms) on average than the initial ($M = 1,406, SE = 53.40$), the peak ($M = 1,337, SE = 53$), and the establisher ($M = 1,018, SE = 54.50$), in that order. For the age group, children ($M = 1,419, SE = 44.30$) showed more total time (ms) on average than adults ($M = 1,222, SE = 45$). There was no significant effect of the interaction between age group and frame conditions, $F(3, 726) = .33, p = .807$. The random effects showed that subjects intercept variance was marginally significant ($\sigma^2_{\text{subject}} = 21,061.070$, Wald $Z = 1.940, p = .052$) and comic intercept variance was not significant ($\sigma^2_{\text{comic}} = 750.201$, Wald $Z = 0.219, p = .827$). The estimated residual variance was 483,709.602 (Wald $Z = 19.016, p < .001$). Total fixation times are shown in Figure 5.

Regressions in (times revisited). Results only showed significant main effects of the frame variable, $F(1, 380) = 73.44, p < .001$, where the peak ($M = 1.38, SE = .1$) showed more regressions in than the release ($M = .70, SE = .1$). There was no other significant effect, neither of the age group condition, $F(1, 407) = 0.37, p = .54$, nor of the interaction between age group and frame conditions, $F(1, 380) = 2.25, p = .13$. The random effects showed that subjects intercept variance was significant ($\sigma^2_{\text{subject}} = 0.055$, Wald $Z = 2.067, p = 0.039$) and comic intercept variance was not significant ($\sigma^2_{\text{comic}} = 0.033$, Wald $Z = 1.417, p = .157$). The estimated residual variance was 0.656 (Wald $Z = 13.789, p < .001$). The number of regressions in is shown in Figure 6.

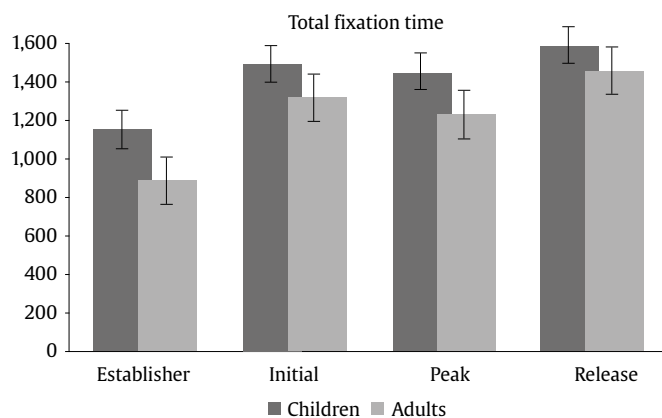


Figure 5. Total Fixation Time for Children and Adults and all the Four Frames for the Coherent Condition: 1 (establisher), 2 (initial), 3 (peak), and 4 (release).

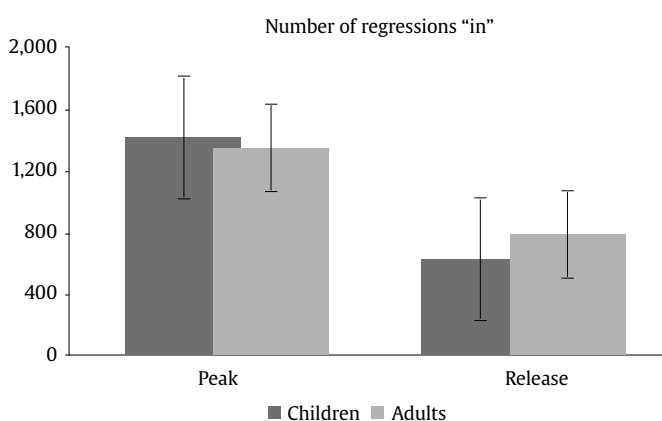


Figure 6. Number of Regressions in for Children and Adults and the Two Final Frames for the Coherent Conditions: 3 (peak), and 4 (release).

Complete analysis version (coherence vs. incoherence)

Mean fixation duration. The main effect of age group was significant ($p < .001$), as children's fixation durations ($M = 375, SE = 12.80$) were longer than adults' fixation durations ($M = 338, SE = 12.80$). The interaction effect between coherence and frame was also significant ($p < .001$). The three-factor interaction was not significant, $F(2, 1150) = 2.00, p = .135$. Table 2 and Figure 7 show the results.

Table 2. Means and SE for Age Group, Coherence, and Frame for Mean Fixation Duration

Group	Coherence	Frame	Mean (SE)
Children	Coherent	Peak	356 (15.00)
		Release	362 (15.10)
	Incoherent peak	Peak	385 (15.00)
		Release	375 (15.10)
	Incoherent release	Peak	347 (15.00)
		Release	425 (15.10)
Adults	Coherent	Peak	323 (15.10)
		Release	347 (15.10)
	Incoherent peak	Peak	346 (14.90)
		Release	334 (15.10)
	Incoherent release	Peak	316 (15.00)
		Release	363 (15.00)

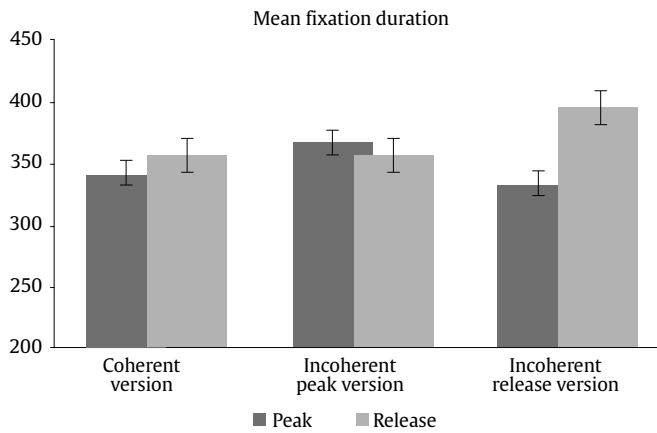


Figure 7. Mean Fixation Durations for Children and Adults together and the Two Final Frames for the Three Coherent Conditions: Peak (dark grey column) and Release (light grey column).

The interaction effect was clear as incoherent peak showed a different pattern than incoherent release. Simple effects showed that peak and release fixation duration means did not differ significantly in the coherent and incoherent peak conditions, while they did significantly differ in the incoherent release condition. The mean difference for the coherent condition was 15 ($p = .090$), being the release the frame with longer durations. The mean difference for the incoherent peak condition was 12 ($p = .176$), being the peak the one with longer fixations. The differences between coherent and incoherent peak conditions only differ in 3 ms. However, the mean difference for the incoherent release condition was 63 ($p < .001$), being the release the one with longer fixations and longer than any other frames in any other conditions. The differences between coherent and incoherent release conditions differ in 48 ms. Thus, comparing the differences between both frames using the coherent condition as a baseline, the distance with incoherent release condition is higher than that with incoherent peak condition.

Number of fixations. The interaction effect between coherence and frame was significant again ($p < .001$). The three-factor interaction was not significant, $F(2, 1149) = 0.46$; $p = .634$. Table 3 and Figure 8 show the results.

Table 3. Means and SE for Age Group, Coherence, and Frame for Number of Fixations

Group	Coherence	Frame	Mean (SE)
Children	Coherent	Peak	4.1 (0.20)
		Release	4.4 (0.20)
	Incoherent peak	Peak	5.9 (0.20)
		Release	3.5 (0.20)
	Incoherent release	Peak	4.0 (0.20)
		Release	5.3 (0.20)
Adults	Coherent	Peak	3.8 (0.20)
		Release	4.3 (0.20)
	Incoherent peak	Peak	5.5 (0.20)
		Release	3.2 (0.20)
	Incoherent release	Peak	4.0 (0.20)
		Release	5.0 (0.20)

The interaction effect was clear again. Simple effects showed that number of fixations on peak and release differed significantly in the coherent, incoherent peak, and incoherent release conditions. However, these differences were not under the same significance level: while the p level for the coherent condition was $p = .046$ (mean difference = .38, $p < .05$), it was $p < .001$ for both the incoherent conditions (mean difference for the incoherent peak = 2.33 and mean

difference for the incoherent release = 1.1). Besides, as we can see in Figure 8, despite the significances of the differences, the pattern, again, was: for the coherent condition and the incoherent release condition, the release was the frame with more number of fixations, while for the incoherent peak condition the opposite was happening, it was the peak the one with more fixations. Also, the distance between the difference in coherent and incoherent peak conditions (1.95 fixations) is higher than that in coherent and incoherent release conditions (0.63 fixations).

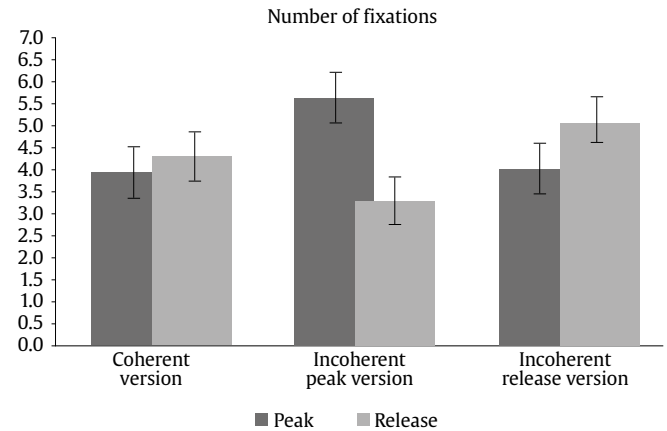


Figure 8. Number of Fixations for Children and Adults together and the Two Final Frames for the Three Coherence Conditions: Peak (dark grey column) and Release (light grey column).

Total fixation time. The main effect of age group was significant ($p = .005$) once again, as children's total fixation time ($M = 1,684$, $SE = 79.50$) was longer than adults' total time ($M = 1,442$, $SE = 79$). The interaction effect between coherence and frame was significant again ($p < .001$). The three-factor interaction was not significant, $F(2, 1149) = 2.14$; $p = .119$. Table 4 and Figure 9 show the results.

Table 4. Means and SE for Age Group, Coherence, and Frame for Total Fixation Time

Group	Coherence	Frame	Mean (SE)
Children	Coherent	Peak	1,448 (101.70)
		Release	1,585 (102.00)
	Incoherent peak	Peak	2,202 (101.70)
		Release	1,301 (102.00)
	Incoherent release	Peak	1,406 (101.70)
		Release	2,160 (102.80)
Adults	Coherent	Peak	1,217 (102.50)
		Release	1,449 (102.40)
	Incoherent peak	Peak	1,860 (100.40)
		Release	1,071 (102.40)
	Incoherent release	Peak	1,276 (101.50)
		Release	1,781 (101.30)

The interaction effect was once more, very clear. Simple effects showed that peak and release fixation duration means differ significantly in the coherent (mean difference = 184, $p = .009$), incoherent peak (mean difference = 845, $p < .001$) and incoherent release (mean difference = 629, $p < .001$) conditions. Besides, as we can see from Figure 9, the pattern was the same as for mean fixation duration and number of fixations: for the coherent condition and the incoherent release condition, the release was the frame with more total fixation time, while for the incoherent peak condition it was the peak the one with a higher total time. Once more, comparing the differences between both frames using the coherent condition

baseline, the distance with incoherent peak condition (661 ms) is higher than that with incoherent release condition (445 ms).

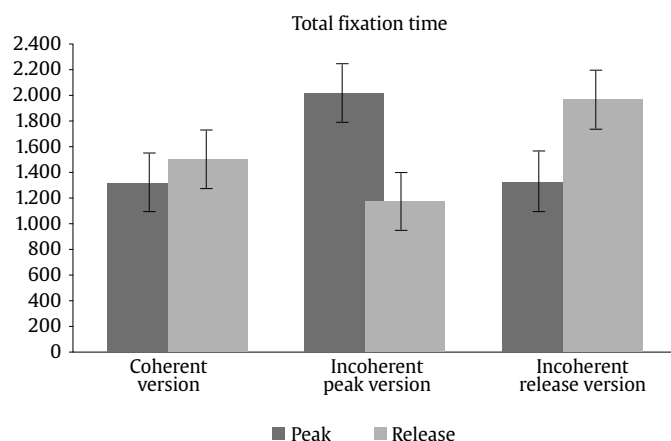


Figure 9. Total Fixation Time for Children and Adults together and the Two Final Frames for the Three Coherence Conditions: Peak (dark grey column) and Release (light grey column).

Regressions in (times revisited). Only frame main effect was significantly different from zero, $F(1, 1229) = 263.10$, $p < .001$. However, non-significant differences in number of regressions in were found, nor for age group condition, $F(1, 1242) = 2.06$, $p = .152$, neither for coherence condition, $F(2, 1239) = 0.72$, $p = .489$. There was a significant interaction between coherence and frame, $F(2, 1229) = 7.34$, $p = .001$. The three-factor interaction was not significant, $F(2, 1229) = 0.42$, $p = .658$. The estimated means and standard deviations (SD) are shown in Table 5 and Figure 10.

Table 5. Means and SE for Age Group, Coherence, and Frame for Number of Regressions in

Group	Coherence	Frame	Mean (SE)
Children	Coherent	Peak	1.4 (0.10)
		Release	0.6 (0.10)
	Incoherent peak	Peak	1.4 (0.10)
		Release	0.5 (0.10)
	Incoherent release	Peak	1.3 (0.10)
		Release	0.7 (0.10)
Adults	Coherent	Peak	1.4 (0.10)
		Release	0.8 (0.10)
	Incoherent peak	Peak	1.4 (0.10)
		Release	0.5 (0.10)
	Incoherent release	Peak	1.3 (0.10)
		Release	0.8 (0.10)

The interaction effect, as shown in Figure 10, was clear as incoherent peak showed a different pattern than incoherent release. Simple effects showed that all differences between peak and release were significantly different for the three coherence conditions. However, this difference was higher for the incoherent peak condition ($M = 0.93$, $SE = 0.10$) than for the other two conditions, coherent ($M = 0.68$, $SE = 0.10$) and incoherent release ($M = -0.53$, $SE = 0.10$), being the peak higher than the release in all conditions. This time, using the coherent condition as a baseline, the distance with incoherent peak condition (0.25 regressions) is less than that with incoherent release condition (1.21 regressions).

Curiously, neither coherent peak differed from incoherent peak ($MD = 0.07$, $p = .367$), nor coherent release differed from incoherent release ($MD = 0.07$, $p = .367$). However, incoherent peak significantly differed from coherent peak of the incoherent release

condition ($MD = 0.15$, $p = .048$), being higher the NoRI—number of regressions in—for the peak on incoherent peak condition, and incoherent release significantly differed from coherent release of the incoherent peak condition ($MD = 0.26$, $p < .001$), being higher the NoRI on release of the incoherent release condition. In addition, coherent release significantly differed from coherent release of the incoherent peak condition ($MD = 0.19$, $p = .015$), being higher the NoRI for the release on coherent condition.

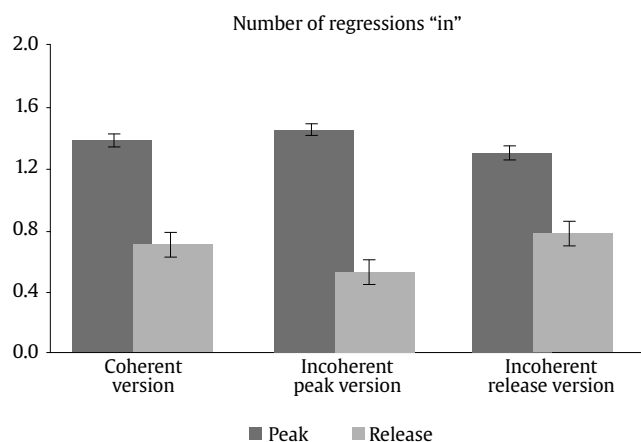


Figure 10. Number of Regressions in (NoRI) for Children and Adults together and the Two Final Frames for the Three Coherence Conditions: Peak (dark grey column) and Release (light grey column).

Conclusion and Implications

Some interesting patterns regarding the “coherent” version only—which means normal comic reading—were found. First finding is that children, despite they reach very good percentages of success and almost the same level of understanding than adults—72% vs. 88%—, spend more time when reading. Thus, their reading abilities seem to be less efficient, as happens with beginner readers. Therefore, to answer our first research question, children had more and longer fixations than adults, so it seemed that their eye movements’ patterns showed more effort to reach the whole comprehension of the stories.

Furthermore, this variable did not interact with the others in any of the analyses that were made. This gives us some clues: first, to address the second research question, there is no difference between eye movements patterns on peak and release due to the age differences of the two participant groups; second, this difference between children and adults remains the same for all the coherence versions, which leads to the conclusion that children are sensitive to coherence changes in the same way as adults are. Therefore, taking in account for a moment the coherence conditions, this, as an answer to the fifth research question, points out that children and adults are not affected differently by incoherence. Thus, this experiment seems to be proving that there is a control of eye movements in order to reach the best understanding of the stories, like some authors already argued (Luke, Nuthmann & Henderson, 2013; Schotter, Tran, & Rayner, 2014) and we advanced in previous sections of this paper. Nevertheless, it is a new—and surprising—finding that this fact is similar for both, children—assuming that they are less skilled—and adults—assumed to be more skilled—, and that there is no difference at all between them. Of course, these conclusions were drawn from the absence of interaction between conditions of the age group and the other variables’ conditions, so they should be taken carefully.

It is not so central, but also worth mentioning that establishments or initials—never both of them but always one—, are often among

the most and longest visited frames, so it seems that people need to figure out the beginning of each story along with the release of them. These findings agree with some authors' point of view (e.g., Gernsbacher, 1996; Gernsbacher, Varner, & Faust, 1990). Gernsbacher (1996) postulated a theory about comprehension to build cohesive mental representations called the Structure Building Framework. According to this theory, the first step to construct comprehension of units larger than a word or image is laying a foundation. To do this, readers and viewers spend more time reading the first words of clauses, sentences, or paragraphs or viewing the first images of comic strips or picture stories (Gernsbacher, 1996). Thus, children and adults of this experiment could be setting this foundation of the representation in memory by spending more time and number of fixations on the establishers or the initials. This is also coherent with the "advantage of first mention" idea (Carreiras, Gernsbacher, & Villa, 1995), where the first mentioned participant is accessed first, regardless of its semantic role or syntactic position. Moreover, Gernsbacher (1996) also defends that this effect does not happen when the coherence of the story is called into question. Yet in this study, despite not being echoed in the Results section not to make them too long, no increment in the number or length of the fixations was found in establishers and initials when the coherence was manipulated, so this importance of first ideas seems to be a relatively stable effect.

These patterns of eye movements could set a baseline where the release of the story seems to lead to some wrap-up processes, giving the conclusion—and "releasing" the tension—to the rest of the story, like they release the tension in Cohn's (2013b) comics, as the author finds out. In van den Broek et al. (1996), they found for films that for adults the forward connections an event has have a stronger impact on the recall of events than backward connections. This means that the connections of an event with its consequences influence memory more than the connections with its antecedents; same happens for written narratives (van den Broek & Lorch, 1993). This could be in line with the findings about the release of the story as a conclusive process, where participants try to give a global sense to the whole story. Still, going one step further, an important finding in this study is that the processing of the peak and release, respectively, does not occur in isolation but, in fact, the processing of the one is influenced by the processing of the other. Therefore, conclusions about these components should take into account that they are mutually dependent and they interact. This statement will be detailed in further paragraphs.

Concerning the "complete version" analysis, there were also some interesting findings.

With regard to fixations (duration and frequency), it was found that the coherent/incoherent paradigm has an effect over the eye movement patterns of people. All incoherent versions increase their number and duration of fixations for their incoherent frames (peak or release), as expected from the beginning so, to address the third research question, incoherent frames seem to be more complex than coherent frames. Besides, what we find more surprising about this new outcome is the effect of the incoherent frames over the processing of other parts or frames of the stories. Some coherent frames in their incoherent versions of the story decrease in number and duration of fixations, what leads to answer the fourth research question, as the incoherence interacts with the coherent frames in such a way.

Thus, incoherence draws reader's attention, indeed, as it would be expected; nevertheless, it also takes attention out of the other frames in the incoherent peak version. It is also worth mentioning that, for incoherent release, the effect is not significant, although it follows the same direction. Therefore, it may be thought that this could be an effect of the sample size and maybe with a larger sample this effect could be significant. However, it could also simply be as it is: the effect appears for the incoherent peak version and not for the incoherent release version, but there is not enough data to answer

these questions, so they will be addressed in further research.

As far as data show, then, people's eye movements patterns tend to focus on semantically incoherent parts of the comic strips, captivating their eyes more and longer than to other coherent parts of the scenes. Thus, as previous authors pointed out (Becker, Pashler, & Lubin, 2007; Clifton, Staub, & Rayner, 2007; Harris, Nefs, & Grafton, 2008; Rayner, 2009; Staub & Rayner, 2007), semantical incongruence attracts viewers' attention. Implausible events yield longer times (Joseph et al., 2008; Rayner et al., 2004), as it happens in this study, even if the context is a fantasy-like context (Warren, McConnell, & Rayner, 2008). Besides, ambiguous information gets longer fixations (Altmann, Garnham, & Dennis, 1992; Rayner & Sereno, 1994). Even more, some studies about errors and motivation (Cimpian, 2010; Zentall & Morris, 2012) show that the threat of making mistakes raise the number of fixations people make. Maybe this could be an explanation of the high number of fixations on the incoherent information, as people want to be sure about what they are viewing, not to make any mistakes when asked.

Completing the discussion about the study, results from regressions will be now approached. As can be seen from data, peak always receives more regressions than release. This may be because release is the last picture, so participants do not have to leave it to finish viewing, as it happens with peak. Furthermore, despite all the changes that will be next mentioned, regressions' patterns seem to be relatively stable, as Rayner (2009) and other authors previously mentioned, as they seem to keep this general pattern regardless of these small differences we mention.

However, what happens when we change the coherence of the story structure? Are regressions also sensitive to these changes, as it happens with fixations? Results seem to show that, indeed, incoherence takes attention, again, out of the other frames for the incoherent peak condition. However, this does not happen for incoherent release condition, but the effect, once more, follows the same direction. Incoherent peak/release do not receive more regressions than their coherent versions, though they receive more than coherent peak/release from the opposite incoherent conditions; consequently, changing the coherence of both makes the other receive fewer regressions than usual, again. Besides, it is also important to mention that when the peak is incoherent, the release receives fewer regressions than the other two releases. Therefore, the incoherence of both pictures disturbs the reading patterns, so both affect comprehension, as ambiguous information constrain people to make more regressions, like in previous theory (Altmann et al., 1992; Rayner & Sereno, 1994). Thus, for "regressions in", there might be an interaction of both incoherent conditions, but only when something more is incoherent. However, there is a new discovery, not so central, but also interesting: only incoherence of peak affected release in such a way that eye movement patterns are even below its "usual" "coherent" condition; therefore, peak may be moderating the eye movements patterns on the release in a different way that vice versa.

This experiment was started trying to understand reading patterns and see if they differ between children and adults while reading comics, following a coherent/incoherent paradigm. Our objective was to find eye-tracking patterns that were accurate, reliable, and related to reading comprehension and higher order cognitive processes.

Some of the outcomes we reached with this experiment show surprising conclusions and we have found significant results supporting this general idea of eye movements as a measure that is sensitive to narrative coherence and to different maturity levels. Thus, eye movements develop and change while people grow up, and respond to some high level cognitive strategies, reflecting different processes of deep comprehension. However, this is a new path that researchers have not come to explore deeply. There is not much information about it, yet we hope to have paved the way or, at least, shed a bit more light on reading comprehension and eye movements' research.

Conflict of Interest

The authors of this article declare no conflict of interest.

References

- Aghababian, V., & Nazir, T. A. (2000). Developing normal reading skills: Aspects of the visual processes underlying word recognition. *Journal of Experimental Child Psychology*, 76(2), 123-150. <https://doi.org/10.1006/jecp.1999.2540>
- Altmann, G. T., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language*, 31, 685-712. [https://doi.org/10.1016/0749-596x\(92\)90035-v](https://doi.org/10.1016/0749-596x(92)90035-v)
- Becker, M. W., Pashler, H., & Lubin, J. (2007). Object-intrinsic oddities draw early saccades. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 20-30. <https://doi.org/10.1037/0096-1523.33.1.20>
- Blythe, H. I. (2014). Developmental changes in eye movements and visual information encoding associated with learning to read. *Current Directions in Psychological Science*, 23, 201-207. <https://doi.org/10.1177/0963721414530145>
- Blythe, H. I., Liversedge, S. P., Joseph, H. S., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research*, 49, 1583-1591. <https://doi.org/10.1016/j.visres.2009.03.015>
- Carreiras, M., Gernsbacher, M. A., & Villa, V. (1995). The advantage of first mention in Spanish. *Psychonomic Bulletin & Review*, 2, 124-129. <https://doi.org/10.3758/bf03214418>
- Cimpian, A. (2010). The impact of generic language about ability on children's achievement motivation. *Developmental Psychology*, 46, 1333-1340. <https://doi.org/10.1037/a0019665>
- Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye Movements* (pp. 341-371). Amsterdam, The Netherlands: Elsevier. <https://doi.org/10.1016/b978-008044980-7/50017-3>
- Cohn, N. (2013a). *The visual language of comics: Introduction to the structure and cognition of sequential images*. London, UK: Bloomsbury. <https://doi.org/10.5040/9781472542175>
- Cohn, N. (2013b). Visual narrative structure. *Cognitive Science*, 37, 413-452. <https://doi.org/10.1111/cogs.12016>
- Cohn, N. (2014b). The architecture of visual narrative comprehension: The interaction of narrative structure and page layout in understanding comics. *Frontiers in Psychology*, 5, 1-9. <https://doi.org/10.3389/fpsyg.2014.00680>
- Cohn, N. (2014a). You're a good structure, Charlie Brown: The distribution of narrative categories in comic strips. *Cognitive Science*, 38, 1317-1359. <https://doi.org/10.1111/cogs.12116>
- Cohn, N., & Bender, P. (2017). Drawing the line between constituent structure and coherence relations in visual narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 289-301. <https://doi.org/10.1037/xlm0000290>
- 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., 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>
- Gernsbacher, M. A. (1996). The structure-building framework: What it is, what it might also be, and why. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 289-311). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 430-445. <https://doi.org/10.1037/0278-7393.16.3.430>
- Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology*, 102(2), 167-181. <https://doi.org/10.1016/j.jecp.2008.04.002>
- Harris, J. M., Nefs, H. T., & Grafton, C. E. (2008). Binocular vision and motion-in-depth. *Spatial Vision*, 21, 531-547. <https://doi.org/10.1163/156856808786451462>
- Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods*, 39, 101-117. <https://doi.org/10.3758/BF03192848>
- Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research*, 49, 2948-2959. <https://doi.org/10.1016/j.visres.2009.09.012>
- Hyönä, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction*, 20(2), 172-176. <https://doi.org/10.1016/j.learninstruc.2009.02.013>
- Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 1430-1440. <https://doi.org/10.1037/0278-7393.21.6.1430>
- Jian, Y. C., Wu, C. J., & Su, J. H. (2014). Learners' eye movements during construction of mechanical kinematic representations from static diagrams. *Learning and Instruction*, 32, 51-62. <https://doi.org/10.1016/j.learninstruc.2014.01.005>
- Joseph, H. S., Liversedge, S. P., Blythe, H. I., White, S. J., Gathercole, S. E., & Rayner, K. (2008). Children's and adults' processing of anomaly and implausibility during reading: Evidence from eye movements. *The Quarterly Journal of Experimental Psychology*, 61, 708-723. <https://doi.org/10.1080/17470210701400657>
- Joseph, H. S., Nation, K., & Liversedge, S. P. (2013). Using eye movements to investigate word frequency effects in children's sentence reading. *School Psychology Review*, 42, 207-222.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95, 163-182. <https://doi.org/10.1037/0033-295x.95.2.163>
- König, P., Wilming, N., Kietzmann, T. C., Ossandón, J. P., Onat, S., Ehinger, B. V., ... Kaspar, K. (2016). Eye movements as a window to cognitive processes. *Journal of Eye Movements Research*, 9(5), 1-16. <https://doi.org/10.16910/jemr.9.5.3>
- Leininger, M., & Rayner, K. (2017). What we know about skilled, beginning, and older readers from monitoring their eye movements. In J. A. León & I. Escudero (Eds.), *Reading comprehension in educational settings. Series of studies in written language and literacy* (pp. 1-27). Amsterdam, The Netherlands: John Benjamins Publishing Company. <https://doi.org/10.1075/swll.1601lei>
- León, J. A. (1986). *La memoria de los niños a través de los cuentos: un análisis experimental*. Madrid, Spain: UNED.
- León, J. A. (2003). *Conocimiento y discurso: claves para inferir y comprender*. Madrid, Spain: Pirámide.
- León, J. A., Dávalos, M. T., Escudero, I., Olmos, R., Morera, Y., & Froufe, M. (2015). Effects of valence and causal direction in the emotion inferences processing during reading: Evidence from a lexical decision task. *Anales de Psicología*, 31, 677-686. <https://doi.org/10.6018/analesps.31.2.167391>
- León, J. A., & Escudero, I. (2003). Protocolos verbales en el estudio de las inferencias: Una metodología emergente. In J. A. León (Coord.) (2003), *Conocimiento y discurso. Claves para inferir y comprender*, (pp. 100-119). Madrid, España: Pirámide.
- León, J. A., & Escudero, I. (2017). Causal inferences in reading comprehension: State of the art. In J. A. León & I. Escudero (Eds.) (2017), *Reading comprehension in educational settings. Series of studies in written language and literacy* (pp. 63-99). Amsterdam, The Netherlands: John Benjamins Publishing Company. <https://doi.org/10.1075/swll.1603leo>
- León, J. A., & Pérez, O. (2001). The influence of prior knowledge on the time course of clinical diagnosis inferences: A comparison of experts and novices. *Discourse Processes*, 31, 187-213. https://doi.org/10.1207/S15326950DP3102_04
- León, J. A., Solari, M., Olmos, R., & Escudero, I. (2011). La generación de inferencias dentro de un contexto social. Un análisis de la comprensión lectora a través de protocolos verbales y una tarea de resumen oral. *Revista de Investigación Educativa*, 29, 13-42.
- Luke, S. G., Nuthmann, A., & Henderson, J. M. (2013). Eye movement control in scene viewing and reading: Evidence from the stimulus onset delay paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, 39, 10-15. <https://doi.org/10.1037/a0030392>
- Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 9(1), 111-151. [https://doi.org/10.1016/0010-0285\(77\)90006-8](https://doi.org/10.1016/0010-0285(77)90006-8)
- Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education*, 60(1), 95-109. <https://doi.org/10.1016/j.compedu.2012.07.011>
- Pardo, A., & San Martín, R. (2015). *Análisis de datos en Psicología III*. Madrid, Spain: Pirámide.
- Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, 41, 211-236. [https://doi.org/10.1016/0022-0965\(86\)90037-8](https://doi.org/10.1016/0022-0965(86)90037-8)
- Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin*, 85, 618-660. <https://doi.org/10.1037/0033-2909.85.3.618>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372. <https://doi.org/10.1037/0033-2909.124.3.372>
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 62, 1457-1506.
- Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical

- ambiguity. *Memory & Cognition*, 14, 191-201. <https://doi.org/10.3758/bf03197692>
- Rayner, K., & Sereno, S. C. (1994). Regressive eye movements and sentence parsing: On the use of regression-contingent analyses. *Memory & Cognition*, 22, 281-285. <https://doi.org/10.3758/bf03200855>
- Rumelhart, D. E. (1975). Notes on a schema for stories. In D. E. Rumelhart, D. G. Bobrow, & A. Collins (1975), *Representation and understanding*. (pp. 211-236). New York, NY: Academic Press. <https://doi.org/10.1016/b978-0-12-108550-6.50013-6>
- Schotter, E. R., Tran, R., & Rayner, K. (2014). Don't believe what you read (only once) comprehension is supported by regressions during reading. *Psychological Science*, 25, 1218-1226. <https://doi.org/10.1177/0956797614531148>
- Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In M. G. Gaskell (Ed.) (2007). *The Oxford handbook of psycholinguistics* (pp. 327-342). New York, NY: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198568971.013.0019>
- Stein, N. L. & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *New directions in discourse processing* (pp. 53-119). Norwood, NJ: Ablex.
- Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 9(1), 77-110. [https://doi.org/10.1016/0010-0285\(77\)90005-6](https://doi.org/10.1016/0010-0285(77)90005-6)
- Van den Broek, P., & Lorch Jr, R. F. (1993). Network representations of causal relations in memory for narrative texts: Evidence from primed recognition. *Discourse Processes*, 16, 75-98. <https://doi.org/10.1080/01638539309544830>
- Van den Broek, P., Lorch, E. P., & Thurlow, R. (1996). Children's and adults' memory for television stories: The role of causal factors, story-grammar categories, and hierarchical level. *Child Development*, 67, 3010-3028 <https://doi.org/10.2307/1131764>
- Van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes*, 39, 299-316. <https://doi.org/10.1080/0163853x.2005.9651685>
- Vorstius, C., Radach, R., & Lonigan, C. J. (2014). Eye movements in developing readers: A comparison of silent and oral sentence reading. *Visual Cognition*, 22, 458-485. <https://doi.org/10.1080/13506285.2014.881445>
- Warren, T., McConnell, K., & Rayner, K. (2008). Effects of context on eye movements when reading about possible and impossible events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1001-1010. <https://doi.org/10.1037/0278-7393.34.4.1001>
- Yang, J., Wang, S., Chen, H. C., & Rayner, K. (2009). The time course of semantic and syntactic processing in Chinese sentence comprehension: Evidence from eye movements. *Memory & Cognition*, 37, 1164-1176. <https://doi.org/10.3758/mc.37.8.1164>
- Zentall, S. R., & Morris, B. J. (2012). A critical eye: Praise directed toward traits increases children's eye fixations on errors and decreases motivation. *Psychonomic Bulletin & Review*, 19, 1073-1077. <https://doi.org/10.3758/s13423-012-0294-y>
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162-185. <https://doi.org/10.1037/0033-2909.123.2.162>