

Conceptual Blending in Computer Games: Integrating Fiction and Meaning

Marco Caracciolo (University of Bologna, Italy)

marco.caracciolo2@unibo.it

Abstract

In this paper, I will build on Fauconnier and Turner's (2002) *The Way We Think* to introduce my own category of "conceptual blending." This concept can be illustrated by referring to a number of cognitive-scientific sources, and accords with Fauconnier and Turner's single most insightful claim: the idea that we live in a blended world.

Moreover, I'll mention the structural (as opposed to mimetic) resemblance between the way we experience the world and the way we experience aesthetic artifacts, including computer games. By doing so, I will pave the way for what I like to think of as an "enactivist aesthetics."

As to computer games, I'll try to show that they construct blended worlds—that is, worlds structurally similar to the one we live in. In particular, computer games integrate an environment (a world, a fiction) with some affordances. "Affordance" is, obviously, Gibson's (1979) term for an opportunity of interaction; as we'll see, it is also the blanket term through which I intend to capture both the rules of the game and, in case there is one, the narrative it tells.

This integration occurs at the level of the interface and, in particular, of what I will call the 2½-D interface. Eventually, turning to my main case study, *Assassin's Creed*, I'll contend that the interface actually shown on the screen refers to a larger hermeneutic space where the interaction between player and game takes place. This hermeneutic space is *also* blended.

Getting Started

Historically, the idea of "conceptual blending" has its roots in metaphor theory. Fauconnier and Turner's (2002) big book on conceptual blending is, in more than one way, descended from Lakoff and Johnson's (1980) classic. Yet, Fauconnier and Turner's emphasis seems to have shifted markedly: their book is less about metaphors, more about blends. What does this mean? Whereas metaphors are, basically, slender projections from a source to a target domain (they are dual structures), blends are giant creatures, sprawling across *at least* four mental spaces: two input spaces, roughly corresponding to target and source in metaphors; a generic space, which contains the abstract mental structure underlying both input spaces; the blend, where the inputs are brought together (despite their often clashing organizing frames) in order to develop emergent structure.

It would be pointless, here, to lay out the minute details of Fauconnier and Turner's treatment of conceptual blending. As I will make clear shortly, what I'm really interested in is the subtype of conceptual blending that I will call "ceptual blending." For now, let me just point out the two most remarkable insights provided by Fauconnier and Turner's book. The first concerns emergent structure. Consider the personification of death as "the Grim Reaper": Fauconnier and Turner (2002: 291-295; see also Turner 1996: 77-82) have taken pains to show how it results from a rather intricate integration network. I'd like to call attention to the fact that the Grim Reaper is usually depicted wielding a scythe: this tool helps us "unpack" the blend by pointing to one of its input spaces (someone reaping plants). By itself, harvesting has positive connotations in our culture; but scythes do not, even when abstracted from the Grim Reaper. What does this mean? It means that, culturally, the image of the Grim Reaper is so strong that the tool it is associated with has negative overtones. But it is only in the blend that scythes are deadly weapons: in the death input space, no one is literally harvested to death, nor are scythes normally used to kill people (hence, the reaping input space contains plants, not humans). In Fauconnier and Turner's terms, the scythe points to the "emergent structure" developed within the blend (the Grim Reaper is a harvester of souls), which cannot be found in either input space. When we see a scythe, the Grim Reaper integration network is automatically triggered because of its large degree of cultural "entrenchment"; the scythe is *almost* naturally associated with death. All in all, emergent structure is emergent meaning: we'll have to keep this in mind.

As to Fauconnier and Turner's second insight, it dawns on them a few pages before the end of the book. And this is probably its single most profound passage: "Blending is not something we do in addition to living in the world; it is our means of living in the world. Living in the human world is 'living in the blend' or, rather, living in many coordinated blends" (2002: 390). Of course, my example seems to prove this, but at the relatively high level of culture: the sight of a scythe is sufficient to activate the Grim Reaper blend. As I'll try to make clear in the following pages, various cognitive scientists have claimed that even our basic-level perception of the world is blended. In other words: despite being common in literature, advertisements, cockpits and other superfluous things, blending is a basic instrument of our being in the world. It is part of our cognitive toolset; it plays a vital role in our meaning-making.

My last words belong to a blend where COGNITIVE FACULTIES ARE TOOLS. Accordingly, I'll start my discussion of ceptual blending by revisiting Heidegger's famous contention that a tool "bears with it that referential totality within which [it] is encountered" (1978: 99). He added: "*with* this thing, for instance, which is ready-to-hand, and which we accordingly call a 'hammer', there is an involvement in hammering; with hammering, there is an involvement in making something fast" (1978: 116, emphasis in the original). The meaning of things is what we can do with them, what they are means of (at least on this informal reading of Heidegger's words). This idea resonates with the recent attempts in cognitive science (and in particular in "enactivist" approaches) to show that perception and concepts are intertwined: on the one hand, Alva Noë has argued that perception depends on our practical, "proto-conceptual" (2004: 183) understanding of our sensorimotor skills. For instance, an occluded object is perceived as virtually present because we know that, if we changed perspective by moving our eyes or our body, it would be in full sight. As simple as it sounds, this statement commits the author to a rejection of the view of perceptions as static, snapshot-like internal representations (the prevailing view in first-generation cognitive science). But if perceptions are, to some extent, conceptual, concepts are ultimately based on "embodied simulations" (Gallese and Lakoff 2005) or image schemas (see Johnson 2007: 136-145).

These findings (resulting from both neuroscientific and cognitive-linguistic research) would seem to confirm Heidegger’s contention that objects are characterized by their readiness-to-hand. The perception of a hammer triggers the concepts of hammering and of making something fast, because concepts are nothing else than “simulated potential actions” or “plans” to act (Gallese 2005: 25; see also Gallese and Lakoff 2005: 6). The same would happen if someone read the word “hammer” or saw a depicted hammer: as Freedberg and Gallese write, “[t]he observation of a graspable object leads to the simulation of the motor act that the object affords” (2007: 200). Like this one:



Figure 1

Screenshot from *Fallout 3*
(Bethesda Softworks, 2008)

This is a hammer in a computer game. It doesn’t just *look like* a hammer, *it is* one, as the interface tells us. Let’s consider what kind of interaction this hammer affords. We can pick it up by pressing the button “E” on the keyboard. And we can pick it up by pressing the button “Z” on the keyboard. What? Obviously, I’m playing on two different meanings of the phrase “pick it up”: if we press “E,” we add the hammer to the protagonist’s inventory, as suggested by the interface; if we press “Z,” we watch the hammer hover in mid-air and are allowed to reposition it in the game world by looking around and dropping it wherever we want. We are given two options, both of which seem relatively inconsequential in this case, since this is a humble hammer, it can’t even be equipped—that is to say, used as a weapon—in *Fallout 3*. Suppose it’s a powerful weapon then, or simply one of the “Super Sledge Hammers” wielded by some Super Mutants in the ruins of Washington D.C.. In this game, you can hold a weapon in your hand only if you’ve previously added it to your inventory by “collecting” it (pressing “E”). Weapons are meant to extend the reach and increase the strenght of our body: they add to our physical abilities, and this is perhaps why they figure so significantly in computer games. (How many of them are really about collecting objects, such as weapons or armor, that improve our existing abilities or provide new ones?) On the other hand, some forms of emergent gameplay (Juul 2005: 76) take advantage of the “grab” function: for instance, containers such as barrels can be stacked and climbed to reach hidden recesses, or valuable objects can be moved to a position from which it is easier to steal them.¹

At first, the choice between “grab” and “collect” may appear awkward, since in the real world grabbing, collecting, and holding something in your hand belong to a sensorimotor continuum, they are not distinct tasks. However, this choice is dictated by the level of abstraction (see Juul 2007) implicit in any computer game. Players adjust rapidly to it: with some practice, they come to understand that all objects in *Fallout 3* can be interacted with in

¹ Both techniques are described here: <http://fallout.wikia.com/wiki/Grab> (accessed August 6th, 2009).

two ways. It is no chance that the two affordances (Gibson's (1979) famous term for "opportunities of interaction with the environment") are introduced in the tutorial, when the protagonist is a small child: the learning curve is thus fictionally motivated. Although it is probably true that most meaningful actions are performed through equipped weapons, both ways of interacting with objects enhance our structural coupling with the environment; they improve our survival rate either directly (we collect a weapon and start shooting at enemies), or indirectly, by allowing us to obtain potentially beneficial objects, as in the techniques described above. My point is, however, that both affordances become rapidly integrated into the object itself: when experienced players see a weapon, they'll recognize immediately what its affordances are; they'll pick it up by pressing "E" without thinking twice. Likewise, when they realize that a certain valuable object can't be reached, they'll cast about for something to jump upon and move it by pressing "Z." And even if, for some reason, the players decide not to do anything, they remain aware of the object's affordances, its "meanings" in the game. The actions one *could* perform with an object, I contend, become one with it because they activate the relevant embodied simulations.

This is, in short, what I call "ceptual blending." Fauconnier and Turner (2002) never really employ this term, but the idea behind it is largely inspired by their work.² In particular, Turner writes: "Whenever we see something *as* something—when we look at the street and see *a woman getting into a car*—we are blending our sensory experience with abstract conceptual structure" (1996: 112, emphasis in the original). Where do we draw the line between our perceptual world and our concepts? The line can be drawn and *has been* drawn, but what is sure—if the body of research mentioned earlier is correct—is that we don't live by it. Since the very world we live in is blended, we can never disentangle ourselves from the blending of concepts and percepts: we are entangled or, as Heidegger writes, "involved" in our blended world. Therefore, we should steer clear of the "natural attitude" according to which concepts are in our mind and percepts are in the world. This would imply assuming a representational stance on cognition, since we can "take in" perceived objects only by mentally representing them. But this myth has been shattered by recent cognitive science. There's no need to download the world to our mind to make sense of it: we do that just as well by interacting with the world itself, which is—in Brooks's famous phrase—"its own best model" (1991: 583). This is the main thrust of the enactive approach to cognition. The phenomenon I call "ceptual blending" shows that our cognition is off-loaded onto the environment: as we'll see, we live by enacting a world. In a way, ceptual blending is neatly captured by an object in a computer game: its perception is blended with the concept we frame of it, which in turn is based on the range of actions we could carry out with it.

Doors

We know, however, that aesthetic artifacts like novels, paintings and computer games place constraints on their users; or, to put it the other way around, they give them some affordances and not others. Think of a recent game like *Grand Theft Auto 4* (Rockstar North, 2008): it displays a perceptively rich, even luxuriant world—only to severely limit the player's interaction with it. For instance, there is no apparent reason why we can smash only a very few shop windows in the whole Liberty City. Aarseth (2005) has made a similar point in relation to doors in games: those we can't interact with are, he contends, "fictional"; those we can open are "virtual" or "simulational." Are we to conclude that game worlds are not fully

² The term "ception" is used to indicate both perception and conception (and to question the validity of their distinction) by Talmy (1996).

blended, that “fictional” doors only look like doors, but don’t do what doors are supposed to do? I don’t think so. Or, perhaps, they are poorly blended, if we assume that in-game doors should behave just like their real-world counterparts. A mimetic fallacy may be lurking in Aarseth’s argument: after all, from a real-world perspective, what can be said to be really “simulational” or “virtual”? In most games, a simulated door (in Aarseth’s sense) can be opened; but can it be unhinged? Can we shoot at it and peep through the bullet hole? Can we scrape off its varnish? No, and—at least from my point of view—this is not even what video games are about. We don’t need a double of the real world: one’s more than enough. As game designer Frank Lantz (quoted in Juul 2007: 511-512) has put it, in a total simulation there would be nothing to do, except perhaps reinventing chess! We need constraints or affordances (some doors can be opened, other can’t) because only within their boundaries can our interaction with the game be truly meaningful.

This is why I tend to be skeptical of binary dichotomies such as Aarseth’s (or Frasca’s (2003) similar distinction between simulation and representation). Of course, these dichotomies might be valid at some level: for instance, they are probably convenient for discussing how game worlds have been intentionally designed by those who made them. I would only suggest to find less metaphysically committed terms, since Aarseth’s distinction boils down to what objects can be interacted with, and what can’t. But—I contend—it plays absolutely no role in the player’s experience, for the simple reason that there’s no way to distinguish “simulational” and “fictional” objects *a priori*, without actually playing the game (i.e. getting to know its affordances). Take the prologue of *Max Payne 2* (Remedy Entertainment, 2003): Max wakes up at a hospital in a bad shape; he’s received multiple gunshot wounds and, what is more, has a “blindspot” in his head, “a bullet-shaped hole where the answers should be.”³ These words point out Max Payne’s most urgent problem: making sense of his present situation, clearing up the radical doubts surrounding him. Better than any theoretical statement, this beginning shows that sense-making and interpretation are *not* (just) scholarly hobbyhorses: they are the survival strategies by which we cope with our environment. Describing a cellular automata, Varela, Thompson, and Rosch have argued that even its fairly simple structural coupling with its environment involves “a minimal kind of interpretation . . . where interpretation is understood widely to mean the enactment of a domain of distinctions out of a background” (1991: 156). Our world, they add, does not pre-exist us: it is enacted in our interaction with it. (Consider how much play-time goes to opening closets or other containers in order to see what’s inside, and consider that objects found there—such as ammunition or first aid kits—could literally save your life: isn’t this a textbook example of what “structural coupling with the environment” means?)

Now, consider that Max Payne’s predicament is directly reflected in the player’s. When we start playing *Max Payne 2*, we are given no explicit goals or missions. The game’s narrative unfolds, emerges from the player’s attempts at doing something meaningful within the game world. This results inevitably in a trial-and-error process: the player tries to familiarize with the affordances of the game world. (Think of the importance of “tutorials” at the beginning of most games.) Not all actions are allowed, and not all doors (to return to my criticism of Aarseth’s dichotomy) can be opened. For instance, this door can’t:

³ From the game script, <http://faqs.ign.com/articles/446/446198p1.html> (accessed August 6th, 2009).

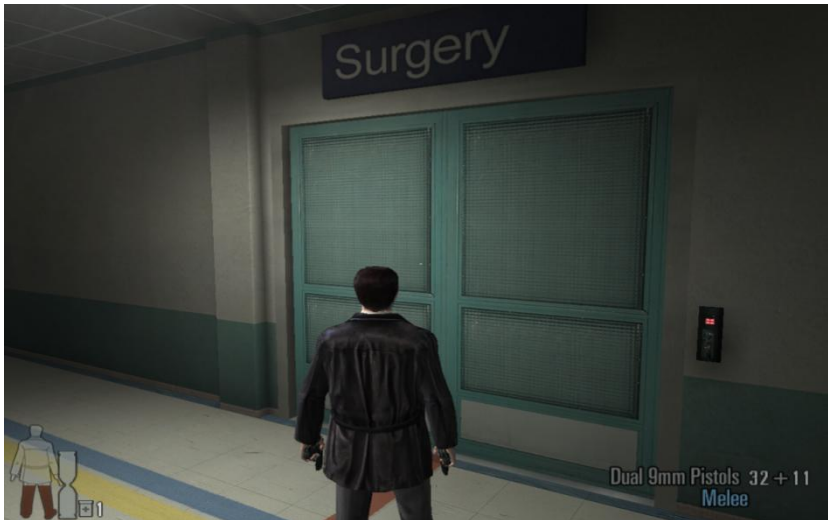


Figure 2

Screenshot from *Max Payne 2*
(Remedy Entertainment, 2003)

But how is the player supposed to know that this door won't open? The red light on the switch next to the door could be a clue; indeed, it becomes one – but only when the player has understood that, in this game, the red light signals that the door is closed. Hypothetically, it could mean that the area beyond it is unsafe. But since we are at the very beginning of the game, there's no other way than actually trying to open it. Is this door simulational/virtual or fictional? We must concede at least that it is virtual until one attempts to open it, since the player takes the computer game as a blend in which, in Fauconnier and Turner's (2002) terms, an outer-space relation of Representation (we see a door-like shape *as* a door) is compressed into an Identity relation (we see a door *in* the screen⁴). The door *must be* virtual because all objects trigger in those who perceive them specific "plans" to act (remember Heidegger's hammer); thus, our concept of "door" (which is blended with the door we are perceiving) is related to a series of (virtual) actions, including "opening."⁵

The door, however, doesn't open. Will the player slap his forehead at this point, exclaiming: "Of course it doesn't open, it's fictional!"? I find that picture hardly believable. But I won't deploy here the naïve (yet to a certain extent reasonable) argument that doors can be closed in computer games as well as in real life, and that if Max is unable to open the door, it just means it's closed.⁶ As I said, players don't know beforehand what is fictional and what is virtual. This is because their actions determine but at the same time *are determined* by their immediate surroundings: more than the real world (and certainly more evidently), game worlds are oriented, narratively or otherwise. This is a more abstract way of saying that, soon or later, Max Payne will find himself inside the elevator shaft, and the prologue will end. If this is to happen, the door to the surgery ward must remain closed: for instance, we can safely assume that opening it would put the character on the wrong track.

We've found out that binary oppositions such as "fictional" vs. "virtual/simulational" are too one-dimensional to account for the player's phenomenological experience of the game. Intuitively, the objects we encounter in game worlds do not belong to two ontologically distinct categories (fictions and simulations), but make up a continuous, highly integrated (blended) whole. The door shown in Figure 2 doesn't look any different after realizing that it

⁴ I refer to Wollheim's (1980) key opposition between "seeing-as" and "seeing-in."

⁵ This happens because we rely on Marie-Laure Ryan's (1991) "principle of minimal departure" to interpret fictional worlds: if we see a door in a computer game, we understand it as being as similar to real-world doors as the game allows us to.

⁶ Besides, this argument wouldn't apply to the unbreakable shop windows of *Grand Theft Auto 4*.

is closed (i.e. it can't be opened). It's not even true that we are not allowed to interact with it, as Aarseth's argument goes, because in fact we do every time we try to open it. From my perspective, even inferring from the red light that the door must be closed is a way of interacting with the door. So, if we wanted to preserve Aarseth's terminology at all costs, we could say that *all* doors in computer games are simulated, because their recognition as doors triggers specific action simulations (as defined by Gallese (2005)). If we didn't see a door *in* the screen, why would we even attempt to open it? The fact that only some doors can be opened can be accounted for at another level: if the player's interaction with the game is to be meaningful at all, the game world must be oriented. All in all, we shouldn't take Varela, Thompson, and Rosch's (1991) claim that the world doesn't pre-exist us as implying that we enjoy absolute freedom, since we create our own world (an extreme form of subjectivism). On the contrary, an organism "brings forth a world" (Varela, Thompson, and Rosch 1991: 205) because its anatomical structure and the world's affordances arise "codependently." What does this mean? That neither pre-exists and pre-structures the other: they are interlocked, and the world exists only *in* the interaction between its affordances and our sensorimotor capacities; to put it more briefly, it is enacted. Johnson (2007: 118) warns against the use of the term "interaction," for—he writes—"it carries the misleading implication that there are two or more independent 'things' that are interacting." We'll have to remember his caveat, but this concept is so pregnant in the context of computer games that the payoffs of using it easily outweigh its shortcomings. It is hard to suspend our "natural attitude" towards computer games (the idea that there is, dualistically, a player on one side of the screen and a game world on the other), but everyone will concur that the player's experience actually unfolds as an interaction. So far we've explored only one side of the screen, the world's affordances (some doors open, others don't), abstracting from the how the player acts on the screen to interact with the game world. In the next section, I'll try to factor it in.

Two and a Half Dimensions

Computer games place constraints on the environment's affordances, as we've seen; but they place them on the character's sensorimotor skills as well (and consequently on the player's, if we consider that the way we interact with games is *also* sensorimotor). For instance, in the so-called "zombie mode" of *Call to Duty: World at War* (Treyarch, 2008) you can't jump over objects. This is probably intended to preclude forms of emergent gameplay that would make it far too easy to defeat the hordes of zombies (such as shooting from where they can't reach you). From the player's perspective, however, it's almost impossible to say if this constraint is put on the character or on the game world: you *can* jump, but you can't jump over objects. And no player, I believe, would ever quit the game because you can't jump over objects; players usually make do with what they have, as long as the game is enjoyable. This shows that even though players transport to game worlds with real-world assumptions, and even though these assumptions facilitate their adaptation to the game at first, players are more than willing to abandon them whenever the game responds differently. This shows, additionally, that you can't really draw a line between affordances and sensorimotor skills, because they are in a reciprocal, symbiotic relationship.

At what level does this relationship exist? It's time to introduce another key concept: the interface. It is at this level, I claim, that ceptual blending actually occurs. Much work has been done on interface as metaphor (Rohrer 1995; Fauconnier and Turner 2002: 22-24) and also on conceptual blending in human-computer interaction (Dourish 2001; Imaz and Benyon 2007). Dourish, in particular, has argued that "Embodied Interaction is the creation, manipulation, and sharing of meaning through engaged interaction with artifacts" (2001: 126). I'd like to

bring the discussion one step further by examining how interface-driven ceptual blending works in aesthetic artifacts like computer games. What I find interesting about computer games is that their GUI has an ontic rather than an epistemic function (I owe this distinction to Brey 2005): unlike most computer programs, they do not extend and enhance our cognitive faculties directly, but they project worlds by generating an enactive (or interactive) blend that structurally resembles the way we are engaged with the real world. As we've seen, both in reality and in computer games we enact a world by interpreting it; but since game worlds are so small and unambiguous (as compared to the one we live in), interpretation itself is foregrounded.⁷ Moreover, computer games share this feature with all works of art—and this is why I keep calling them “aesthetic artifacts,” even if this idea would undoubtedly meet with strong resistance in some environments. Building on John Dewey’s 1934 *Art as Experience* (2005), Mark Johnson is one of the very few philosophers who have made a case for the relationship between art and embodied meaning: “we must realize—he writes (2007: 212)—that aesthetics is about the condition of experience as such, and art is a culmination of the possibility of meaning in experience.” For obvious reasons, I can’t expand on this point here, but I’d like to keep it as a theoretical backdrop to my argument.



Figure 3
Screenshot from *World of Warcraft* (Blizzard Entertainment, 2004)

Consider the graphical interface presented in Figure 3. Needless to say, there is a stark contrast between the three-dimensionality of the game world and the two-dimensionality of the window, the map, and the buttons at the edges of the screen. As we might expect, these interface elements are based on various cultural interfaces (in Lev Manovich’s (2001) phrase): the space inside the window is designed to resemble the yellowed page of an old book, whereas the window itself and the buttons exploit the usual HCI conventions (whose direct historical antecedents are, according to Manovich (2001: 92), the control panels on cars and planes). The map is, obviously, another pre-digital instrument referred to here. All of this is pretty ordinary, however: there is indeed some conceptual blending at work, but nothing too

⁷ Among the theorists who have drawn attention to the nexus between computer games and meaning are Salen and Zimmerman (2004), Juul (2005), and Ferri (2007). Towards the end of his *Half-Real*, Juul writes: “Games—like stories—are things we use to relate to death and disaster. . . . [They] are playgrounds where players can experiment with doing things they would or would not normally do” (2005: 193). I completely agree with him on this point. Ferri (2007) addresses the same problem from a semiotic perspective.

sophisticated. (The bulky interface is probably motivated by the large degree of control required by the MMOG genre.)

More interesting is, from my perspective, the yellow circle on the ground and the exclamation mark floating over the goblin. These elements do not seem to belong to the game world, since they are clearly symbolic (the exclamation point marks off quest givers) or indexical (the yellow circle indicates the selected character) in nature. They are prone to disappear or change under certain circumstances. Besides, characters do not seem to notice them, nor do they impede our movements. But the point is that these elements do not belong to the interface either, since they are clearly positioned *within* the game world and are, at least to a certain extent, represented three-dimensionally. To mimic Marr's (1982) three-level theory of vision, we could say that, if the game world is (in the games I'll take into consideration here) 3-D and the regular interface is 2-D, this is a 2½-D interface. It is *in* the world but clearly not *of* that world. Other examples would include crosshairs at the center of the screen and all kinds of pointers. These interface elements have a common purpose: they are for the guidance of action, they are meant to assist the player's perception of the game world.

Take a crosshair, for instance. As I've mentioned before, Alva Noë's (2004) enactive approach hinges on the concept of the virtuality of our perceptions: although we never take in the world all at once and in full detail, as if in a snapshot, we still experience it as virtually present because we know that a slight movement of our body or eyes would be sufficient to reveal what was previously out of sight. In other words, we make up for the defects of our vision (we have a limited field of vision, we don't see colors at the borders, we are change and different blind, and so on) by *enacting* perception: we perceive by actively exploring the world. This is precisely what happens in some computer games. In the online multiplayer of *Call to Duty: World at War*, a single shot can kill you. This implies that, to be successful in the game, you can't afford looking in the same direction for more than a few seconds; if you want to survive, you need to move both yourself and your "eyes" continuously in order to spot hidden enemies and take cover. Paradoxically, off-screen space is all that matters here: good players take in the environment in very quick glances; they feel an almost natural urge to run and look around. This results in a frantic gameplay and in a turmoil of on-screen images. But this flux is structured by the crosshair at the center of the screen: you always know what you're looking at and, even more important in a game like this, you always know what you're firing at. In a way, the crosshair is (or should be) the sole focus of the player's attention: it assists his meaningful interaction (his structural coupling) with the game world.

More in general, whether the game is first- or third-person, the crosshair works as a structural backdrop to the player's virtual vision. At the same time, it is part of the (hidden) layer at which ceptual blending actually occurs. Imagine a counterfactual scenario: we play a customized version of *Call to Duty: World at War* in which we never know where our bullets will land; they are *totally* random. We could gun down a player we don't even see, or two players could look at each other and pull the trigger in vain. In such a game, which obviously doesn't need a crosshair, there would be no way to *make sense* of what goes on in the game world. This is the function of the 2½-D interface: it allows our action to be meaningful within the game world, it blends meaning and perception (or rather, it shows that they are blended from the start). The same applies to the yellow circle under the goblin in Figure 3: the character is selected, which means that we designate the goblin as the object of the actions we carry out (by pressing a button on the keyboard or on the 2-D interface). All in all, the 2½-D interface off-loads our cognition onto the environment.

Before moving on, I'd like to draw attention to the other 2½-D element in Figure 3, the exclamation mark. Everything that has been said so far holds true for it, except that the background of meaning on which that symbol is projected is larger and, possibly, more comprehensive. So far, I have followed Johnson (2007) in adopting an embodied and non-propositional view of meaning: in a game like *Call to Duty: World at War*, shooting at an enemy *is* meaning-making, just as Varela, Thompson and Rosch's (1991) cellular automata can be said to "interpret" its environment by enacting it. But if it is true that meaning sometimes arises from basic sensorimotor activities, it also comes in more sophisticated flavors, as with propositional and linguistic meaning (which we usually regard as meaning proper). The idea that narrative is a "powerful tool for rendering the world cognizable, manageable, and rememberable" (Herman 2002: 24) has become almost commonplace today. And it goes without saying that narratives rely on and generate conceptual (propositional and linguistic) meaning.

Now, the exclamation mark over the goblin's head sets up a complex blend: a generic symbol is conventionally associated with the idea that the character has a quest for you. By using symbols such as this within the game world, the game achieves a seamless integration between the narrativity of quests and the game world itself. It is well known that most MMOG players do not even bother reading the background information and details of the quests they accept, since they are not really interested in the narrative *per se*. However, in a way, they are involved in the basic narrative that unfolds *within the game world* as a constant succession of grey exclamation mark ("too early for this quest"), yellow exclamation mark ("I have a quest for you"), grey question mark ("come back when you have completed the quest"), and yellow question mark ("quest completed, you may claim your reward"). Quest symbols in *World of Warcraft* condense usually elaborate stories into successions of elementary states, adhering to Fauconnier and Turner's (2002) golden rule ("achieve human scale") for the player's benefit. Thus, when arriving at a new location, players accept all of the available quests before embarking on one of them: this can happen because quest givers are clearly marked off from other non-playing characters. Not unlike crosshairs and cursors, these symbols blend (at the level of the 2½-D interface) meaning with the game world, except that here meaning stems from an overarching narrative.

Anomalous Interfaces

Before coming to the conclusions I'd like to examine two recent games where the 2½-D interface figures prominently. The first of them is *Dead Space* (EA Redwood Shores, 2008), a survival horror game repeatedly praised for its immersive quality.⁸ This proves that, despite having been famously (and convincingly) debunked by Salen and Zimmerman (2004), the so-called "immersive fallacy" is still alive and kicking. But then again, what this game accomplishes is truly remarkable, as Figure 4 shows.

⁸ For instance, from Jeff Haynes's review for IGN: "Because you're never removed from the action, you feel much more immersed in the world" (<http://pc.ign.com/articles/923/923212p2.html>, accessed August 6th, 2009).



Figure 4

Screenshot from *Dead Space*
(EA Redwood Shores, 2008)

This almost looks like the standard 2-D interface of an inventory, but moving the camera around it shows that it is in fact a holographic projection *internal* to the game world. (Of course, this trick is convincing because of the sci-fi setting of the game.) As a consequence, the character always turns his head towards the cursor, as if he were actually looking at the selected object (or rather at its icon). Of course, it seems odd that someone should need a sophisticated holographic panel to use an object that he supposedly carries with him. But the impression most players get from this GUI design is that they are acting *directly* on the game world, without the mediation of a 2-D interface layer (which—it is usually thought—spoils the immersion because it adds a meta-fictional layer, exposing the game as game and not as world, in Marie-Laure Ryan’s (2001) terms). The player’s interaction with the interface is replaced by the character’s, so that the player is apparently excluded from the game world: since there is no 2-D interface players can act upon *qua* players, they seem to vanish altogether. Instead, as a result of the great degree of embodiment afforded by third-person games, players project onto the character’s body: as Laurie Taylor (2003) points out from a Lacanian perspective, third-person games allow players to act *within* the screen, not *on* the screen. Returning to the framework I’ve developed in the previous pages, the structural coupling between the character and his environment almost overlaps with the structural coupling between the player and the game. This is an impressive example of ceptual blending: not only is the player’s interaction with the interface blended with the character’s interaction with the game world (which happens in most games), but the interface itself is fully integrated into the game world. Take the “medium air can” icon shown in Figure 4: in a standard 2-D interface, it would prompt a basic blend between an interface element and a game world object; in *Dead Space*, that interface element is a game world object in itself, because (unlike the 2½-D interfaces we’ve considered so far) it can be perceived by the protagonist. Hence, the unique interface of *Dead Space* comes close to being a 3-D interface.

It should be noted that this 3-D interface revolves around two virtual bodies: the character’s body, which is “virtual” both in the conventional sense (computer-generated) and in the one described by Noë (2004) (it determines the sensorimotor patterns through which the game world is perceived); the reader’s body, which can be seen as virtual because it runs “embodied simulations” (Gallese 2005) of the character’s actions (as shown on the screen).⁹ These two bodies are, again, blended at some level. But *at what level*, we may ask, since players are

⁹ Gallese contends (personal communication, June 26th, 2009) that the concept of “virtual body” is unnecessary, since embodied simulations are rooted in our real body. But I believe that the ontological divide between real and fictional worlds calls for such a distinction. The problem is, obviously, too big to be addressed here.

made to believe that they are acting *within* the game world? Indeed, in *Dead Space* there appears to be no intermediate layer between the game world and the real world. We don't want to give in to the immersive fallacy, both because it's theoretically unsound and because no player seriously believes he's literally inside the game. Immersion is just a metaphor through which we strive to understand the relation between an aesthetic artifact and its user; its main shortcoming is that it tries to conceal that relation behind a wrongheaded notion of "how it feels to be immersed in the real world." However, we must recognize that we've come to a standstill: the sole focus on immersion of *Dead Space* doesn't give us any additional insight into the problem of the interface. It is time to turn to my second case study.

Its gameplay *per se* may be repetitive, but *Assassin's Creed* (Ubisoft Montreal, 2007) features an impressive packaging (at least for the theoretically minded): the story of Altaïr, a member of a secret brotherhood in 12th century Palestine, is relived by a present-day bartender thanks to a machine (called "the Animus") designed to discover mnemonic traces of an ancestor in someone's DNA. This Chinese-box structure, however, is more than a background to the actual game: even though the primary focus of *Assassin's Creed* is on the embedded world, the frame was clearly meant to be part of the game itself, and even plays an important role (narratively speaking) in the game's ending. Figure 5 shows what the protagonist, Desmond, sees from his horizontal position in the Animus when he's not "connected" to his genetic memories.

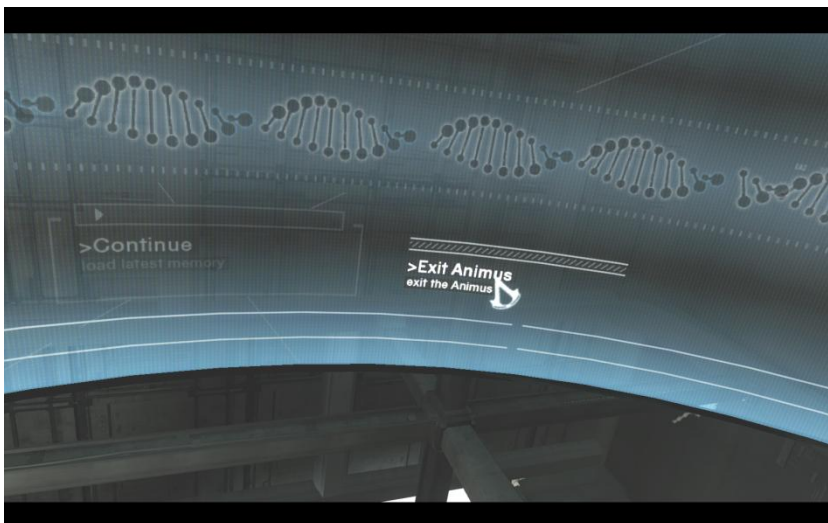


Figure 5

Screenshot from *Assassin's Creed* (Ubisoft Montreal, 2007)

The Chinese-box structure of *Assassin's Creed* is reflected in the complex procedure required to quit the game: when, after my first game session, I hit Esc and selected "Exit Memory" from the menu, I didn't find myself in front of my desktop, as I had expected, but in front of the Animus, as in Figure 5. At this point I selected "Exit Animus" in the hope that the game would eventually close, only to see Desmond standing beside the Animus in the primary game world! To quit the game for good, you need to hit Esc yet another time, log into your game profile and choose "Exit Game" from the main menu. Now, this procedure is certainly less obtrusive in the console version (most console players would just turn off the machine), but in the PC version it seems awkward—to say the least. (Of course, there are other, more forceful methods to quit, even on a PC, but this is the only one built into the game itself.)

The Animus is a perfect metaphor of the interface as the space where the interaction between player and game takes place. Naturally, adding to the game such a conspicuous meta-fictional layer (even if disguised as a frame narrative) could be said to disrupt our sense of immersion. But as long as it remains at the borders, so to speak, of the main action (set in the 12th century)

it is no different from an alarm clock beeping too far away to be heard distinctly. Since players have to pass through this layer in order to actually play the game, their sense of immersion could even be heightened by their identification with Desmond reclining in his virtual reality machine. But this, however, doesn't happen; unlike *Dead Space*, this is no immersive game—because of the glitches. These “glitches,” as they are called by the game instructions, are hexagonal (nucleotide-like) shapes, strings of characters, or other abstract forms that from time to time flash on the screen while the player is controlling Altaïr. Indeed, they are everywhere in the game world: they appear during loading times and when Altaïr completes a mission, they mark the bonuses scattered throughout the game world, they signal that by pressing any key players can change their viewing angle during cut-scenes, they figure on the virtual walls that delimit inaccessible parts of cities (see Figure 6).



Figure 6

Screenshot from *Assassin's Creed* (Ubisoft Montreal, 2007)

All in all, glitches are what Wolfgang Iser (1978) called “alien associations”: they continuously remind us that this is just a game by drawing attention to the fictionality of the game world. To put it in Iser's words, “consistency-building brings in its wake all those elements that cannot be integrated into the Gestalt of the moment” (1978: 126). Glitches are spectacular examples of 2½-D interfaces: they are *within* the game world but do not belong to it. For instance, why are there small, identical flags throughout the Holy Land? Needless to say, the internal logic of the game world offers no explanation. It is obvious that, since this is a game, the flags are meant to be collected: this is why they are surrounded by glitches. These glitches arise, then, from the rules of the game and indicate the constraints put on the game world (or its affordances): in a way, they are material anchors (see Fauconnier and Turner 2002: 195) for a megablend, which ultimately coincides with the game itself. I'll try to further clarify this point. As Jesper Juul (2005) has indicated, fiction and rules can be regarded as two complementary aspects of computer games. In *Assassin's Creed*, limiting ourselves to the intradiegetic level, we have one fictional world and one set of constraints. I prefer to speak of constraints/affordances rather than rules because I assume that *even narrative* belongs to this domain: after all, not everything can happen in computer games, but what *does* happen has been previously built into them, not unlike rules. However, my point is that we are never given the game world as separate from its affordances: it is within a blend of rules and story (on the one hand) and fiction (on the other) that meaning emerges. In *Assassin's Creed*, the in-game disturbance is precisely a material anchor, a concrete manifestation of that integration process. But how does meaning emerge? Not by itself. Aesthetic artifacts are pointless without a user, since meaning springs from the interaction between the artifact itself and its user (reader, spectator or player). Just as Fauconnier and Turner (2002) have argued that

language doesn't construct meaning by composition, but *prompts* for its imaginative reconstruction, we may conclude that aesthetic artifacts are props for meaning-making. Thus, in a way, playing a game is making sense of a blended world by enacting it. Juul came very close to this realization when he wrote that "The player navigates . . . two levels, playing video games in the half-real zone between the fiction and the rules" (2005: 202).

The upshot is that, theoretically, there must be an interface between the game-as-blend and the player. As we've seen, the blending process becomes more tangible at the level of what I've called the 2½-D interface: 2½-D elements are almost perfect blends of conceptual organization (the rules, the narrative) and perceptual givenness (the world as fiction). We should be careful to equate the "transcendental" interface made necessary by the interaction between player and game with the one shown on the screen, though. In a similar fashion, Iser warned not to confuse the blankness of the text with the actual blank space between words: those circumscribed blanks are just an actualization of the text's blankness, which is an "abstract, somewhat idealized way in order to explain the pivot on which the interaction between text and reader turns" (1978: 202). As *Dead Space* demonstrates, even a game in which the interface is seamlessly integrated into the fiction necessitates a space where the blend is made to overlap with the player. This is what I call a hermeneutic space. Our relation to it is not one of immersion, it is one of *enactment*: it results from our structural coupling with an environment.¹⁰ We unfold the meaning of an aesthetic artifact by enacting a blend between the world it projects and the constraints it places on the world itself.

What We've Found Out

I will now try to recap the main points I've made in these pages, pursuing a top-down strategy, from my overall conception of computer games to our immediate concern, ceptual blending. To start with, I've tried to draw a preparatory sketch of an "enactivist aesthetics"; this approach seems to be very appropriate to the game medium but could, at least in principle, be extended to other, more traditional media. I take "enactivism" in Varela, Thompson and Rosch's (1991) sense: to go past the typically Western dichotomy of object and subject (and the corresponding philosophical positions, "objectivism" and "subjectivism"), the authors have drawn on the Buddhist doctrine of the "codependent arising" of both. World and consciousness are given together, as the result of an interaction (an enaction). We are structurally coupled with our environment in a way that co-determines both our sensorimotor capacities and the affordances of the environment. Moreover, we enact "from a world of randomness a domain of distinctions . . . that has relevance for the structure of the system" (Varela, Thompson, and Rosch 1991: 155), and this brings into play a further dimension: meaning and interpretation. Far from being an exclusively human property, meaning can be found anywhere in the animal world as a way of meaningfully coping with an environment; humans only bring it to a higher level of sophistication by exploiting their conceptual and linguistic skills.

Following Johnson (2007), I see a structural resemblance between the enactive view of cognition and the way we make sense of aesthetic artifacts (including computer games).

¹⁰ There is obviously no space to discuss this point in detail, but I believe that Iser's conception of our engagement with texts (and, more in general, with aesthetic artifacts) as "entanglement" (1978: 126-129) and not immersion could cast doubt on Marie-Laure Ryan's (2001) rigid distinction between two metaphors: text-as-world (which lays great emphasis on immersion) and text-as-game (which foregrounds interaction). Computer games show that when we are engaged with a game we enact a blend that integrates an environment (a world) and its affordances (the rules of the game, its narrative).

These artifacts, I claim, construct worlds and offer us opportunities of interaction with them in much the same way as our world does; what changes, however, is that they are necessarily limited in extent (they are bounded by rules, stories, frames, and so on), so that our interaction with them appears to be particularly intense and self-conscious. Thus, on the one hand we have the quasi-ontological “feel” of aesthetic artifacts; on the other, the constraints put on them by their creators. But this is an *a posteriori* distinction, if not an illusory one altogether, since we are never given a fiction without some constraints.

I believe that Jesper Juul (and most computer game scholars) could be wrong in denying all games a fictional, world-creating element.¹¹ In the first place, we shouldn’t merge the abstract vs. representational dichotomy into the non-fictional vs. fictional one. Nor should we commit a mimetic fallacy. I can only touch on this point here, but let me say that not all worlds look like ours. Our world looks the way it does because we have certain sensorimotor capacities, but it must appear vastly different to a fly. A computer game like *Tetris* (1984) *does* construct a world, because it occurs in space and time (Salen and Zimmerman’s (2004) “magic circle”). This world may be two-dimensional and very sketchy, unlike the 3-D worlds of the latest computer games, but I don’t see any convincing reason why we shouldn’t regard it as one. It is true that some games place a stronger emphasis on the rules, others on the world, but this doesn’t mean that some games construct no world at all. Conceiving the worlds of computer games as environments in which games take place could possibly defuse the skepticism surrounding the *worldness* (a term I use interchangeably with “fictionality”) of all games.

It is at this point that “blending” comes in. My contention is that computer games construct blended worlds, i.e. they integrate the rules and, sometimes, a narrative¹² into an environment by turning them into constraints/affordances. Hence my initial discussion of the ready-to-handedness of fictional objects within computer games. This ready-to-handedness exists because games prompt what I have called, taking liberties with Fauconnier and Turner’s (2002) book, “ceptual blending”: they build on our tendency to blend perceived objects with the concepts of the actions that can be performed with them. In other words, the key to ceptual blending is the “virtuality” of both our perceptions and concepts: on the one hand, as Noë (2004) has shown, we have no need to store all the perceptual details of the world because they are already in the world, ready to be retrieved by simple eye and body movements. The world is virtually present to us, but present nonetheless. On the other hand, concepts have been characterized by Gallese (2005; see also Gallese and Lakoff 2005) as “embodied simulations” or “simulated potential actions” (where “potential” could be easily replaced with “virtual”). Computer games build on this nexus of simulation and virtuality.

There’s another blending at work, however. We wouldn’t be able to enact the blended worlds of computer games if *we* didn’t interact with them somehow. Beside the world of the work of art, writes Kendall Walton (1990), there is also the world of the game we play with it. My impression is that this “world of the game we play with games” is *also* blended. We tend to fuse the game in itself with the meaning we attribute to it. At another level, we tend to identify with our “avatar.” This is because playing is—fundamentally—meaning-making, and we see the meaning *in* the game while in fact it only exists in a blend. Moreover, some games use specialized interface elements to signal and “anchor” the integration between the blended

¹¹ Of course, not everyone will agree that “fictional” means “world-creating.” Yet this view is implicit in most definitions of fiction inspired by possible worlds theory. Ruth Ronen (1994: 82-83), for instance, writes that “the fictional . . . does not refer to an inner structure but to a type of relation: a relation maintained between what is contained within the literary text and what lies beyond its boundaries.” What lies beyond its boundaries is, precisely, what I call a world.

¹² Whereas I see *all* games as world-creating, I believe that narrative is a dimension games can do without.

world of the game and its meaning. This is what I've labeled the 2½-D interface, which includes all the elements that are present in the game world without belonging to it. To provide one last, straightforward example, *Bioshock* (2K Boston/2K Australia, 2007) has an option for highlighting usable objects, as shown in Figure 7.



Figure 7

Screenshot from *Bioshock* (2K Boston/2K Australia, 2007)

The wheel clearly stands out from its surroundings, but where does the highlight come from? Although it marks off objects in the game world, we immediately understand it as belonging to another dimension. And this is only in part because usable objects don't glow in the real world; rather, it is because we realize that the highlight is a material anchor for two recursive blends: between the rules of the game and the game world at one level, and between the meaning prompted by the game and the character's interaction with his environment at another. In other words: the rules of the game stipulate that we can pick up the wheel because it enables us to make sense of the game by enhancing the protagonist's structural coupling with the game world. A big accomplishment for a small gameplay option.

Eventually, I've tried to show that games establish a hermeneutic space where the player can interact with the blended world of the game, and vice versa. But we shouldn't confuse the interface actually shown on the screen (be it 2-D, 2½-D, or even 3-D) with this space, which could be described as the backdrop on which the game (as meaning-making activity) is projected. The real interface is just an aspect of the "transcendental" interface players navigate (to quote Juul again) in their interaction with the game. At the end of the day, it is only through this interface that we can enact games and—possibly—all aesthetic artifacts. And by insisting on the concept of "enaction" I wish to stress, again, the structural (as opposed to mimetic) resemblance between our experience of the real world and our experience of the blended worlds projected by aesthetic artifacts.

Games

ASSASSIN'S CREED. Ubisoft Montreal, PC, 2007.
BIOSHOCK. 2K Boston/2K Australia, PC, 2007.
CALL TO DUTY: WORLD AT WAR. Treyarch, PC, 2008.
DEAD SPACE. EA Redwood Shores, PC, 2008.
FALLOUT 3. Bethesda Softworks, PC, 2008.
GRAND THEFT AUTO 4. Rockstar North, PC, 2008.
MAX PAYNE 2. Remedy Entertainment, PC, 2003.
WORLD OF WARCRAFT. Blizzard Entertainment, PC, 2004.

References

- Aarseth, E. (2005). Doors and Perception: Fiction vs Simulation in Games. In *Proceedings of the 6th Digital Arts and Culture Conference*. Copenhagen. http://www.luisfilipeteixeira.com/fileManager/file/fiction_Aarseth_jan2006.pdf (accessed August 6th, 2009).
- Brey, P. (2005). The Epistemology and Ontology of Human–Computer Interaction. *Minds and Machines* 15, no. 3-4: 383–398.
- Brooks, R. A. (1991). Intelligence Without Reason. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, 569-595. Sydney. <http://ijcai.org/Past%20Proceedings/IJCAI-91-VOL1/PDF/089.pdf> (accessed August 6th, 2009).
- Dewey, J. (2005). *Art as Experience*. New York: Perigee Books.
- Dourish, P. (2001). *Where the Action Is: The Foundations of Embodied Interaction*. Cambridge, Mass.: MIT Press.
- Fauconnier, G., and Turner, M. (2002). *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.
- Ferri, G. (2007). Narrating Machines and Interactive Matrices: A Semiotic Common Ground for Game Studies. In *Situated Play. Proceedings of the DiGRA 2007 Conference*. Tokyo. <http://www.digra.org/dl/db/07311.02554.pdf> (accessed August 6th, 2009).
- Frasca, G. (2003). Simulation versus Narrative: Introduction to Ludology. In *The Video Game Theory Reader*, ed. Mark J. P. Wolf and Bernard Perron, 221-235. New York: Routledge.
- Freedberg, D., and Gallese, V. (2007). Motion, Emotion and Empathy in Esthetic Experience. *Trends in Cognitive Sciences* 11, no. 5: 197-202.
- Gallese, V. (2005). Embodied Simulation: From Neurons to Phenomenal Experience. *Phenomenology and the Cognitive Sciences* 4, no. 1: 23-48.
- Gallese, V., and Lakoff, G. (2005). The Brain's Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge. *Cognitive Neuropsychology*, no. 22: 455-479.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Heidegger, M. (1978). *Being and Time*. Trans. John Macquarrie and Edward Robinson. Oxford: Wiley-Blackwell.
- Herman, D. (2002). *Story Logic: Problems and Possibilities of Narrative*. Lincoln and London: University of Nebraska Press.
- Imaz, M., and Benyon, D. (2007). *Designing with Blends: Conceptual Foundations of Human-Computer Interaction and Software Engineering*. Cambridge, Mass.: MIT Press.

- Iser, W. (1978). *The Act of Reading: A Theory of Aesthetic Response*. Baltimore and London: Johns Hopkins University Press.
- Johnson, M. (2007). *The Meaning of the Body: Aesthetics of Human Understanding*. Chicago and London: University of Chicago Press.
- Juul, J. (2005). *Half-Real: Video Games between Real Rules and Fictional Worlds*. Cambridge, Mass.: MIT Press.
- . (2007). A Certain Level of Abstraction. In *Situated Play. Proceedings of the DiGRA 2007 Conference*. Tokyo. <http://www.digra.org/dl/db/07312.29390.pdf> (accessed August 6th, 2009).
- Lakoff, G., and Johnson, M. (1980). *Metaphors We Live By*. Chicago and London: University of Chicago Press.
- Manovich, L. (2001). *The Language of New Media*. Cambridge, Mass.: MIT Press.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman.
- Noë, A. (2004). *Action in Perception*. Cambridge, Mass.: MIT Press.
- Rohrer, T. (1995). Metaphors We Compute By: Bringing Magic into Interface Design. <http://www.uoregon.edu/~uophil/metaphor/gui4web.htm> (accessed August 6th, 2009).
- Ronen, R. (1994). *Possible Worlds in Literary Theory*. Cambridge: Cambridge University Press.
- Ryan, M.-L. (1991). *Possible Worlds, Artificial Intelligence and Narrative Theory*. Bloomington and Indianapolis: Indiana University Press.
- . (2001). *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media*. Baltimore and London: Johns Hopkins University Press.
- Salen, K., and Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. Cambridge, Mass.: MIT Press.
- Talmy, L. (1996). Fictive Motion in Language and "Ception." In *Language and Space*, ed. Paul Bloom, Mary A. Peterson, Lynn Nadel, and Merrill F. Garrett, 211-276. Cambridge, Mass.: MIT Press.
- Taylor, L. (2003). When Seams Fall Apart: Video Game Space and the Player. *Game Studies* 3, no. 2 (December). <http://gamestudies.org/0302/taylor/> (accessed August 6th, 2009).
- Turner, M. (1996). *The Literary Mind: The Origins of Thought and Language*. New York and Oxford: Oxford University Press.
- Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, Mass.: MIT Press.
- Walton, K. (1990). *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, Mass., and London: Harvard University Press.
- Wollheim, R. (1980). Seeing-as, Seeing-in, and Pictorial Representation. In *Art and Its Objects*, 205-226. Cambridge: Cambridge University Press.