Adding Suspense to a Story Generation System through a Cognitive Model of the Impact of Affective Terms

Pablo Delatorre[1], Carlos León[2], Manuel Palomo-Duarte[1], and Pablo Gervás[2]

[1] University of Cadiz, Spain
pablo.delatorre@uca.es, manuel.palomo@uca.es
http://esingenieria.uca.es/
[2] University Complutense of Madrid, Spain
cleon@ucm.es, pgervas@sip.ucm.es
http://nil.fdi.ucm.es/

Abstract. Suspense is known to play a crucial role in storytelling phenomena in general, and computational storytelling systems in particular. While several story generation algorithms have addressed suspense, they have usually done so either by focusing on the narratology-related aspects, or by providing functional approximations of the potential response. In this paper, following the overall objective of devising a cognitive model of suspenseful stories, we have adapted a story generation system so that it includes elements represented by affective terms extracted from experiments. In this way, we put the focus on the cognitive aspects. The obtained results suggest that the generation model works and that evaluators perceive levels of suspense comparable to the ones obtained in the previous contributions.

Keywords: suspense, story generation, computational creativity, cognitive modeling

1 Introduction

The long-standing goal of achieving creative behavior in an automatic story generation system requires, at least, mimicking aspects of narrative that humans commonly use and thus expect when they experience storytelling. Among these aspects, suspense is one of the most prominent and effective ones. Actually, it is a key factor in a wide range of narrative media: together with coherence and thematic complexity, suspense explains 54% of the variance in interest of a narrative, making the single greatest contribution explaining roughly 34% [30, p. 436, 444].

Being such an important aspect, several computational systems producing suspenseful stories have been created. Many of them have focused on partial perspectives. They mostly tend to focus on cognitive-related aspects by analyzing the escape strategy of the protagonists, as can be seen in the work of Gerrig & Bernardo (1994) [15]–. For example, the unawareness of the threat [34, p. 95], the

proximity [5, p. 73] and importance [36, p. 63] of the outcome; or the environment features [33, p. 4], among others. From the point of the narrative theories, those and other properties work together to evoke the emotion of suspense. We believe that in order to build robust suspenseful story generation systems, all these cognitive aspects must be leveraged and be the focus of the computational model.

With this objective in mind, we have previously proposed an architecture that tries to address the cognitive aspects of suspense as a whole [10]. Furthermore, we have provided evidence that general affective responses of the audience to the elements in the scene influence suspense when this elements are showed in the scene [11].

In this paper we describe the process and the results of adding a cognitive model of the influence of decorative elements versus other scene affective elements –threat features and resources– in the perception of suspense to an automatic storytelling system. To achieve this, the architecture has been built as integrated part of the generator Stella [19].

In order to test the results, we have developed a stripped-down version of Stella. This version is able to produce stories that include several elements and actions that trigger specific cognitive responses in the audience. With these generated stories, an experiment with $N = 37$ human subjects was conducted. The subjects rated their perceived suspense from an small core of suspenseful stories based on a classical threat-discover and escape sequence.

From the set of stories generated by the story generator, four representative scenes were selected by hand. The selection was carried out according to the affective values of a set of features: threat profession, threat physical aspect, threat weapon and decorative elements. The results suggest that including explicit, validated cognitive model of the reaction to suspenseful improves generation results and, more importantly, provides a new set of parameters for story generation.

The paper is structured as follows: Section 2 describes the related previous work on suspense through emotions evoked by story concepts, and how suspense is currently treated by automatic storytellings. Section 3 explains the inclusion of the architecture model in Stella. Section 4 describes the experiment, whose results are detailed in Section 5. Finally, Section 6 and Section 7, respectively, discuss and summarize these results.

2 Related Work

This section tries to provide a general description of the state of the art in the related fields: emotions and their relation with decorative elements (sec. 2.1), automatic suspense generation (sec. 2.2) and the story generation system that was adapted for this research (sec. 2.3).

2.1 Emotions associated to concepts and impact of decorative elements

The impact in suspense of the elements in the scene seems to be strongly related to their semantics and the emotional characteristics that, on average, they

trigger on the audience [13, p. 28]. This hypothesis was demonstrated in [11]. This evidence makes it possible for a computational system to include this information for predicting suspenseful reactions to decorative elements. Affective Norms for English Words (ANEW) was used to make the predictions, being this an extensive list that contains a number of emotional aspects of the included terms.

The American ANEW was created as the result of an experiment of Bradley and Lang (1999). Participants were asked for the emotional affection of 1034 words. To evaluate the set of words, participants selected each dimension painting in a 9-point rating scale represented by the Self-Assessment Manikin (SAM) [2]. Devised by Lang in 1980, the SAM model is a non-verbal, pictorial assessment technique that directly measures the emotion conceptualizing it in three dimensions: valence (or pleasure, ranging from *pleasant* to *unpleasant*), arousal (ranging from *calm* to *excited*) and dominance (or control, ranging from *in control* to *out of control*) associated with a person's affective reaction to a wide variety of stimuli [16, p. 39] [2, p. 49]. ANEW experiment has been replicated for several languages as French, Finnish, Dutch, Portuguese or Italian [20,14,22,31,21].

Emotional valence describes the extent to which something cause a positive or a negative emotion [8, p. 79]. In terms of the story, an element has a negative valence when it pushes towards a negative outcome. It has been extensively investigated the paradox in that texts with negative valence are perceived as more amusing than texts with neutral or positive valence.

The second dimension is the arousal [1], that refers the intensity of the emotion [8, p. 79]. This dimension seems to have a similar effect on the audience that the pattern found in negative valence. So, the higher the discomfort during the tension phase, the higher the pleasure in the moment of resolution [17, p. 82]. Novelists and narratologists agree with that the duration of this intensity has an important role in this tension. "Suspense" comes from the world "suspend". Its etymology suggest that the more suspense is wanted, the longer suspend the scene is needed [24, p. 106]. Presenting the outcome a little later than expected [9, p. 325] is a key that relates suspense and timing.

Finally, the third dimension, called dominance, control or power, reflects the degree of control an individual feels over a specific stimulus and extends from out of control to in control [21, p. 888].

In Delatorre et al. (2017), results show that these affective dimensions make it possible to suspense in a scene through the modification of the decorative elements present in it. Thus, decorative elements, even when not influencing the narrative plot, impact the perception of suspense. This impact is related to the emotional features of these objects [11, p. 297]. Concretely, the experiments support that valence and dominance are moderately correlated with suspense ($\rho_{val} = -0.579$, $p < 0.01$; $\rho_{dom} = -0.423$, $p < 0.01$) for textual stories.

2.2 Automatic Suspense Generation Systems

This section has summarized the treatment of suspense in the main computational narrative systems. Given the scope of the proposed model of suspense, the review focuses on generative systems.

MEXICA [27] is a program that writes short stories about the Mexicas, the old inhabitants of what today is Mexico City [27, p. 2]. These stories are represented as clusters of emotional links and tensions between characters, progressing during development, and whose operators, intensity and predefined texts are customizable. In MEXICA, it is assumed that a story is interesting when it includes degradation-improvement processes (i.e., conflict and resolution) [27, p. 4]. Throughout the history, emotional links among the characters vary as a result of their interactions; so, *princess healed jaguar knight* produces the effect of increasing a positive emotion (gratitude) from the knight to the princess.

MEXICA is an exception in the use of positive emotions to implement the narrative tension. The system works with two predefined types of emotion: brotherly love and amorous love, both ranging from -3 (negative emotion) to 3 (positive emotion). Additionally, ten types of tension are defined (actor dead, love competition, health normal...), which are generated based on the type and emotional value of each character. The stories search degradation-improvement curves through actions that transform the extent of the tensions.

MINSTREL [35], meanwhile, is a complex program that writes short stories about Arthurian legends, implemented on a case-based problem-solver where past cases are stored in an episodic memory [28, p. 4]. MINSTREL recognizes narrative tension plots and tries to increase the suspense by adding more emotionally charged scenes, by storing a simple ranking which tells when such inclusion is reasonable; for example, when the action is preserving a life. It uses two strategies for generating suspense: via character emotion and via character escape. In the first one, text includes a sentence that describes the fear of the character about the immediate threat. The second one adds another sentence that reports a failed character's attempt [35, p. 123–126].

Another initiative is Suspenser [7], that creates stories with the objective of increasing the reader's suspense. It provides an intermediate layer between the fabula generation and the discourse generation, which selects the steps of the plot according to their *value of importance* for the final goal. For this and based on the Gerrig & Bernardo's assumption[3], Suspenser uses a set of heuristics grounded in the number of paths available for the character to reach its goal, considering optimal the probability of protagonists' success as 1/100 [6, p. 59].

Also based in Gerrig & Bernardo's work, Dramatis proposes an implementation of a system to evaluate suspense in stories that utilizes a memory model and a goal selection process [26, p. 5], assuming that the reader, when faced with a narrative, evaluates the set of possible future states in order to find the best option for the protagonist. With a similar target, Dramatis generates escape plans attempting to "break" the causal links that would reach non-desired

[3] "Readers feel suspense when led to believe that the quantity or quality of path through the hero's problem space has become diminished". [15]

goals (typically, the character death) and the reader could predict more easily. To do this, the memory model assigns more relevance to the elements recently narrated than to those mentioned at the beginning of the story.

Finally, we review IDtension [32], a drama project which comes up in order to demonstrate the possibility of combining narrative and interactivity. Unlike approaches based in character's chances or the course of the actions, it conceives the stories based on narrative properties (conflict or suspense).

Suspense is treated by IDtension as a reaction to the *obstacles* (conflicts), and is correlated to the risk of facing every expected obstacle (high or low risk, without intermediate values). The narrative effects of the tension are calculated by six criteria: ethical consistency, motivational consistency, relevance, cognitive load (influence in the story), characterization and conflict. Also, the condition is managed by a series of actions as accepting, refusing, congratulate, etc., available for use on / among the characters.

With respect to our goals, the review of the above systems has exposed the aforementioned comparative limitation: we can observe that none of them takes into account a general cognitive theories as explicit part of the model, neither physical aspect, resources nor environmental issues.

2.3 The Story Generation System Stella

Stella is a story generation system based on a hybrid model in which exhaustive, non-deterministic simulation is controlled by a narrative layer [19].

Stella models stories as time-ordered sequences of states. Each state contains a detailed representation of each of the entities that populate it: physical information, emotions, intentions, knowledge about the world, and others. The simulation is carried out non-deterministically. On each generation step, the current state $s^{current}$ is expanded and all its potential next states $\{s_1^{next}, s_2^{next}, \ldots, s_n^{next}\}$ are generated. This means that, for each non-deterministic option for each next value of each attribute for each entity, a new path is created. This produces a vast generative space of stories.

The narrative layer receives this huge space and uses narrative information to identify the best alternatives. In order to do this, Stella uses several techniques. The system allows to set conditions for the output stories. For instance, it is possible to filter out those stories which do not include a murder, or search for stories in which there is a love scene. Stella can also control the generative process by setting narrative aspects like length, amount of interaction between characters and other general aspects of narrative.

Additionally, Stella can drive the generation by controlling the evolution of user-defined dimensions, which are represented as curves [18]. Each of these dimensions are given to the system in the form of objective curves, and Stella controls the generation so that the resulting curve for each of the dimensions matches the objective with a given error margin.

3 Including Affective Terms and their Impact in the Story Generation System

As previously introduced, the main objective of this research is to include a partial cognitive model of suspense in a story generation system. This section explains how the Stella story generation system has been adapted to generate short suspenseful stories by including elements and actions that convey affective response related to suspense.

3.1 Stellite: A stripped-down version of Stella

As introduced in section 2.3, the story generation system Stella addresses many characteristics of storytelling. This implies a complex execution model and an elaborated data model. These two aspects make it difficult to isolate and control a specific subset of the generation parameters, which is crucial if the output is to be tested against human evaluators.

In order to run a more controlled environment, a new, stripped-down version of Stella has been created. In this version, only the core generation engine and the knowledge model have been kept. The curves generation and matching engine have been removed, and the generation constraints and objectives have been made simpler.

The simulation engine has also been simplified. The physics subsystem in Stella is based on Newtonian equations, and Stellite only performs simple computations for movement in a discrete world. Instead of having a full physics engine, in Stellite characters move only to interesting locations (i.e. hard-coded landmarks). Besides, manipulation in the full version of Stella lets characters grab and store things, and do multiple things at the same time if they are possible. In Stellite, characters can only grab or release one item. Instead of developed behavioural models, they have simple traits (like "being a murderer") and straightforward objectives ("escaping" or "killing") which are now specific symbols with semantics intended to guide the generation.

With this added information, Stellite performs story generation. This process creates stories as ordered sequences of states. The generation in carried out non-deterministically, in such a way that for each current state $s^{current}$ all possible next states $\{s_1^{next}, s_2^{next}, \ldots, s_n^{next}\}$ are explored, in the same way that original Stella does. The expansion of the nodes uses the non-deterministic simulation engine to produce all candidate new states. On each new state, all characters are updated: they receive the world information, reason, update their objectives, and act accordingly.

Compared to Stella, the set of output stories is noticeable smaller. This reduces the chances to find a highly original story, and makes it impossible to generate at a very fine detail, but the generation is faster and the output stories are all coherent.

3.2 Including Affective Terms in Stellite

As introduced in Section 2.1, we have previously found a correlation between the affective aspects of terms and the suspenseful feeling they trigger [12]. That study was carried out using ANEW, which is a validated mapping from terms to these affective values. The ANEW model relates terms with their valence, arousal and dominance [3]. Since the evaluation of the system has been carried out with native Spanish speakers –see Section 4.1–, the Spanish version of ANEW [29] was used. Although there are some differences between the Spanish and the English version of ANEW, the terms in both studies are the same.

ANEW includes 1034 words. For implementing Stellite and testing the effect of terms, we have selected a subset justified by the category of the terms: professions (20) –from *terrorist* (valence of 1.51) to *writer* (6.47)–, physical adjectives (30) –from *sick* (valence of 1.61) to *elegant* (7.23)–, handheld weapons (9) –from *pistol* (valence of 1.83) to *bottle* (5.10)– and decorative elements (25) —from *corpse* (valence of 1.41) to *flower* (7.34); this one coming from the set already collected in [11]–.

In order to model professions in Stellite, we have selected names that represent tasks which may be performed by the corresponding characters. Physical adjectives are considered as visually perceptible features. The set of chosen handheld weapons includes resources that can be used by a character to attack another character. Likewise, decorative elements are defined as those entities present in a scene or story which do not play a role in the main plot, and could therefore be interchanged by others or removed without any relevant change in the narrative structure of a story. Accordingly, no entity in Stellite interacts with them.

4 Experiment

This section describes the experiment and the methodology that was applied to extract the information about the existing differences between audiences of interactive and non-interactive stories. Section 5 describes the results.

4.1 Participants

The experiment took place in the Computer Science Faculty of University of Cadiz (Spain). The experiment was publicly announced to all students, and those wanting to take part in it voluntary enrolled. The objective and the process were explained to them and they all accepted the conditions. Thirty seven undergraduate students ($N = 37$), 32 males (86.49%) and 5 females (13.51%) participated. All the participants were students of Computer Science degree. Their ages ranged from 21 to 33 years ($mean = 23.92$, $stdev = 3.15$). All participants were Spanish native speakers. There was no compensation for participating in the experiment.

The low proportion of volunteer female subjects made the experiment unbalanced in terms of participants' gender. This issue is analyzed in Section 5 and discussed in Section 6.

4.2 Story

Stellite was parameterized to generate stories about one character trying to escape and one "murderer" trying to kill him. The world was formed by a simple map including a bedroom, a corridor and a room with an exit door. The generator was configured to generate stories with 6 events.

By combining character decisions, affective terms and different endings, Stellite generated 10275 stories. Several stories did not include any suspenseful event, and some others did finish without a clear outcome. In order for the evaluation to be doable and useful, we decided to choose one instance in which there was some suspense. It is obvious that this selection affects the overall results, and this is discussed in Section 6.

The text was generated from the structured representation. The content was filtered and rendered in Spanish with simple text templates. The chosen story is shown next[4]:

> *John walked down the corridor, trying not to make noises. Suddenly, he heard something behind. When he turned back, he saw the terrorist. The terrorist was sick and carried a pistol. John tried to escape through the door. On the floor, there was a flower. John opened the door. The terrorist chased John.*

In order to provide variation, we selected four versions of the previous short story in Stellite. For selecting these versions, we combined decorative elements and profession-aspect-weapon (threat), and crossed maximum and minimum values of valence. The next list details the selection:

1. The threat is a sick terrorist carrying a pistol. There is a corpse on the floor: minimum valence for scene, minimum valence for decoration.
2. The threat (the murderer) is a sick terrorist carrying a pistol. There is a flower on the floor: minimum valence for scene, maximum valence for decoration.
3. The threat is an elegant writer carrying a hammer. There is a flower on the floor: maximum valence for scene, minimum valence for decoration.
4. The threat is an elegant writer carrying a hammer. There is a corpse on the floor: maximum valence for scene, maximum valence for decoration.

Table 1 shows ANEW valences for these versions.

4.3 Method

The experiment was run as a *paper-and-pencil* session in one single classroom. Participants were randomly placed, keeping an empty table between each pair.

[4] In the original Spanish: "*Juan caminaba por el pasillo tratando de no hacer ruido. De repente, oyó algo a su espalda. Cuando se volvió, vio al terrorista. El terrorista estaba enfermo y llevaba una pistola. Juan intentó escapar por una puerta. Junto a la puerta, en el suelo, había una flor. Juan abrió la puerta. El terrorista persiguió a Juan.*"

Table 1. Valences of each story version

version	elegant	writer	hammer	sick	terrorist	pistol	flower	corpse	total valence
1	-	-	-	1.61	1.51	1.83	-	1.41	6.36
2	-	-	-	1.61	1.51	1.83	7.34	-	12.29
3	7.23	6.47	4.38	-	-	-	-	1.41	19.49
4	7.23	6.47	4.38	-	-	-	7.34	-	25.42

A single demographic survey was filled by each participant. The experiment was explained. Subjects were warned that the scene that they were about to read could have been written either by a human or by a computer. After this point, the content was presented, handing the evaluator over a sheet with one version of the story. Versions were equally distributed.

The test invited the participant to imagine the scene and rate the level of suspense using a 1 to 9 range (according to the SAM model, which is used by ANEW). Additionally, subjects were asked to answer the question: "Do you think this story has been written by a human?".

5 Results

This section details the results obtained from the experiments described in previous Section 4, comparing the suspense of the different versions of the scene. For all measures, the criteria for statistical significance was set at $\alpha = 0.05$.

Globally, results show a moderate-high downhill correlation between reported suspense and valence of the scene ($\rho = -0.749$, $p < 0.000$). There are three different significant groups identified. There were no significant differences between version 2 and 3. Differences by story elements are showed[5] in Fig. 1.

When comparing the effect of the threat with decorative elements, results show that both influence reported suspense. However, the effect of decoration ($Z = 2.394$, $p < 0.02$) is weaker than the effect of the threat features ($Z = 4.945$, $p < 0.000$). Illustratively, Table 2 shows the mean and standard deviation for each dimension, where a sightly higher standard deviation is observed when the decorative element is a flower.

This effect of decorative elements may be contrasted by analyzing the responses on whether the question had been written by a human or not. Actually, while significant influence of the threat features was not found ($Z = 0.176$, $p = 0.860$), the decorative elements used in the story clearly influence it ($Z = 2.129$, $p < 0.05$): 72.22% of the subjects reported that the creator was a machine when the decorator was the flower; in contrast, only 36.84% of the subject reported that opinion when the decorator was the corpse. This may be influenced by the coherence of the element in the story, as will be discussed in Section 6. Tak-

[5] *Terrorist* includes *sick* and *pistol*. Likewise, *writer* includes *elegant* and *hammer*. It has been reduced due to the limited space.

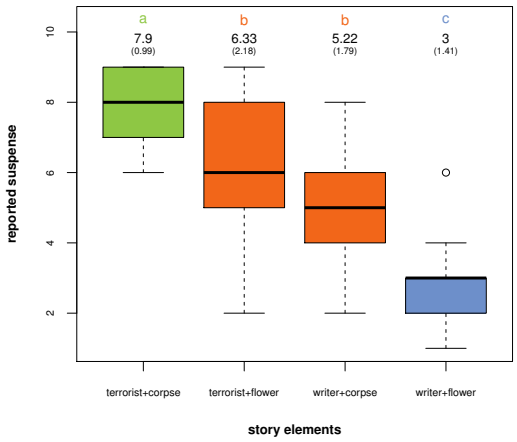

Fig. 1. Groups of reported suspense by story elements.

Table 2. Threat features and decoration effect comparison

dimension	mean	std
threat feature		
sick terrorist carrying a pistol	7.15	1.80
elegant writer carrying a hammer	4.11	1.93
decorative element		
corpse	6.63	1.95
flower	4.67	2.47

ing into account all the answers, 54.05% of subjects thought that the story was generated by a machine, 35.13% when the flower was in scene.

Finally, no significant influence of gender in general reported suspense has been found ($Z = -0.135$, $p = 0.893$). The same result has been obtained when analyzing the impact of the affective elements by itself.

6 Discussion

Despite of the promising results, some issues need to be addressed.

In contrast to the idea that audience gender influences suspense [23,4,25], no significant differences were reported in the analysis. It may be due to the small number of female subjects (5), so we cannot obviate the gender impact. Although it does not invalidate the general conclusions of this study, this effect must be reviewed more closely in further research.

The seed story used for the experiment, while generated by the story generation system, was chosen by hand. Automatically identifying which story would provide the best outcome (or a list of valid set of answers in terms of suspense) would require a computational system at least as complex as the story generation system. This would have made it impossible to run the proposed texts. However, further research considers addressing this issue and provide a model of identifying promising stories.

It is perceptible that most subjects consider the story as created by a computer when the scene includes the flower. We understand that this effect is due to a sort of *learned-inconsistence*: while corpse is a self-explained typical decorator in a suspenseful chase scene, focusing a flower may be decontextualized.

In fact, there are combinations of elements that have not been tested and could evoke a similar perception –for example, what would happen if the terrorist is carrying the hammer–. However, checking the consistence of affective elements working together, despite it is an essential mechanism of an automatic generator of suspense, it is outside the scope of this paper.

7 Conclusions

The paper has described an original system that produces short stories by including suspenseful decorative elements in them. The decorative elements and their expected impact in humans were previously validated, and we have used this information to compare the performance of the story generation system against the expected value with human evaluators.

Additionally, other evaluable affective elements, which were not tested previously –threat features and resources–, have been included in Stellite. The influence found in the reported suspense by the experimental subjects supports the evidence that the effect of the emotional affectivity is not limited to the decorative elements. It leads to analyze the particularities of this effect and quantify it, in comparison with the decorative elements, as immediate future work. This will be addressed as part of the future work.

Acknowledgements

This work has been funded by the Andalusian Government under the University of Cadiz programme for Researching and Innovation in Education 2015/2016 (SOL-201500054211-TRA), and by the IDiLyCo project (TIN2015-66655-R) funded by the Spanish Ministry of Economy, Industry and Competitiveness.

The research presented in this paper has been carried out as part of the first author's work in the PhD Programme in Computer Science of Universidad Complutense de Madrid (RD 99/2011) and is part of his PhD dissertation.

References

1. Daniel E Berlyne. Conflict, arousal, and curiosity. *McGraw-Hill Series in Psychology*, XII, 1960.

2. Margaret M Bradley and Peter J Lang. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59, 1994.
3. Margaret M Bradley and Peter J Lang. Affective norms for english words (ANEW): Instruction manual and affective ratings. Technical report, Technical Report C-1, The Center for Research in Psychophysiology, University of Florida, 1999.
4. Leslie A Burton, Laura Rabin, Susan Bernstein Vardy, Jonathan Frohlich, Gwinne Wyatt, Diana Dimitri, Shimon Constante, and Elan Guterman. Gender differences in implicit and explicit memory for affective passages. *Brain and Cognition*, 54(3):218–224, 2004.
5. Andrew Caplin and John Leahy. Psychological expected utility theory and anticipatory feelings. *Quarterly Journal of economics*, pages 55–79, 2001.
6. Yun Gyung Cheong. *A computational model of narrative generation of suspense*. PhD thesis, North Carolina State University, 2007.
7. Yun-Gyung Cheong and R Michael Young. A computational model of narrative generation for suspense. In *Association for the Advancement of Artificial Intelligence (AAAI Journal)*, pages 1906–1907, 2006.
8. Francesca MM Citron, Marcus A Gray, Hugo D Critchley, Brendan S Weekes, and Evelyn C Ferstl. Emotional valence and arousal affect reading in an interactive way: neuroimaging evidence for an approach-withdrawal framework. *Neuropsychologia*, 56:79–89, 2014.
9. Minet de Wied, Ed SH Tan, and Nico Henry Frijda. Duration experience under conditions of suspense in films. *NATO ASI series. Time, action and cognition: Towards bridging the gap*, pages 325–336, 1992.
10. Pablo Delatorre, Barbara Arfè, Pablo Gervás, and Manuel Palomo-Duarte. A component-based architecture for suspense modelling. In *Proceedings of AISB 2016's Third International Symposium on Computational Creativity (CC 2016)*, pages 32–39, Sheffield (UK), 2016.
11. Pablo Delatorre, Carlos León, Pablo Gervás, and Manuel Palomo-Duarte. A computational model of the cognitive impact of decorative elements on the perception of suspense. *Connection Science*, 29(4):295–331, 2017.
12. Pablo Delatorre, Carlos León, Alberto Salguero, Cristina Mateo-Gil, and Pablo Gervás. Impact of interactivity on information management for suspense in storytelling. In Joanna Bryson, Marina De Vos, and Julian Padget, editors, *Proceedings of AISB 2017's Fourth International Symposium on Computational Creativity (CC 2017)*, pages 110–116, Bath (UK), 2017. Society with AI.
13. Pablo Delatorre, Manuel Palomo-Duarte, and Pablo Gervás. Formalising suspense from immersive environments. In David Camacho, Marco Antonio Gómez-Martín, and Pedro Antonio González-Calero, editors, *Proceedings of 3rd Congreso de la Sociedad Española para las Ciencias del Videojuego (CoSECiVi 2016)*, volume 1682 of *CEUR Workshop Proceedings*, pages 25–36, Barcelona, Spain, 2016.
14. Tiina M Eilola and Jelena Havelka. Affective norms for 210 British English and Finnish nouns. *Behavior Research Methods*, 42(1):134–140, 2010.
15. Richard J Gerrig and Allan BI Bernardo. Readers as problem-solvers in the experience of suspense. *Poetics*, 22(6):459–472, 1994.
16. Peter J Lang, Margaret M Bradley, and Bruce N Cuthbert. International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, pages 39–58, 1997.
17. Moritz Lehne. *Emotional Experiences of Tension and Suspense*. PhD thesis, Freie Universität Berlin, Germany, 2014.

18. Carlos León and Pablo Gervás. Prototyping the Use of Plot Curves to Guide Story Generation. In *Third Workshop on Computational Models of Narrative, 2012 Language Resources and Evaluation Conference (LREC'2012)*, Istambul, Turkey, 2012.
19. Carlos León and Pablo Gervás. Creativity in story generation from the ground up: Nondeterministic simulation driven by narrative. In *5th International Conference on Computational Creativity, ICCC*, pages 201–210, 2014.
20. Catherine Monnier and Arielle Syssau. Affective norms for French words (FAN). *Behavior research methods*, 46(4):1128–1137, 2014.
21. Maria Montefinese, Ettore Ambrosini, Beth Fairfield, and Nicola Mammarella. The adaptation of the Affective Norms for English Words (ANEW) for Italian. *Behavior research methods*, 46(3):887–903, 2014.
22. Agnes Moors, Jan De Houwer, Dirk Hermans, Sabine Wanmaker, Kevin van Schie, Anne-Laura Van Harmelen, Maarten De Schryver, Jeffrey De Winne, and Marc Brysbaert. Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior research methods*, 45(1):169–177, 2013.
23. Justin M Nolan and Gery W Ryan. Fear and loathing at the Cineplex: gender differences in descriptions and perceptions of slasher films. *Sex Roles*, 42(1-2):39–56, 2000.
24. Raymond Obstfeld. *Novelist's Essential Guide to Crafting Scenes*. Novelists Essentials. F+W Media, 2000.
25. Mary Beth Oliver and Meghan Sanders. The appeal of horror and suspense. *The horror film*, pages 242–260, 2004.
26. Brian O'Neill. *A computational model of suspense for the augmentation of intelligent story generation*. PhD thesis, Georgia Institute of Technology, 2013.
27. Rafael Pérez y Pérez. Employing emotions to drive plot generation in a computer-based storyteller. *Cognitive Systems Research*, 8(2):89–109, 2007.
28. Rafael Pérez y Pérez and Mike Sharples. Three computer-based models of storytelling: BRUTUS, MINSTREL and MEXICA. *Knowledge-based systems*, 17(1):15–29, 2004.
29. Jaime Redondo, Isabel Fraga, Isabel Padrón, and Montserrat Comesaña. The Spanish adaptation of ANEW (affective norms for English words). *Behavior research methods*, 39(3):600–605, 2007.
30. Gregory Schraw, Terri Flowerday, and Stephen Lehman. Increasing situational interest in the classroom. *Educational Psychology Review*, 13(3):211–224, 2001.
31. Ana Paula Soares, Montserrat Comesaña, Ana P Pinheiro, Alberto Simões, and Carla Sofia Frade. The adaptation of the affective norms for English words (ANEW) for European Portuguese. *Behavior research methods*, 44(1):256–269, 2012.
32. Nicolas Szilas. IDtension: a narrative engine for interactive drama. In *Proceedings of the Technologies for Interactive Digital Storytelling and Entertainment (TIDSE) Conference*, volume 3, pages 187–203, 2003.
33. Angela Tinwell, Mark Grimshaw, and Andrew Williams. Uncanny behaviour in survival horror games. *Journal of Gaming & Virtual Worlds*, 2(1):3–25, 2010.
34. François Truffaut and Helen Scott. *El cine según Hitchcock*. Alianza Editorial (Madrid), 1974. Reedition of Le Cinéma selon Hitchcock, Robert Laffont Éditions (Paris, 1966).
35. Scott R. Turner. *The Creative Process: A Computer Model of Storytelling and Creativity*. Taylor & Francis, 2014.
36. Peter Vorderer and Silvia Knobloch. Conflict and suspense in drama. *Media entertainment: The psychology of its appeal*, pages 59–72, 2000.