



ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

New Techno-Humanities

journal homepage: www.elsevier.com/locate/techum

Beyond interactivity and immersion. A kinetic reconceptualization for virtual reality and video games

Enzo D'Armenio*

Marie Skłodowska-Curie Postdoctoral Researcher (IF), University of Liège, Place Crockerill 3, 4000 Liege, Belgium

ARTICLE INFO

Keywords:

Interactive media
Digital media
Visual media
Semiotics
Aesthetics

ABSTRACT

In this paper, we aim to propose a theory for understanding and analyzing the experience of meaning in video games and virtual reality. By focusing on the viewer's movements on the interface and their transformation into movements within virtual worlds, we will attempt to overcome assumptions based on the concepts of interaction and immersion that are at the core of academic thinking. It is not clear what distinguishes the interaction of digital media from the interpretive interaction of traditional media. Similarly, the concept of immersion risks generating confusion between sensory involvement and emotional and narrative involvement. For this reason, we will understand video games and VR as movement-images: experiences that build their meaning and storytelling according to two interrelated sets of qualities. On the one hand, there are the visual qualities that are widely studied in semiotics and in art history, while on the other hand, there are the kinetic qualities, those concerning movement. By postulating the existence of kinetic diagrams, i.e. relations between movements on the interface and virtual movements, we will identify the way in which video games and virtual reality build emerging narratives, combining visual worlds with figurative, abstract or thematic movements. Some examples of virtual experiences — including the cases of *Tetris*, *Half-Life: Alyx*, and *Carne y Arena* — will allow us to show how kinetic syntax and visual syntax are respectively articulated in video games and virtual reality.

1. Introduction

The spread of digital media such as video games, 360-degree cinema, as well as virtual and augmented reality have opened up a new field of research (Harris, 2019, Evans, 2019) in the humanities. Multiple approaches are possible to attempt to understand their commercial success, their use in art and advertising, their influence on violent behavior, and their ideological impact. However, analytical efforts that attempt to analyze how these experiences build meaning, and the distinctive forms of narrative through which they do so, are still limited.

In an attempt to explain the specificities of the experience of these media, labels such as “interactive” or “immersive” media have been used (Ryan 1999; Calleja 2011), starting from the hypothesis of a direct correlation between the increased sensory involvement and engagement of viewers (Catticalà and Eugeni 2020). However, the notion of interactivity is generic and problematic, because the difference between the interpretive interaction entailed by traditional media and the interaction afforded by immersive media has not yet been fully clarified. At the same time, it is unclear how these media, by building perceptual experiences that engage multiple senses, are more immersive on a narrative, emotional, and identity-related level. The consequence is that the specificity of the interaction and immersion built by these new media

needs to be further investigated, before being put into relation with the possible immersion of the viewer and before explaining their storytelling resources.

In this article, we intend to challenge both of these notions because in our view, they prevent us from identifying the mechanisms that articulate the construction of meaning generated by experiences in virtual reality (VR) and video games (VGs). Through a radical re-reading of Gilles Deleuze's theory of cinema (1986, 1989) and of the epistemological assumptions of visual semiotics (Greimas 1989; Dondro 2020), we intend to reconceive interactive experiences as particular types of images that are based on the systematic exploitation of the qualities of movement. Movement has so far been studied almost exclusively with reference to film editing and in particular, to its narrative resources. With the concept of movement-image, Gilles Deleuze (1986) had appropriately recognized that cinematic images are based on an unbreakable relationship between the qualities of images and the qualities of movement, but his proposal did not take into account images outside of the cinematic domain, nor did he explore further the articulations between visual and kinetic properties. More recently, authors such as Pierluigi Basso (2019) and Andrea Pinotti (2017) have defined the images of digital media as environment-images, for the way in which they break with the mimetic tradition of visual representation so as to become

* Correspondence author.

E-mail address: enzo.darmenio@uliege.be<https://doi.org/10.1016/j.techum.2022.04.003>

Available online xxx

2664-3294/© 2022 The Author(s). Published by Elsevier Ltd on behalf of Shanghai Jiao Tong University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)Please cite this article as: E. D'Armenio, Beyond interactivity and immersion. A kinetic reconceptualization for virtual reality and video games, *New Techno-Humanities*, <https://doi.org/10.1016/j.techum.2022.04.003>



Figure 1. A screenshot of the video game *Assassin's Creed 2* (Ubisoft, 2009).

reprogrammable and, foremost, explorable. However, the central role of movement within these virtual explorations has never been taken into account. Our hypothesis is that immersive and interactive media need to be conceived of as movement-images, meaning systems that obey a deep fusion of two syntaxes: the visual syntax, which has already been studied in semiotics and visual studies, and which pertains to the qualities of still images, and a syntax never addressed before, that is, a kinetic syntax which articulates the qualities of the movement itself. The result is that interactive media develop the dual syntax of the movement-image into a complex form: on the one hand, they build traditional forms of storytelling, relying on the conventions of literary, visual or cinematic genres. On the other hand, another kind of storytelling emerges from the practice of interactive movement, adding another narrative layer, that is, the thematic movements performable by the spectator.

In the first part, we will explain the reasons why the concepts of interaction and immersiveness are not fully valid to explain the peculiarity of VR and VG experiences, and why they do not really help to understand the way VR and VGs build meaning and peculiar forms of narration. Instead, we will start from the more general concept of explorable images and then will propose a new interpretive criterion: that of visual movement. VGs and VR are characterized by the requirement for the user to perform movements on the interface, which are then translated into movements within the virtual worlds. To understand the construction of the meaning of these experiences, rather than talking about interaction or immersion, it is necessary to analyze the correlation between the visual and kinetic syntaxes through which particular forms of interaction and immersion are generated.

In the second part, we will start from the tools of visual semiotics in order to elaborate an original methodology capable of describing the visual and kinetic syntaxes underlying interactive experiences. Starting from the case of videogames, we will distinguish the way in which movements on the interface and movements in virtual worlds constitute kinetic diagrams capable of building a second track of meaning in addition to the traditional narration realized through audiovisual devices: a narration which emerges through the practice of play, through figures, themes, and rhythms directly expressed through movement.

In the third part, we will apply the theoretical and methodological model to experiences in VR. They represent an important variation of this formula because they are constructed around a single body, that of the spectator, capable of interacting directly within virtual worlds through the use of a helmet and motion sensors. We will study how the impression of inhabiting a single continuous body is built upon the partial correspondence between movements upon the interface, including perceptive movements, and movements within the virtual world.

Overall, the diagrammatic movement-images allow us to go beyond the theories that describe these media as interactive, opening up the investigation of the kinetic viewer's participation and its semantic and storytelling results.

1.1. *Beyond interaction and immersion: explorable images as movement-images*

By "interactive medium" is usually meant the difference between a traditional medium, for example a book or a movie, and a digital medium such as a video game or a session in VR or AR. If the former does not allow for intervention in the unfolding of the text, the latter, on the contrary, allows the viewer to intervene and actively participate in the construction of meaning. The degree of participation within the interactive media would be much more significant so as to open the field to a new type of experience, one in which the viewer is to some extent also co-creator of the meaning of the text.

However, if we go back to a now classic concept of semiotics and narratology, the concept of the model reader elaborated by Umberto Eco (1979), this difference does not appear fundamental¹. When faced with a written text, readers continuously make interpretive moves, try to anticipate events, identify with certain situations and let themselves be deceived by false leads set up by the text in order to experience surprising effects. The same happens during the so-called interactive experiences, where an interpretive interaction is always at work. Therefore, the difference between the interaction specific to digital media and the interpretive interaction of traditional media should be clarified.

The concept of immersive media instead intends to emphasize the degree of sensory involvement activated by video games and VR compared to traditional media. The former would be able to build an effect of presence based on the stimulation of multiple sensory channels: going beyond the simple cognitive involvement of literature and the auditory-visual regime of cinema, adding a tactile and bodily dimension, these experiences would be globally more immersive and thus engaging. However, the correlation between perceptual participation and overall viewer involvement is not so obvious. A book can involve us deeply in its plot, activate mechanisms of identification with or empathy towards its characters – make us feel emotions. There are few elements that allow sensory involvement to be directly correlated with identity, emotional or narrative involvement. Nevertheless, as with interaction,

¹ The idea of an ideal reader built by the text, who anticipates and prepares the interpretive and sensitive moves of the empirical reader, has been explored, namely, by Wolfgang Iser with the notion of implicit reader (1980).



Figure 2. A recreation of one of the earliest PC versions of *Tetris* (The Tetris Company).

it seems to us that involvement is a characteristic that interactive media share with traditional media. In other words, due to the fact that interpretive interaction and emotional, identity-related, and narrative involvement pertain as much to digital media as they do to traditional media, the expressive and narrative specificities of VGs and VR need to be identified elsewhere.

In our view, we need to take a step back and ask ourselves what kind of expressivity these technologies and experiences embody. In order to clarify the particular nature of interaction and immersion, we first propose to understand them as characterizing a particular type of images, endowed with specific expressive resources. The famous dichotomy proposed by Lessing, 2013 Gotthold Ephraim Lessing in *Laocoön* (1767) allows us to identify some paths of investigation. As is well known, Lessing distinguished the arts of time from the arts of space. Poetry belongs to the first group by virtue of two main characteristics: its primary mode of expression is temporal, being constituted by successions of words, and access to meaning is categorical, because it starts from the general cognitive contents expressed by language to reach sensitive emotions on the basis of particular concatenations. The plastic arts, on the contrary, among which Lessing included painting and sculpture, have a predominantly spatial mode of expression, because they sensitively express specific occurrences, related to the shapes and volumes that constitute them, rather than expressing general concepts through verbal occurrences. In a sense, these two arts build an inverse path of meaning: from cognitive content to sensible perceptions through verbal successions for the temporal arts; from sensible relations to general concepts and emotions for the spatial arts. This does not mean that temporal arts cannot express spatial concepts and vice versa, but to do so, they must use specific procedures, exploiting their secondary mode of expression. According to Lessing and the generous commentary by Jean Petitot (2004), painting and sculpture can rely on non-generic configurations, showing various phases of an action frozen in space in order to convey a narrative unfolding and even general concepts, as in the case of allegories in painting. Similarly, the interplay of rhyme and rhetorical figures allows poetry to build sensitive images, transcending the generality and temporality of their primary mode of expression.

As regards this dichotomy, not only interactive media, but also audiovisual media constitute a significant exception, because they are able to rely on other modes of expression than the purely spatial ones of still images. In his works dedicated to cinema, Gilles Deleuze rightly insisted on the particular case of audiovisual media, understanding them as an inseparable fusion of image and movement: “In short, cinema does not give us an image to which movement is added, it immediately gives us a movement-image. It does give us a section, but a section which is mobile, not an immobile section + abstract movement” (Deleuze 1986, 2). By isolating this theoretical passage from Deleuze’s thought on cinema, one could already assert that in order to understand audiovisual documents, one must take into account the articulations between the qualities of the image (the system of points of view, the plastic and figurative readings) and the qualities of movement (the movements within the shot, between the shots, the rhythms, the tensions, the accelerations, the hesitations, etc.). However, these assumptions were used by Deleuze exclusively to explain cinematic imagery and to propose his classification of the forms of montage that have followed one another since the classical age. The relationship between the qualities of the image and the qualities of movement has not been further studied outside of cinema.

Understanding this connection is, in our view, crucial to the analysis of interactive and immersive media, not only because movement is a central element of the expressiveness of their images, but because it extends to the constraints of kinetic participation submitted to the spectator. Andrea Pinotti has recently insisted on the explorable character of digital images that become true environment-images:

Immersivity and interactivity in virtual environments are able to elicit in the user an intense feeling of “being there”, namely of being embodied in an independent and self-referential world. Images are consequently transformed into habitable environments, which tend to negate themselves as representational images of something — i.e., as icons: they are veritable “an-icons”. (Pinotti 2017, 1).

Our hypothesis is that interactive experiences are grounded in the kinetic participation required from the viewer in order to access and inhabit environmental images. For this reason, we will understand them as movement-images, that is, as the intrinsic fusion of visual and kinetic qualities, in a formula that reverses Deleuze’s original formulation in



Figure 3. A screenshot of *Half-Life: Alyx* (Valve Corporation, 2020).

French (*image-mouvement*) and which insists on the kinetic character of this duality. The advantages of this over the concepts of immersion and interaction are at least twofold. With respect to interaction, kinetic participation emphasizes an exclusive character of the experience of digital media, not replacing but rather complementing the interpretive interaction shared with traditional media. Furthermore, the idea that digital media allow unbounded participation and choice can be challenged: indeed, to access the experience of VGs, VR or AR, viewers cannot participate without any constraint whatsoever, but rather must participate with their movement submitted to the tightly regulated systems of interactive and immersive experience. In fact, each of these expressions involves strictly regulated kinetic systems, which only allow for a very precise margin of intervention by the viewer in terms of types of movements, speed, rhythm, and sequencing of action.

With respect to the concept of immersion, the effect of presence of which Pinotti speaks first of all derives from a kinetic correspondence between the commands performed on the interface by the user and the actions carried out within the game worlds. This is a kinetic identification, similar in a way to the point of view shot in cinema, but realized through movement, which is antecedent to any empathic or narrative immersion effect: it is through movement that the sense of immersion in a virtual environment is realized. Starting from this hypothesis, it is possible to build a theoretical model capable of identifying the specific forms of narrative construction brought into play by the new media, and to specify the particular forms of interpretive interaction and sensory immersion.

2. Analyzing visual and kinetic syntaxes: the form and the body of movement

In order to study visual qualities, semiotics, visual studies and art history have already elaborated very sophisticated models of analysis. While the semiotics of Roland Barthes (1961) limited itself to recognizing the figures present in an image and to attempting ideological analysis based on verbal paraphrases, in compliance with the idea that only verbal language can capture the signification of visual systems, a second and a third generation of semioticians have focused their attention on the expressive specificities of images. Algirdas J. Greimas's article "Figurative Semiotics and the Semiotics of the Plastic Arts" (1989 [1984]) summarizes many of the fundamental achievements of the

emerging visual semiotics resulting from the efforts of a research group composed of authors such as Jean-Marie Floch and Felix Thürlemann.

There are two aspects of visual qualities that we are interested in emphasizing with respect to VGs and VR: first, the way in which images configure access to meaning, and second, the way in which meaning is articulated by fully visual means. With regard to access to meaning, semiotician Jean-François Bordron (2016) has rightly stated that images, as far as bare perceptual experience is concerned, open and project us onto other visual scenes. The presence of frames and, more generally, the edges of the image punctuate the space of representation, isolating the semiotic space of the visual scene they display. In his pioneering article, Greimas details the ways in which this visual access to meaning is achieved in comparison to verbal language. It is not at all a matter of resemblance or mimetic qualities of the figures on the canvas versus the natural world², but the existence of a human reading grid that allows us in the first place to interpret the natural, sensible world, splitting it into objects, processes, and elements. Through this grid, we make sense of the world around us, organizing perceptions and sensations into a fully meaningful form.

It is this grid through which we read which causes the world to signify for us and it does so by allowing us to identify figures as objects, to classify them and link them together, to interpret movements as processes which are attributable or not attributable to subjects, and so on. This grid is of a semantic nature, not visual, auditive, or olfactory. It serves as a "code" for recognition which makes the world intelligible and manageable. Now we can see that it is the projection of this reading grid — a sort of "signified" of the world — onto a painted canvas that allows us to recognize the spectacle it is supposed to represent. (Greimas 1989, 632).

It is this same reading grid that, in the face of an image, allows us to build its meaning, because in encountering the visual traits present on its surface, it configures them as signifiers.

A more attentive examination of the act of semiosis would show that the principal operation constituting it is the selection of a certain num-

² As stressed by Goodman (1968), realism depends above all on the norms in force in a given era. See also the concept of surrogate stimuli developed by Umberto Eco (1999): images are capable of replicating the perceptual effects of real objects or scenes through certain visual features.

Table 1
The procedural semi-symbolism in *Tetris* expressed by the visual syntax.

Semantic categories of content	Plastic and abstract categories of visual expression		
chaos	Vertical	open	multiple
order	Horizontal	closed	unitary

ber of visual features and their subsequent globalization. This is a simultaneous grasping that transforms the bundle of heterogeneous features into a format, that is, into a unit of the signifier. This unit is recognizable, when it is framed by the grid of the signified, as the partial representation of an object from the natural world. (Greimas 1989, 633).

However, the way in which we access images does not tell us much about the ways in which images articulate meaning by their own means. In order to avoid falling into the trap of a verbo-centric model, visual semiotics has elaborated two distinct but mutually influencing readings capable of identifying the expressive specificities of images. The figurative reading concerns the recognition of figures and iconographic themes present in visual works, in a similar way to what has already been proposed by Roland Barthes. Figures, however, can have a variable visual density, ranging from a minimum degree that almost prevents their recognition, to a maximum degree of iconization that ideally builds an effect of realism with respect to a given era and to bare perception. The second type of reading, the plastic reading, was developed for the study of abstract paintings but concerns any type of visual representation. It is a matter of taking into consideration the formal aspects related to the chromatic (colors), eidetic (lines and shapes) and topological (organization of spaces) categories, regardless of whether they go on to constitute recognizable figures. The correlation between semantic and figurative categories on the one hand and formal and plastic categories on the other are capable of describing some of the mechanisms of the articulation of meaning specific to images. It is possible that a particular color takes on a particular meaning within a painting, for example the plastic trait /gold/ associated with the semantic trait /sacred/. This type of association is called symbolism in semiotics. However, it is also possible for a plastic opposition to be correlated with a semantic opposition in an image-specific process of meaning: the opposition between the top of a painting and the bottom can be correlated with the semantic opposition between the sacred and the profane.

Further dimensions in the meaning of images concern the way in which visual documents may not only affirm something by showing it unobstructed, but may also affirm with less intensity, or place obstacles in front of figures, build paths of perceptual discovery or perform genuine negations (Fontanille 1989; Badir and Dondero 2016). Finally, the elaboration of the enunciatory reading (Dondero 2020) has made it possible to describe the way in which the image constructs the viewer's gaze, by staging characters that directly address it, as is often the case in portraits, or by staging events that are autonomous scenes because they do not directly address the viewer. Overall, the plastic, figurative, and enunciatory readings have made it possible to identify meaning mechanisms specific to images and to formalize some of the classic genres in painting and photography (Stoichita 1997).

Interactive media exploit all of these expressive features, which we can summarize using the formula of visual syntax, because like still images, they configure an access to another scene – in this case, to another visual world. Virtual worlds can also be articulated as more or less figurative, or be completely abstract; they can place the viewer in a specific point of view, that of the first or third person, for example, and build well-determined paths for the gaze. They can also use the principles of film editing to tell stories through non-interactive sequences.

However, visual syntax by no means exhausts the expressive qualities of interactive media. In order to understand the ways in which meaning is built, it is necessary to take into account the way in which visual syntax is combined with an original syntax. In addition to visual access and to the articulation of meaning, we need to consider how: a) view-

ers access virtual worlds through their own movements, and b) the expressive results of the articulations between these movements. It is a kinetic syntax that directly concerns the qualities of movement-images. Our hypothesis postulates that it is possible to analyze this syntax on the basis of two general parameters: the body and the form of movement. The body of the movement designates the substrate of the movement – a body which can be either absent from the frame of the image or abstract, present in the form of a human, animal or object, or constitute itself as an indirectly signified body, inferable from the type of movement exhibited. The form of movement concerns the qualities of the movement itself and involves various degrees of abstraction, ranging from purely plastic movements of trajectories, rhythms and accelerations, or from recognizable figurative movements, to thematic movements that can be described through verbal lexicalization.

2.1. Kinetic diagrams in video games

Let's consider the case of kinetic systems set up by video games, before moving on to that of VR. As demonstrated in a previous monograph (D'Armenio, 2014), what characterizes video games is the spectator's obligation to participate in the movement within virtual worlds by moving upon an interface. Usually, the mediation between these two systems of movement takes place through a digital body that we will call a prosthesis: a pointer, an object or a character through which one can explore the virtual worlds and which takes charge of installing a point of view. On the one hand, movements upon the interface are usually abstract and plastic, because they are not expressed through recognizable figures or themes, but through abstract commands such as up, down, left, right, or using other activation keys (A, B, X, Y) and progressive triggers. Movements within the virtual worlds, on the other hand, often translate the abstract movement upon the interface into a system of figurative and thematic movements (running, jumping, shooting, climbing, etc.). Between these two types of movements, a kinetic diagram³ associates the abstract commands upon the interface with more or less figurative and thematic movements within the virtual world. This diagram of kinetic relations establishes the rhythmic and aspectual resonance between the two systems of movement: for example, punctual movements such as jumping within the virtual world may be matched with equally punctual movements upon the interface, while durative movements such as running require prolonged activation of the commands. The result is that interactive media develop the dual syntax of the movement-image into a complex form: on the one hand, they build traditional forms of storytelling, relying on the conventions of literary, visual or cinematic genres. On the other hand, another kind of storytelling emerges from the practice of interactive movement, adding another narrative layer, that is, the thematic movements performable by the spectator. For instance, in the *Assassin's Creed* saga, and in particular in the second chapter (Figure 1),

³ We refer to the definition of diagram proposed by Peirce, 1931–1935 Charles Sanders Peirce (CP 2.277): sensitive representations that in their internal relations replicate the internal relations of another phenomenon. This is the case, for example, of a map that selects some characteristics of the physical space, such as distance and spatial extension, and expresses them visually by replicating some relations of the real object. But it is also the case of geometric demonstrations, with which it is possible to manipulate graphical traits in order to build new knowledge through visual icons instead of categorical concepts. In our hypothesis, the relations between the plastic movements made on the interface are translated into the system of kinetic relations expressed by the movements in the virtual worlds.

Table 2
A schematization of the procedural meaning of *Tetris*.

Kinetic syntax	<i>Plastic form of movement</i>	Increasingly fast automatic descent VS Controlled rotation and displacement		
	<i>Figurative form of movement</i>	Vertical falling VS Horizontal interlocking		
Visual syntax	<i>Visual plastic opposition</i>	Vertical	Open	Multiple
		VS	VS	VS
Procedural meaning	<i>Semantic dynamic</i>	Horizontal	Closed	Unitary
		Clash between chaos and order		

a conspiracy plot in a medieval setting (visual syntax) is set up against a system of commands close to parkour, which allows one to traverse digital cities acrobatically (kinetic syntax).

In this case, the conspiracy thriller provides the background for an acrobatic exploration of cities, with a kinetic obligation to climb monuments and dominate architectures from above, producing vertigo effects in a historical space. It is a very specific and playful way of inhabiting virtual cities, which configures a mixed generic arrangement between a classical detective and espionage type of narration, and a kinetic narration that is rather close to *parkour*. The overlap between visual and kinetic syntaxes builds the peculiarity of this experience. On the one hand, the visual syntax follows the canons of traditional narration, respecting the twists and turns, the investigations, and the motif of the storytelling. On the other hand, the kinetic syntax sets up thematic forms of movement, those of parkour, installing a precise rhythm of exploration, articulated with the adaptable scenarios. The overlap between these two syntaxes is particularly evident in the entity that gives access to the exploration of the virtual world: the digital prosthesis is in this case a character with an identity and a narrative background, but it is at the same time a body of movement that expresses itself kinetically through a precise set of thematic movements.

The so-called genre of walking simulators adopts a very different system of interactive movements, where it is not possible to attack enemies or to defend oneself, but only to investigate a situation by attentively exploring the space, in accordance with a contemplative regime of exploration. The result is a kind of theatrical arrangement of events, where the coupling between the themes imposed by a fixed kinetic system (running, shooting, jumping or simply walking, looking, investigating) and the architecture of virtual cities is central to the meaning of these experiences.

It is also possible for visual and kinetic syntax to work in a kind of dynamic opposition in order to build particular effects of meaning. In *Shadow of the Colossus* (Sony Computer Entertainment, 2005-2018), for example, they build a conflict of modalities through the kinetic obligations: players are required, through movement and actions performed through the character-prosthesis, to do something they do not want to — kill sacred beings called colossi. This conflict between the obligation imposed by the kinetic system through thematic forms of movement (hitting with arrows, with swords, hurting and killing) and the beauty and sacredness of the victims contributes to create a dramatic and desperate effect of *pathos*, based on the correlation between the kinetic obligations imposed by the video game and the somehow innocent sacredness set up by the visual and kinetic system of the colossi to be climbed on and killed.

2.1.1. Semantic concepts through plastic movements: the case of Tetris

The kinetic and visual syntaxes are not, however, always put into the service of the construction of a fulfilled story. In some cases, video games can build more rarefied experiences, in which the meaning is closer to an abstract painting than to a proper narration. This is the case of *Tetris*, the famous puzzle game developed by Aleksej Leonidovič Pažitnov (Figure 2). It is a video game in which, within a vertically stretched space, geometric blocks of seven different shapes and colors

automatically slide from top to bottom and one at a time. The player can rotate these blocks and move them to the right or to the left for the duration of the descent, until they reach the bottom. The goal is to compose horizontal lines with no gaps. When this happens, the line disappears, and the score increases. If the screen is filled with blocks — so if the player was no longer able to make the lines disappear — the game ends⁴.

First of all, we can say that the correspondence of each piece with a single color and a single shape contributes to establish a functional strategy, aimed at keeping the course of the game readable. While the player is placing a piece, the next piece is already displayed in the bottom left corner, and this anticipation allows the player to organize action tactics. Needless to say, the pieces are designed to fit together. Beyond the skills required — a certain manual and coordinative ability, as well as sensory faculties — we believe that it is crucial to go further into the type of intervention that the video game requires of the player. What does it mean to move the pieces to compose continuous horizontal lines? First of all, it means maintaining order. If the player just starts a game, but does not intervene, the blocks start to pile up on top of one another in the center of the empty space, creating a messy vertical column of various pieces, until this column reaches the top of the screen and the game ends.

We can therefore reformulate the analysis of *Tetris* as follows: players install themselves in the experience by temporarily assuming kinetic control of geometrically shaped prostheses. However, the blocks are neither characters nor real figures due to the lack of intentionality and of a recognizable form with respect to the natural world. They are, however, endowed with minimal recognizability, being blocks each with a different shape and color, and they assume a stable form of movement: they continue to descend from top to bottom automatically and, as the game progresses, with increasing speed. The player is impelled to participate by rotating the blocks and moving them, and can only modify the speed by increasing it: with a command on the interface, the player can accelerate the descent of the piece, but cannot by any means slow it down. The goal of the game is to obtain horizontal lines, which at a deep level translates into the semantic value /order/. However, it is not enough to obtain horizontal lines, because the aim of the game presupposes a durative component: it is not enough to achieve order, one must also maintain it.

The fundamental semantic opposition is therefore /order/ vs /disorder/, or better said: /order/ vs /chaos/. The course of the game is

⁴ Tetris is a much-discussed video game within academia. Both the game designer Rouse and Jesper Juul, one of the leading exponents of the ludological current, have indicated Alexey Pajitnov's title as exemplary of their theoretical approach. In particular, the former believes that Tetris does not contain any kind of narrative, and therefore video games do not need stories at all (Rouse 2001, 215), while Juul (2001) reaches the same conclusion through a more mediated reasoning. His opinion is that texts with a narrative component are those that can be translated into another language. With respect to these hypotheses, we intend to propose a different reading: while it is true that Tetris does not tell a story, our analysis aims to show that it is nonetheless endowed with semantic features expressed directly through the articulation of visual syntax and kinetic syntax.

Table 3The classification of immersive media proposed by [Catricalà and Eugeni \(2020, 83\)](#).

Type of media	User's enunciative configuration	Mode of technologically modified presence
Bystanding media	Enunciatee (absence of co- enunciation)	Emptiness (atonic orientation + atonic grasp)
Bystanding-immersive media	Enunciatee + haptic-per- ceptive and/or haptic-narrative co-enunciator	Inanity (atonic orientation + tonic grasp)
Moderate immersive media	Enunciatee + sensorimotor- perceptive co-enunciator	Lack (tonic orientation + atonic grasp)
Radical immersive media	Enunciatee + sensorimotor- perceptive and sensorimotor- narrative co-enunciator	Fullness (tonic orientation + tonic grasp)

continuously sanctioned at a visual level: the more order there is on the screen, the better one is playing. As for the score, it is nothing but a second-level value indicator: a valence. It tells us how much the value of our object is worth: in this case, how much our order is worth ([Table 1](#)).

The dual purpose of the video game — to flatten chaos, to make smooth and linear and horizontal what is angular and vertical, but also to do it well and to sustain it — illustrates the hybrid functioning of interactive meaning. It is a narration of a fundamental semantic opposition, realized through visual and kinetic mechanisms. If we focus on the visual syntax, we can find a semi-symbolism underlying the global functioning of *Tetris*. The opposition between the plastic visual categories /vertical/, /open/ and /multiple/ on the one hand, and /horizontal/, /closed/ and /unitary/ on the other, corresponds to the semantic opposition between /order/ and /chaos/. This opposition is in place during any game and provides its abstract static reading. In other words, *Tetris* primarily organizes a dynamic visual narrative organized according to plastic polarities that are related to a fundamental semantic opposition.

However, it is only by also taking the kinetic syntax into consideration that it is possible to analyze the dynamic meaning of *Tetris*. We have seen that the bodies of movement are geometric blocks of different shapes and colors, and that they are neither characters nor real figures to identify with. The player's movement is scattered through the exertion of control over each of the blocks, without an empathic mechanism being activated with regards to characters and figures equipped with a narrative background. However, it is at the level of the form of movement that a real semantic and kinetic struggle takes place, a struggle between the automatic functioning of the videogame and the intervention required from the player. The form of the movement is shared between an automatic vertical sliding on the one hand, and the rotation and horizontal movement performed according to the player's intervention on the other: the player must try to respond to the increasingly fast sliding of the pieces. This is a real semantic micro-scenario resulting from a kinetic opposition in action, a kinetic clash in which the form of the movement assumes two fundamental figures, one being the kinetic figure of the automatic vertical fall, the other being the horizontal interlock performed by the player ([Table 2](#)).

It is only through the analysis of the kinetic syntax that it is possible to grasp the dynamic meaning of the interactive experience: it is not only a matter of the graphical opposition between order and chaos, but also of the durative and tensive opposition between them, embodied in two repeating and extreme figures of movement. For chaos, the figure of movement is one of vertical falling; for order, it is the figure of horizontal interlocking.

Furthermore, this durative component installs tensive mechanisms. As the player accumulates points, the game accelerates the speed of the descent of the pieces, and the same happens to the music, with a sound loop that acquires more and more rhythm. At a certain point, the speed of the game overcomes any manual skill, and all the games — sooner or later — end up with a full screen, with the victory of vertical descent and chaos.

Therefore, we can not only say that *Tetris*, despite being an abstract video game, has a semantic component, but also that it is effective precisely because it is based on a universal dynamic of meaning: the struggle against chaos. It is a semantic parable expressed by a dynamic that is both visual and kinetic: the clash between the figures of movement linked to the vertical fall, which are increasingly fast and urgent, and the

horizontal interlocking performed by the spectator nevertheless bound to succumb.

3. *The promise of a single kinetic body in virtual reality*

It is possible to apply this theoretical model, based on the dual visual and kinetic syntax, to understand meaning in VR experiences. However, to understand how the two syntaxes are arranged, we need to briefly return to the question of immersion. We have already observed how [Andrea Pinotti](#) opposes two regimes of images to explain immersive experiences in digital media.

“Presentification” rather than representation is here the key issue. Subjects relating to an-icons are no longer visual observers in front of images isolated from the real world by a framing device: they become experiencers living in a quasi-world that offers multisensory stimuli and allows sensorimotor affordances and interactions. ([Pinotti 2017, 1](#))

Rather than building representations that aim to achieve a form of iconic mimicry with respect to the perception of the world, an aim common to a great proportion of Western art, VR on the contrary configures a veritable experience, based on an effect of presence within environment-images.

Other contributions have rightly pointed out that the peculiarity of VR and digital media is to be found in the ways through which this presence is constructed. In a recent article, [Eugeni and Catricalà \(2020\)](#) proposed a typology of immersive media based on the degree of perceptual involvement of the viewers, the degree of incidence of their actions on the narrative, and the degree of freedom in exploring and manipulating the environment.

However, while these two proposals frame one of the strengths and novelties of VR, we may find that the particular perceptual immersion of these experiences is once again centered around movement. Indeed, as [Pinotti](#) explains in his article, VR seems to embody in an exemplary way the mechanisms of remediation described by [Bolter and Grusin in their now classic volume Remediation \(1999\)](#). VR in fact promises to build a transparent and immediate experience, allowing viewers to move directly in a virtual space through their own body. Behind this transparency, however, lies a hypermediation, due to the technological equipment required to build the effect of transparent presence. Not only is it necessary to have a computer powerful enough to process virtual environments, but also to have a sophisticated virtual helmet equipped with screens capable of isolating the surrounding environment in favor of a virtual environment, placed a few centimeters from the viewer's eyes. The VR helmet is also equipped with motion sensors capable of recording and reacting to head movements, orienting the presentation of the virtual world accordingly. Finally, joystick-like interfaces, which are also equipped with buttons and controls in a videogame-like manner, are still needed, and so are accelerometers and gyroscopes capable of detecting movement. Finally, systems based on cameras or other motion sensors must be placed in the room in order to map the virtual space to the real space.

The transparency obtained through technological hypermediation does not, however, modify the fundamental vector of interaction and immersion generated by VR: it is once again the movement of the spectator within the real space that is translated, through the series of interfaces we have described, into movement within the virtual world. The fundamental difference lies in two aspects: the first concerns the expulsion of the real world from the perceptual field of the subject who is immersed

in the virtual world by means of the VR helmet. Second, the centrality of movement and kinetic syntax is complemented by another type of movement, one around which the immersion effect is built: it is the perceptual movement, which is partially integrated into the experience, on account of the correspondence of the movement of the player's head and the movement of the perceptual point of view within the virtual world. In other words, the kinetic syntax of the VR experience promises an almost perfect correspondence between the movement of the player's body and the movement within the virtual environment, going on to constitute a single body capable of moving directly in a simulated space. This is the promise of a single body of movement (the player's) and a form of movement shaped with respect to the directionality of the bodily actions. However, this arrangement does not invalidate the fact that every VR experience entails a particular kinetic obligation, constituted around a fixed and specific repertoire of performable actions (walking, running, observing, picking up, throwing, and so on). In short, it is once again movement that builds the effect of immersion: a movement that seems to construct a single body, associated with the inclusion of perceptual movements in a virtual environment surrounding the viewer.

This promise of immersion realized through the correspondence of the perceptual and concrete movements of the viewer with those of the virtual prosthesis entails important consequences with respect to the video game medium. Since performable actions are tied to the viewer's body, the kinetic diagram does not provide for abstraction between movements made on the interface and movements within the virtual world. If, in the *Super Mario* video game saga, pressing a button several times allows one to perform acrobatic jumps, in VR experiences, the possibilities of action, being linked to the viewer's body, do not allow the building of comparable gestures: the price of motor and perceptual involvement is the inability to transcend the physical possibilities of a human user.

Half-Life: Alyx (Valve Corporation, 2020), one of the most elaborate VR experiences released to date, provides insight into how the setup of virtual spaces and kinetic systems have been reconfigured. Being the latest chapter in a critically acclaimed first-person shooter saga, this is the first episode to be entirely presented in VR, and it therefore allows for a direct comparison with the video game's previous episodes. At the structural level, there are two fundamental differences. First of all, the degree of freedom granted to the player in exploring the environment is greatly diminished, with the progression taking place "on rails", that is to say, with architectural obstacles imposing a strictly obligatory path. Though this kinetic organization represents a limitation compared to the previous chapters, on the other hand, the meticulous layout of the exploration path builds a more scenic experience, with each glimpse and encounter being the result of a more stringent but also more effective architectural direction.

This limitation, however, is accompanied by a greater degree of manipulability of the environment, with drawers and cabinets to open, objects to collect and combine, and more generally, movements that directly involve the player's hands and arms. Although classic commands such as buttons to push on the interface for grabbing and throwing remain present, the correspondence between the actions of the body, and in particular of the hands and head, and the actions of the virtual prosthesis is very high. One example in particular allows us to identify the fundamental difference with video games: whereas in the previous chapters, it was enough to press a button to reload a weapon, in *Half-Life: Alyx*, a command is needed to throw the empty magazine, so one must bring the hand to the back by operating the "grab" command and then bring the new magazine to the weapon (Figure 3).

The body and the form of movement are profoundly modified: from a body as a unitary kinetic entity to be operated with plastic commands that activate complex actions (running, jumping, climbing), here the body becomes articulated into limbs and perceptual movements whereby simply exploring the environment, grasping objects and manipulating them becomes the heart of the experience. The form of movement is modified by moving from an emerging kinetic macro-

narrative, based on complex figurative and thematic movements, to micro-narratives based on proxemic manipulation and perceptual movements⁵.

This is not to imply that VR experiences are any less rich and interesting than video game texts, but simply that their meaning and type of narratives are organized around different kinetic diagrams, which take advantage of the perceptual-motor involvement due to sensory and bodily movements. Many experiences in VR build the effect of perceptual and motor immersion in view of a cognitive contemplation, without relying on complex concatenations of thematic or figurative movements. These limitations can also build peculiar semantic effects. In *Carne y Arena* by Alejandro G. Iñárritu, an opposition is set up between the effect of presence due to perceptual movements and the impossibility of carrying out actions that have an impact. With the aim of capturing the sense of powerlessness and invisibility of Mexican migrants in the desert, this experience requires viewers to take off their shoes and to walk barefoot in the sand. However, the virtual world is neither manipulable nor friendly: despite the presence of other migrants who share the journey, the experience configures an invisible body of movement, an anonymous character that is ignored and powerless, who walks in a daze, without really being able to intervene in the action, and who is pushed back by armed police.

4. Conclusions

In this article, we proposed a theoretical reconceptualization of experiences generated by media such as video games and VR. The main effort has been to challenge and attempt to clarify the notions of interaction and immersion that are at the heart of contemporary academic debate. On the one hand, we challenged the idea that interaction is an aspect that distinguishes digital media from traditional media, since both generate interpretive interaction. On the other hand, we questioned the idea that sensory immersion corresponds to greater viewer involvement, since traditional media such as films and novels can build very strong involvement effects through the mechanisms of narration, identification, and empathy.

To distinguish VG and VR interaction from interpretive interaction, and to study the particular forms of immersive storytelling they generate, we started from the expressive capacities of images and from the recent proposal to understand digital media as image-environments to be explored. Just as images construct a visual access to another scene beyond concrete experience and exploit formal and plastic qualities to articulate particular meanings, so do VGs and VR construct a double access and a double articulation of meaning. On the one hand, access is certainly visual, resulting from the setup of environments that are more or less visually dense. On the other hand, access is also kinetic, because the viewer must move on the interface by means of commands that are translated into movements within the virtual environment. We therefore proposed to adapt the concept of *image-movement* (movement-image) proposed by Deleuze to analyze experiences in VR and VGs. This involves taking into account two different but deeply fused expressive qualities: on the one hand, the visual syntax typical of images, and on the other hand, the kinetic syntax stemming from the articulation of performable movements. With respect to audiovisual media, movement

⁵ Augmented reality constitutes a particular case because one does not find a plastic movement on an interface which is translated in a virtual figurative movement; on the contrary, the figurative movement is autonomous – it does not derive from a prosthesis controlled by the spectator – because it is mainly accomplished by a virtual actor who invades the space of the interface and, through it, the space of the city. At this point, a highly regulated kinetic interaction and a kind of synchronization between plastic and figurative movement opens up. The case of *Pokémon Go* is in this sense enlightening, because it is the virtual figures which, by moving in the city, initiate an interaction that is simultaneously real and virtual.



must be actively produced by the spectator in accordance with systems that are highly regulated in terms of types of performable actions, speed, rhythm, action sequences. For this reason, we have proposed to understand these experiences as movement-images. The kinetic syntax, in particular, can be described on the basis of the body and the form of the movement. While the former concerns the virtual entity that performs the movement (character, object or cursor; absent or indirect body), the latter concerns the analysis of the movement itself. In fact, the movement can be purely formal and plastic (rhythms, accelerations) or build figures, until it becomes openly thematic and lexicalizable (jumping, shooting, running, climbing).

Video games in particular exhibit a diagrammatic relationship between two kinetic systems: movements on the interface are usually abstract and plastic, whereas their realization in game worlds is often figurative and thematic. By taking advantage of visual syntax, this medium can utilize visual genres and cinematic editing and construct characterized architectural environments. However, another narrative track proves more important: the narrative built in practice from the figures, themes, and rhythms of the movement itself. We have seen that while in *Assassin's Creed*, this dual syntax results in a mixture of narrative and acrobatic exploration, in *Tetris*, it is the procedural enactment of a fundamental opposition built with abstract figures and movements: the opposition between order and chaos.

VR experiences build a different kinetic diagram, given first of all by the translation of the perceptual movements of the head of the spectator into virtual movements, which contribute to generate an effect of sensory immersion. The fundamental difference lies in the construction of the promise of a single body of movement, that of the spectator, capable of moving directly in virtual worlds, thanks to sensors, gyroscopes and the helmet that isolates from the real world. However, this setup comes at a price, because since there is no abstraction between the commands on the interface and the virtual movements, the performable actions, and thus also the narrative that emerges pragmatically, remains limited to actions performable by an average human body. Except in rare cases that rely on video game-like configurations in which a key on the interface allows for complex actions, acrobatic movements are excluded in favor of micro-interactions with the environment or of contemplative exploration experiences.

Overall, our proposal equips the investigation of interactive and immersive media with new tools serving to identify the kinetic forms of storytelling and meaning experiences set up by digital media. A more detailed study of richer corpora would allow us to formalize the variety of these experiences around more stable canons, going beyond the aleatoric nature of the current classification proposed by the specialized press, thanks to the peculiar modalities of articulation between kinetic and visual syntax.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 896835 - IMACTIS.(Fig. 1, Fig. 2, Fig. 3, Table 1, 2, 3)

Bibliography

- Badir, Sémir, Dondero, Maria Giulia, Giulia. 2016. "L'image Peut-elle Nier?" *L'image peut-elle nier?*, Liège: Presses universitaires de Liège.
- Barthes, Roland. 1961. "Le message photographique". *Communications* 1: 127-138.
- Pierluigi, Fossali, Basso, Fossali. 2019. "L'image du devenir: le monde en chiffres et la passion du monitoring". *Signata* 10. doi:10.4000/signata.2261. <http://journals.openedition.org/signata/2261>.
- Bolter, Jay David, Grusin, Richard. 1999. "Remediation: Understanding New Media". *Remediation: Understanding New Media*, Cambridge: MIT Press.
- Bordron, Jean François. 2016. "L'énonciation en image: quelques points de repère". In: Colas-Blaise, Marion, Perrin, Laurent, Maria Tore, Gian (Eds.), *L'énonciation aujourd'hui: Un concept clé des sciences du langage*, 227-239, Limoges: Lambert Lucas.
- Callega, Gordon. 2011. "In-Game". *In-Game: From Immersion to Incorporation*, Cambridge, MA: MIT Press.
- Catricalà, Valentino, Eugeni, Ruggero. 2020. "Technologically Modified Self-Centred Worlds. Modes of Presence as Effects of Sense in Virtual, Augmented, Mixed and Extended Reality". In: Biggio, Federico, Santos, Victoria Dos, Giuliana, Gianmarco (Eds.), *Meaning-Making in extended Reality. Senso e virtualità*, 63-90, Roma: Aracne.
- D'Armenio, Enzo. 2014. *Mondi Paralleli: Ripensare l'interattività nei videogiochi*, Milano: Unicopli.
- Deleuze, Gilles. 1986. *Cinema 1: The Movement Image*. Translated by Hugh Tomlinson and Barbara Habberjam, Minneapolis: University of Minnesota Press.
- Deleuze, Gilles. 1989. "Cinema 2: The Time Image". *Cinema 2: The Time Image*, Translated by Hugh Tomlinson and Robert Galeta, Minneapolis: University of Minnesota Press.
- Dondero, Maria Giulia. 2020. "The Language of Images: The Forms and the Forces". *The Language of Images: The Forms and the Forces*, Cham: Springer.
- Eco, Umberto. 1979. *The Role of the Reader: Explorations in the Semiotics of Texts*, Bloomington: Indiana University Press.
- Eco, Umberto. 1999. "Kant and the Platypus: Essays on Language and Cognition". *Kant and the Platypus: Essays on Language and Cognition*, Translated by Alastair McEwen, San Diego: Harcourt.
- Fontanille, Jacques. 1989. "Les espaces subjectifs. Introduction à la Sémiotique de l'observateur". *Les espaces subjectifs: Introduction à la Sémiotique de l'observateur*, Paris: Hachette.
- Greimas, Algirdas Julien. 1989. "Figurative Semiotics and the Semiotics of the Plastic Arts". *New Literary History* 20 (3): 627-649.
- Goodman, Nelson. 1968. *Languages of Art. An Approach to a Theory of Symbols*, London: Bobbs Merrill.
- Iser, Wolfgang. 1980. *The Act of Reading: Act of Reading: A Theory of Aesthetic Response*, Baltimore: Johns Hopkins University Press.
- Juul, Jesper. 2001. "Games Telling Stories?". A brief note on game and narrative, 1: Game Studies <http://gamestudies.org/0101/juul-gts>.
- Lessing, Gotthold Ephraim. 2013. "Laocoön: An Essay on the Limits of Painting and Poetry". Translated by Ellen Frothingham, New York: Dover Publications.
- Peirce, Charles Sanders. 1931-1935. *Collected Papers of Charles Sanders Peirce 1-6*, Cambridge, MA: Harvard University Press.
- Petitot, Jean. 2004. *Morphologie et Esthétique*, Paris: Maisonneuve et Larose.
- Pinotti, Andrea. 2017. "Self-Negating Images: Towards An-Iconology". In: *Proceedings 2017*, 1, 856. doi:10.3390/proceedings1090856.
- Rouse, Richard. 2001. *Game Design: Theory & Practice*, Plano: Wordware Publishing.
- Ryan, Marie Laure. 1999. "Immersion vs. Interactivity: Virtual Reality and Literary Theory". *SubStance* 28 (89): 110-137. No. 2 <https://www.jstor.org/stable/3685793>.
- Stoichita, Victor. 1997. *The Self-Aware Image: An Insight Into Early Modern Meta-Painting*. Translated by Anne-Marie Glasheen, Cambridge: Cambridge University Press.
- Harris, Blake J.. 2019. *The History of the Future: Oculus, Facebook, and the Revolution That Swept Virtual Reality*, New York: Dey Street Books.
- Evans, Leighton. 2019. *The Re-Emergence of Virtual Reality*, New York-London: Routledge.