Interactions in Multisensory Experiences: Toward a New Timeline Metaphor

Céline Jost*

Paris 8 University, CHArt laboratory, celine.jost@univ-paris8.fr

Brigitte Le Pévédic

Southern Brittany University, LAB-STICC laboratory, brigitte.le-pevedic@univ-ubs.fr

Justin Debloos

Paris 8 University, CHArt laboratory, justin.debloos@gmail.com

Gérard Uzan

Paris 8 University, CHArt laboratory, gerard.uzan@univ-paris8.fr

The PRIM project has the objective to allow everyone to produce multisensory interactions without programming and without the help of programmers. Taking inspiration from the simplicity of use of media editing software, our idea is to investigate how to enrich them to integrate interactions by modifying the editing timeline. After arguing that time and timeline are the major challenge to face before building a new software, this paper presents a new timeline metaphor, easy to understand, and that would allow to produce interactive mulsemedia and multisensory exercises once completed. Promising results show that people easily understand and accept this new timeline, which encourages us to continue in this way.

CCS CONCEPTS • Human-centered computing~Human computer interaction (HCI)~Interactive systems and tools~User interface programming •Software and its engineering~Software notations and tools~Context specific languages~Visual languages

Additional Keywords and Phrases: timeline metaphor, interactive mulsemedia, multisensory exercises, media editing software, music scores

ACM Reference Format:

Céline Jost, Brigitte Le Pévédic, Justin Debloos and Gérard Uzan. 2022. Interactions in Multisensory Experiences: Toward a New Timeline Metaphor. In SensoryX'22: Workshop on Multisensory Experiences, together with IMX 2022: ACM International Conference on Interactive Media Experiences. June 22-24, 2022. Aveiro, Portugal.

---

1 INTRODUCTION

The PRIM project (Playing and Recording with Interactivity and Multisensoriality) aims at gathering a multidisciplinary community to conceive an original software offering users to rapidly and simply create interaction scenarios between human and different digital devices [1].

For instance, in the health community, therapists need tools to create exercises where a patient interacts with the technology [2], [3]. PRIM is interested in cognitive or cerebral stimulation through multisensory exercises because using several senses to stimulate a person increases cognitive performance [4], [5]. However, it requires a lot of time either to learn to program visual languages such as Scratch [6], Choregraph [7], and Blockly [8] or to find funds to pay someone to program exercises. Therapists wish to be able to rapidly create exercises that use functionalities already existing in digital devices. In the context of cognitive stimulation, they could, for instance, create a quiz where a robot asks questions associated with sensory clues (wind, light, vibrations, scent, ...). The patient would answer with different actions: pressing buttons, making a gesture (captured and recognized by a smart camera), validating answer cards with an RFID reader, ...). There are numerous exercises that can be created by combining existing functionalities from digital devices.

In the multimedia community, researchers have developed new authoring tools to create mulsemedia (i.e multiple sensorial media [9]) where it is possible to synchronize sensory effects (wind, light, vibrations, ...), produced by digital devices, with audiovisual content. These tools, especially editing software, such as SEVino [10]–[12] and PlaySEM [13], seem very easy to manipulate and partially meet our requirements. Current research even tends to converge toward our issue with the willingness to create interactive media where humans' actions are taken into account (for example STEVE [14], [15]). The objective is to allow the viewer to act on the movie, for example, by typically reading, pausing, or stopping the film, but also by creating new media where humans' actions would impact the story such as in "the books where you are the hero".

Whether to create multisensory exercises or to create interactive mulsemedia, current systems are too complex [16], requiring a big part of programming [5]. For instance, if the video editor wants to create an alternative end, she/he needs to define an algorithmic condition. And if she/he wants to make a film where the same day is repeated again and again (until the hero makes the expected action), she/he needs to define a loop.

In the PRIM project, our approach is to consider that visual programming languages are too complex [17], [18] and that it could be interesting to investigate **how we could enrich mulsemedia editing software to allow the creation of interactive media**, thus allowing the creation of multisensory exercises. The major challenge is to maintain the same simplicity and usability as the current editing software.

Before investigating this question, we have to address a major issue, about the *timeline*, detailed in this paper. To this aim, section 2 highlights both opposite visions from visual programming languages and mulsemedia editing software and introduces the major challenge which has to be addressed before attempting to enrich current editing software. Section 3 presents a prototype developed to answer the major challenge. Then section 4 describes the methods and results of the evaluation we conducted with the prototype. Last, section 5 discusses results, offers perspectives, and concludes this paper.

2 MAJOR ISSUE: THE TIME

Visual programming languages and media editing software have two different visions. The first category is based on algorithmic while the second one is based on audiovisual content. Algorithmic is a discipline that requires time and works to learn and master. It is based on Boolean logic, thus on electronics, which is not natural for humans. Media editing is easier to manipulate because users have to place elements, such as subtitles or sensory effects, on a timeline that shows the progress of the audiovisual content. It seems to be rather based on human logic. We think that media editing software may be easier to use because it would be based on the humans' mental model (to be verified during the project). Indeed, the user has to identify what she/he wants to add, directly written or drawn in her/his language, and has to select and place it on a timeline that expresses the human's time. Everything here is natural for humans. However, in visual programming languages, humans have to translate what they want to do into the machine language and permanently have to mentally picture what actions will occur at what time.

In the visual programming languages, actions occur through time, of course, but not on the time, like in the editing software, even if it is also possible to program actions at a specific time. However, the reverse is not true since it is impossible to place event actions on a timeline from editing software. There is no tool which makes both temporalities co-exist. When both temporalities exist, such as in Choregraphe [7], they belong to different software areas and use different graphic designs. And, in STEVE [15] which seems to be the latest software, it is possible to create separated timelines that are not synchronized with the audiovisual content time. However, that prevents acting on the media. That emphasizes our idea to make both temporalities co-exist in the same area in order to create interactive media or multisensory exercises.

Thus, we face a major challenge that is: **how to integrate events in a timeline system which is also compatible with real temporality?** This requires that the timeline manipulates both real temporality and event temporality, which raises a new question: **how does this timeline work, and is it easily understood and manipulated by users whose habits will necessarily change?** In this paper, we investigate the first step: the possibility of a timeline to manipulate events.

3 SCENAPROD: A FIRST PROTOTYPE

Our first proposal aims at investigating the acceptability of a timeline based on event temporality. In order to allow users to manipulate our timeline during a creation process, we have created a prototype, inspired by video editing software. This prototype has been developed for the sole purpose of contextualizing our timeline, thus it concerns only users who will create some exercises or interactive medias. The patients or viewers are not expected to see the timeline. Figure 1 shows an example of a creation while being played (useful to simulate the creation)[2]. Note that only the acceptability of the timeline is evaluated here, and that the interface itself is not the main priority.

In this prototype, producers can play a sound, display a text and/or a picture, and ask users to type a key on the keyboard which allows them to manipulate events very simply. Each component can be configured in the grey area. The key point in this paper is the editing area (in green). We propose a dotted line to represent discontinuous time. This line continues to the next line if necessary. There is a contextual menu (the black plus sign) that appears between each component placed to allow making changes on the timeline. In Figure 1, the main timeline is duplicated three times. In this case, each timeline has its own time. Thus, if the producer clicks on the plus sign and on "timeline duplication", a new timeline appears and offers the user the possibility to experience parallelism, very easily, while it is a complex algorithmic concept.

Our idea is that the future software will hide all algorithmic elements in the timeline.

In the evaluation of our timeline, we have purposely decided to cause cognitive dissonance, meaning that components are vertically aligned but have different time values (they can last longer or shorter). For example,

[2] See an example of the execution in: https://youtu.be/VWv_qcOFARA

the component on the first line lasts 2 minutes while the first components of lines 3 and 4 are run instantaneously. Each component being currently run is surrounded by a green border, and each component that has already been run is transparent. Thus, in this program, there is a dissonance between vertical synchronization (X axis) and time synchronization. Components currently running are vertically desynchronized, which is here the most important difference with a time-based timeline in which there is a vertical time synchronization. **Our objective is to evaluate if this kind of time execution can be understood and accepted by people.**

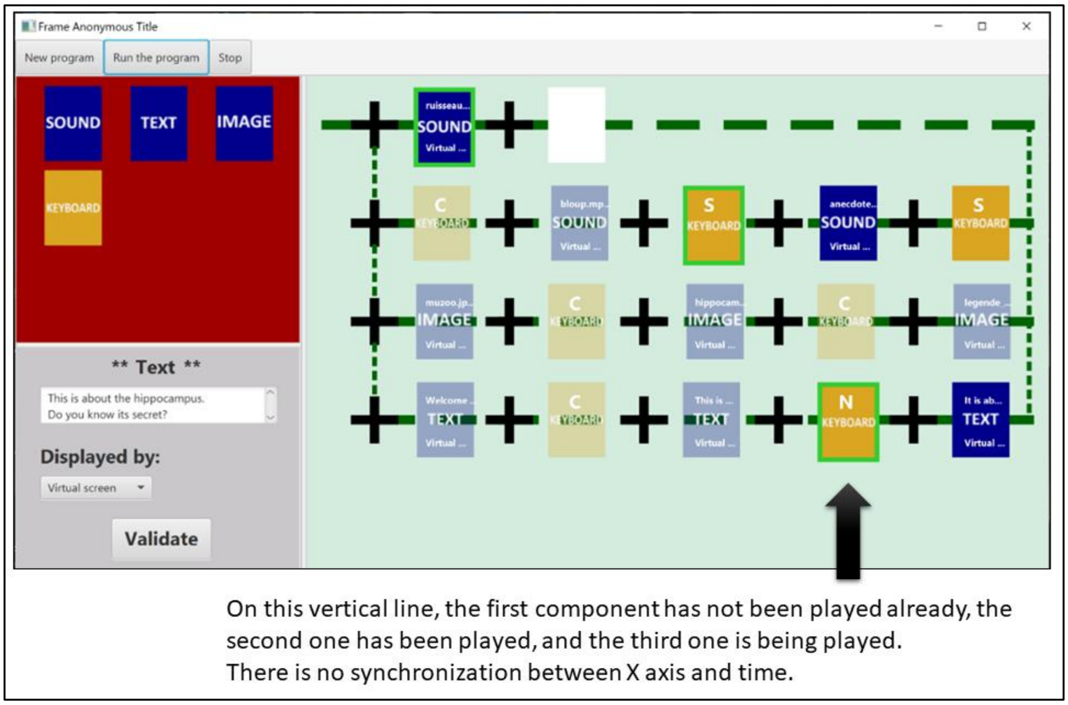

Figure 1: example of one production written with a relative timeline

4 EVALUATION

4.1 Method

The objective of this study was to evaluate how participants got used to our new timeline. To this end, we conducted a remote evaluation where participants had to get connected to the Zoom system and remotely take control of a MacOs computer.

The evaluation was divided into three steps: (1) Participants had to sign a consent form and were informed that there was no recording, that the data would be anonymized, and that they were able to stop anytime. (2) Participants had to follow a tutorial to create, one after the other, three productions with increased difficulty. The tasks had been chosen to make participants experience the execution of the production and to expose them to simultaneous different stimuli that were not visually represented on the same X axis, as explained above. The final objective was to ensure that each participant had seen and experienced what is shown in Figure 1. (3) Participants had to fill out a questionnaire. If we want to be sure that participants evaluate the timeline, we have to measure the prototype usability and to check the user experience. If the user experience is good, that means the prototype do not produce a bad experience thus do not produce bad artefact in the timeline evaluation. That is why the first part of the questionnaire was a System Usability Scale (SUS) [19]. We did not aim at evaluating the prototype, but at validating the absence of a bad user experience. The second part was composed of 10 questions using the same scale as the SUS (5-point Likert scale) and asking specific questions about the software, the timeline, and participants' projection (see Table 1). The third part asked for information about participants and their opinion through open and closed questions.

A session lasted less than 30 minutes. The experimenter was present all the time, giving instructions or answering questions, or helping participants if necessary. The experimenter was also in charge of writing the total duration of the session, the number of asked questions, the number of times when participants were blocked, and any comments from participants, if existing.

We had previously conducted a preliminary evaluation with 5 participants to test the protocol and to ensure that the session lasted less than 30 minutes.

Table 1: Ten questions asked about software, timeline, and participants' projection

N°	Question
Q1	I experienced difficulties to do the asked task.
Q2	I think that the software looks like video editing software
Q3	I think it is required to have computing skills to use this software.
Q4	I think it is difficult to understand how to place components on the timeline.
Q5	I think that this software can be useful for my professional activity.
Q6	After clicking on "execute the program", I think that the progression in the program is visually easy to understand.
Q7	I think that time management is destabilizing.
Q8	I think that it is difficult to understand that each line has its own time.
Q9	I think that it is easy to make a timeline duplication.
Q10	I think that I can create some new programs without help.

4.2 Results

We enrolled 50 participants (31 women, 19 men; mean age: 34.5 years old; standard deviation: 15.4; range: 12 to 75 years old) with a mean of 18 minutes per session (sd: 5.1; range: 9 to 33). Only five participants have had a mental block that required the help of the experimenter (8 times, with a maximum of three for the same person), and 18 participants have asked a total of 47 questions by curiosity (with a maximum of 8 for the same person). Thirty of them had already used video editing software. Eight of them were computer specialists. We recruited participants who had a computer and an Internet connection, coming from various socio-professional categories and different regions.

Concerning the SUS, the mean score is 84 (sd: 8.1; median: 85; min: 62.5; max: 95) which means, according to the interpretation scale, that the prototype is highly acceptable, with a score between good (71.4) and excellent (85.5).

Regarding the ten specific questions asked in addition to the SUS, answers are shown in Figure 2. Questions with an “R” indicate that they have been reverse-coded to make results comparable. For example, for question 1, which is “I experienced difficulties to do the asked task.”, results are 90% strongly disagree and 10% disagree. The expected result is a negation, thus to be able to compare all results, we reversed all sentences where we expected a negative answer. It makes easier the interpretation. Expected results are in green. We can thus considerer that Q1 is “I don’t experience difficulties to do the asked task.” After reverse-coding, it appears that 90% strongly agree and 10% agree, thus meaning that, none of the participants experienced difficulties.

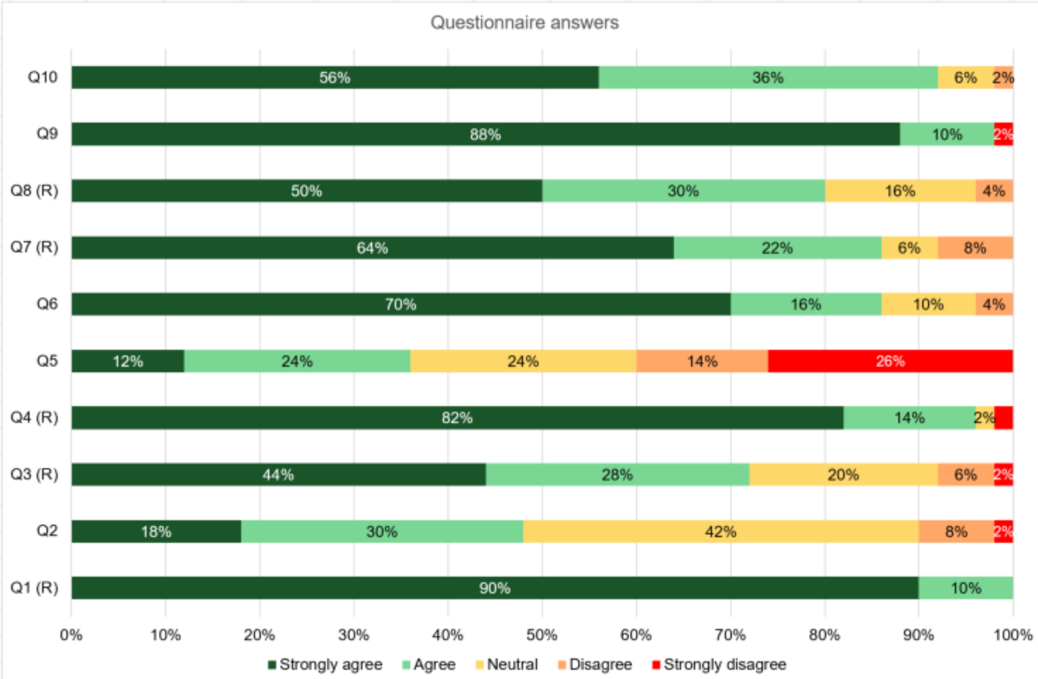

Figure 2: Answers of the ten specific questions

In addition to these ten questions, we asked six yes/no questions:

1. Did you understand that the timeline represents the program progression and not the real-time like in video playing? *Yes: 96%*
2. Do you think it is easy to understand and become familiar with this kind of time management? *Yes: 98%*
3. According to you, is there enough information displayed in the program to summarize it and to remind what it contains? *Yes: 86%*
4. Do you know a similar tool? *Yes: 18%*
5. Do you know another tool which uses reminds you of this tool? *Yes: 60%*
6. In the below screenshot, we can see that the progression is different on the four timelines. Did that disturb you? *Yes: 16%* (The screenshot refers to Figure 2).

Among the 18% of participants who know a similar tool, 67% of them have already used video editing software products. As similar tools, they named tutorial (once), Windows Movie Maker (twice), and Microsoft PowerPoint (three times).

Among the 60% of participants who think our software prototype reminds them of another tool, 73% of them have already used video editing software products. They named video editing (x11), Visual Programming Language (x9), Powerpoint (x6), Videogame (x3), music editing (x2), tutorial, Google suite, H5P, Watch application, communication application.

5 DISCUSSION, PERSPECTIVES AND CONCLUSION

In this paper we presented a first idea of a new timeline metaphor inspired by media editing software with the aim to enrich media editing software. Indeed, our objective is to give the possibility to people to produce some interactive creations without programming. The same software is expected to help therapists creating some multisensory exercises for patients and to help producers creating some interactive medias for viewers.

This new timeline is original because it aims at making users manipulating both the real temporality (when the timeline is based on the time) and the event temporality (when the timeline is based on events which can occur anywhen).

Before designing such a timeline, we conducted a preliminary study to evaluate whether the users can understand and accept an event-based timeline. This study, which was conducted with 50 participants, showed an absence of rejection of the timeline. More than 80% of participants understood its functioning and its time management, and 100% of them thought about being able to make a new production without help.

These results are promising for the PRIM project because they encourage us to keep on in this way. With a user-centered design approach we will investigate what is needed for users to be able to produce some creations with the maximum of possibilities without designing a new visual programming language. The final result has to remain as simple as possible and to look like a simple timeline. We will also investigate how to add hidden algorithmic elements in the timeline such as in music scores. Finally, during the project, we will add digital devices to the prototype to allows the creation of interactive media or multisensory exercises, for example lights, fans, robots, vibrating objects, gamepads, haptic vests, …

ACKNOWLEDGMENTS

This work was funded by several grants from Paris 8 University and by a doctoral grant from the CLI doctoral school. We thank all the members of the PRIM project and all the participants in the evaluation. And we also thank the students who joined us for an internship to participate in the reflection.

REFERENCES

[1] C. Jost *et al.*, “PRIM Project: Playing and Recording with Interactivity and Multisensoriality,” in *IMX 2021: ACM International Conference on Interactive Media Experiences*, 2021.

[2] C. Jost, B. Le Pevedic, O. El Barraj, and G. Uzan, “MulseBox: Portable multisensory interactive device,” in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, 2019, vol. 2019-Octob, pp. 3956–3961.

[3] M. Chetouani *et al.*, “Cognitive services for elderly people: The ROBADOM project,” 2010.

[4] R. Casillas Figueroa, “Personal spaces for multisensory stimulation as support to rehabilitate patients with cognitive disabilities,” in *Proceedings of the XVIII International Conference on Human Computer Interaction - Interacción ’17*, 2017, pp. 1–2.

[5] A. Covaci, L. Zou, I. Tal, G.-M. Muntean, and G. Ghinea, “Is Multimedia Multisensorial? - A Review of Mulsemedia Systems,” *ACM Comput. Surv.*, vol. 51, no. 5, pp. 1–35, 2018.

[6] D. J. Malan, H. H. Leitner, D. J. Malan, and H. H. Leitner, “Scratch for budding computer scientists,” in *Proceedinds of the 38th SIGCSE technical symposium on Computer science education - SIGCSE ’07*, 2007, vol. 39, no. 1, p. 223.

[7] E. Pot, J. Monceaux, R. Gelin, B. Maisonnier, and A. Robotics, “Choregraphe: A graphical tool for humanoid robot programming,” *Proc. - IEEE Int. Work. Robot Hum. Interact. Commun.*, pp. 46–51, 2009.

[8] N. Fraser, “Ten things we’ve learned from Blockly,” *Proc. - 2015 IEEE Blocks Beyond Work. Blocks Beyond 2015*, pp. 49–50, Dec. 2015.

[9] G. Ghinea and O. Ademoye, “A user perspective of olfaction-enhanced mulsemedia,” in *Proceedings of the International Conference on Management of Emergent Digital EcoSystems - MEDES ’10*, 2010, p. 277.

[10] M. Waltl, B. Rainer, C. Timmerer, and H. Hellwagner, “An end-to-end tool chain for Sensory Experience based on MPEG-V,” *Signal*

Process. Image Commun., vol. 28, no. 2, pp. 136–150, Feb. 2013.

[11] C. Timmerer, B. Rainer, and H. Hellwagner, “Sensory effect dataset and test setups Inter-Destination Multimedia Synchronization View project CONCERT View project,” 2012.

[12] M. Waltl, C. Timmerer, and H. Hellwagner, “A test-bed for quality of multimedia experience evaluation of Sensory Effects,” in *2009 International Workshop on Quality of Multimedia Experience*, 2009, pp. 145–150.

[13] E. B. Saleme and C. A. S. Santos, “PlaySEM: a Platform for Rendering MulSeMedia Compatible with MPEG-V,” in *Proceedings of the 21st Brazilian Symposium on Multimedia and the Web - WebMedia ’15*, 2015, pp. 145–148.

[14] D. P. De Mattos, C. Muchaluat-Saade, Débora, and G. Ghinea, “Beyond Multimedia Authoring: On the Need for Mulsemedia Authoring Tools,” *ACM Comput. Surv.*, vol. 54, no. 7, Jul. 2021.

[15] D. P. De Mattos and D. C. Muchaluat-Saade, “Steve: A hypermedia authoring tool based on the simple interactive multimedia model,” *Proc. ACM Symp. Doc. Eng. 2018, DocEng 2018*, Aug. 2018.

[16] E. B. Saleme, C. A. S. Santos, and G. Ghinea, “Coping with the Challenges of Delivering Multiple Sensorial Media,” *IEEE Multimed.*, pp. 1–1, 2018.

[17] E. Coronado, F. Mastrogiovanni, B. Indurkhya, and G. Venture, “Visual Programming Environments for End-User Development of intelligent and social robots, a systematic review,” *J. Comput. Lang.*, vol. 58, p. 100970, Jun. 2020.

[18] K. Howland and J. Good, “Learning to communicate computationally with Flip: A bi-modal programming language for game creation,” *Comput. Educ.*, vol. 80, pp. 224–240, Jan. 2015.

[19] J. Brooke, “System usability scale (SUS): a quick-and-dirty method of system evaluation user information,” *Reading, UK Digit. Equip. co ltd*, vol. 43, pp. 1–7, 1986.