INTERACTIVITY

action > outcome
four modes of interactivity
anatomy of a choice
internal event
external event
space of possibility

The word “interactivity” isn’t just about giving players choices; it pretty much completely defines the game medium.—Warren Spector; RE:PLAY: Game Design + Game Culture
Introducing Interactivity

Play implies interactivity: to play with a game, a toy, a person, an idea, is to interact with it. More specifically, playing a game means making choices within a game system designed to support actions and outcomes in meaningful ways. Every action results in a change affecting the overall system. This process of action and outcome comes about because players interact with the designed system of the game. Interaction takes place across all levels, from the formal interaction of the game's objects and pieces, to the social interaction of players, to the cultural interaction of the game with contexts beyond its space of play.

In games, it is the explicit interaction of the player that allows the game to advance. From the interactivity of choosing a path to selecting a target for destruction to collecting magic stars, the player has agency to initiate and perform a whole range of explicit actions. In some sense, it is these moments of explicit action that define the tone and texture of a specific game experience. To understand this particular quality of games—the element of interaction—we must more completely grasp the slippery terms "interactive," "interaction," and "interactivity."

Defining Interactivity

Perhaps even more than "design" and "systems," debates over the term "interactivity" have run rampant. Interactivity is one of those words that can mean everything and nothing at once. If everything can indeed be considered interactive, then the concept loses its ability to help us solve design problems. In corralling this runaway word, our aim is to try and understand it in its most general sense, but also to identify those very particular aspects of interactivity that are relevant to games. To this end, we look at several definitions of interactivity. We begin with a general question: What is "interaction?" Here are some basic dictionary definitions:

- **interact:** to act on each other; act reciprocally;
- **interactive:** reciprocally active; acting upon or influencing each other; allowing a two-way flow of information between a device and a user, responding to the user's input.1

In the most general terms, interactivity simply describes an active relationship between two things. For our purposes, however, we require a slightly more rigorous definition, one that takes into account the particular nature of games. Instead of asking about interactivity in the abstract, what does it mean to say that something is "interactive?" More specifically, how does interactivity emerge from within a system?

Communications theorist Stephen W. Littlejohn defines interactivity this way: "Part and parcel of a system is the notion of 'relationship'... Interactional systems then, shall be two or more communicants in the process of, or at the level of, defining the nature of their relationship."2 In other words, something is interactive when there is a reciprocal relationship of some kind between two elements in a system. Conversations, databases, games, and social relationships are all interactive in this sense. Furthermore, relationships between elements in a system are defined through interaction.

Following this definition, digital media theorist and entrepreneur Brenda Laurel brings the concept of representation to an understanding of the term: "...something is interactive when people can participate as agents within a representational context. (An agent is 'one who initiates actions.')"3 Laurel's model emphasizes the interpretive component of interactive experiences, framing an interactive system as a representational space.

In an alternative definition of interactivity, theorist Andy Cameron builds on this interpretive dimension by stressing the idea of direct intervention. In his essay "Dissimulations," Cameron writes that

Interactivity means the ability to intervene in a meaningful way within the representation itself, not to read it differently. Thus interac-
itivity in music would mean the ability to change the sound; interactivity in painting to change colors, or make marks; interactivity in film... the ability to change the way the movie comes out.”

Cameron suggests a connection between interactivity and explicit action, a key feature of games and meaningful play. In some sense, it is these moments of explicit action that define the tone and texture of a specific game experience.

A final definition comes from game designer Chris Crawford, who metaphorically defines interactivity in terms of a conversation: “Interactivity: a cyclical process in which two actors alternately listen, think, and speak. The quality of interaction depends on the quality of each of the subtasks (listening, thinking, and speaking).”

While his definition hearkens back to Littlejohn’s relational model, Crawford’s definition stresses the iterative quality of interactivity. He uses the following example for emphasis:

A conversation, in its simplest form, starts out with two people, Joe and Fred. Joe says something to Fred. At this point, the ball is in Fred’s court. He performs three steps in order to hold up his end of the conversation:

Step One: Fred listens to what Joe has to say. He expends the energy to pay attention to Joe’s words. He gathers in all of Joe’s words and assembles them into a coherent whole. This requires an active effort on Fred’s part.

Step Two: Fred thinks about what Joe said. He considers, contemplates, and cogitates. The wheels turn in his mind as Fred develops his response to Joe’s statement.

Step Three: Fred expresses his response back to Joe. He forms his thoughts into words and speaks them.

Now the tables are turned; the ball is in Joe’s court. Joe must listen to what Fred says; Joe must think about it and develop a reaction; then he must express his reaction to Fred. This process cycles back and forth. Thus, a conversation is an iterative process in which each participant in turn listens, thinks, and speaks.

Each of these definitions provides its own critical way of understanding interactivity: it takes place within a system, it is relational, it allows for direct intervention within a representational context, and it is iterative. Yet none of the definitions describes how and where interactivity can take place, and none of them address the relationship between structure and context, two key elements in the construction of meaning. These questions of “how,” “where,” and “by whom” are critical to anyone faced with the challenge of designing interactivity.

In other words, none of these definitions resolve the question of whether or not all media, or even all experiences, are interactive. If interactivity is really so ubiquitous, can it possibly be a useful term for understanding games?

A Multivalent Model of Interactivity

Each of the previous definitions foreground a particular aspect of interaction; in our view, they are all are useful ways of defining interactivity. Rather than try and distill them into a composite definition, we have elected instead to offer a model of interactivity that accommodates each of these definitions. The model presents four modes of interaction, or four different levels of engagement, that a person might have with an interactive system. Most “interactive” activities incorporate some or all of them simultaneously.

Mode 1: Cognitive interactivity; or interpretive participation

This is the psychological, emotional, and intellectual participation between a person and a system. Example: the complex imaginative interaction between a single player and a graphic adventure game.

Mode 2: Functional interactivity; or utilitarian participation

Included here: functional, structural interactions with the material components of the system (whether real or virtual). For example, that graphic adventure you played: how was the interface? How “sticky” were the buttons? What was the response time? How legible was the text on your high-resolution monitor? All of these elements are part of the total experience of interaction.
Mode 3: Explicit interactivity; or participation with designed choices and procedures

This is "interaction" in the obvious sense of the word: overt participation like clicking the non-linear links of a hypertext novel, following the rules of a board game, rearranging the clothing on a set of paper dolls, using the joystick to maneuver Ms. Pac-Man. Included here: choices, random events, dynamic simulations, and other procedures programmed into the interactive experience.

Mode 4: Beyond-the-object-interactivity; or participation within the culture of the object

This is interaction outside the experience of a single designed system. The clearest examples come from fan culture, in which participants co-construct communal realities, using designed systems as the raw material. Will Superman come back to life? Does Kirk love Spock?

Some of these modes occur universally in human experience, such as Mode 1, cognitive interactivity. Yet not all of them do. For our purposes, Mode 3, explicit interactivity, comes closest to defining what we mean when we say that games are "interactive." An experience becomes truly interactive in the sense of Cameron's "direct intervention" only when the participant makes choices that have been designed into the actual structure of the experience.

The rest of this chapter focuses primarily on explicit interactivity and how game designers can create the kinds of choices that result in meaningful play. However, even though we will be focusing on Mode 3, it is important to remember that the other three modes of interactivity are also present as players make explicit choices. For example, choosing whether to fold or not in Poker represents a moment of explicit interactivity. But at the same time, the material quality and size of the cards affect the functional interactivity; the fanciful images on the face cards might engender cognitive interactivity; and notions about what it means to be a suave card shark—or perhaps resentment at being trounced at the Poker table last week—represent forms of cultural participation that lie outside the bounds of the particular game being played.

Interaction, even the explicit interaction of a seemingly straightforward game choice, is never as simple as it appears at first glance. But before we dissect the components of explicit interactive choices, let's pause to consider the role of design itself in creating interactivity.

But Is it "Designed" Interaction?

Interaction comes in many forms. But for the purposes of designing interactivity, it is important to be able to recognize what forms of interactivity designers create. As an example, compare the following two actions: someone dropping an apple on the ground and someone rolling dice on a craps table. Although both are examples of interaction proper, only the second act, the rolling of the dice, is a form of designed interaction.

What about this action has been designed? First, the dice, unlike the apple, are part of a system (a game) in which the interaction between the player and the dice is made meaningful by a set of rules describing their relationship. This relationship, as defined by the rules of Craps, describes the connection between action and outcome—for example, "When the dice are rolled a player counts the number of dots appearing on the face-up sides of the dice." Even this extremely simple rule demonstrates how the act of rolling has meaning within the designed interactive system of the game. Secondly, the interaction is situated within a specific context: a game. Remember that meaningful play is tied not only to the concept of player action and system outcome, but also to a particular context in which the action occurs.

The description of "someone dropping an apple on the ground," on the other hand, does not contain a designed structure or context. What conditions would have to be present to evolve this simple interaction into a designed interaction? The
dropping of the apple does meet baseline criteria for interaction: there is a reciprocal relationship between the elements of the system (such as the person’s hand, the apple, and the ground). But is it a designed interaction? Is the interactivity situated within a specific context? Do we have any ideas about what dropping an apple might “mean” as a form of interaction between a person and an apple? Do we have a sense of the connection between action and outcome?

No. All we know is that an apple has been dropped. What is missing from this description is an explicitly stated context within which the dropping of the apple occurs. If we change the scenario a little by adding a second player and asking the two participants to toss the apple back and forth, we move toward a situation of designed interaction. If we ask the two apple-tossers to count the number of times in a row they caught the apple before dropping it, we add an even fuller context for the interaction. The simple addition of a rule designating that the players quantify their interaction locates the single act of toss-catch within an overall system. Each element in the system is assigned a meaning: the toss, the catch, and the dropped toss. Even in the simplest of contexts, design creates meaning.

**Interaction and Choice**

The careful crafting of player experience through a system of interaction is critical to the design of meaningful play. Yet, just what makes an interactive experience “meaningful”? We have argued that in order to create instances of meaningful play, experience has to incorporate not just explicit interactivity, but meaningful choice. When a player makes a choice in a game, the system responds in some way. The relationship between the player’s choice and the system’s response is one way to characterize the depth and quality of interaction. Such a perspective on interactivity supports the descriptive definition of meaningful play presented in chapter 3.

In considering the way that choices are embedded in game activity, we look at the design of choice on two levels: micro and macro. The micro level represents the small, moment-to-moment choices a player is confronted with during a game. The macro level of choice represents the way these micro-choices join together like a chain to form a larger trajectory of experience. For example, this distinction marks the difference between tactics and strategy in a game such as Go. The tactics of Go concern the tooth-and-nail battles for individual sectors of the board, as individual pieces and small groups expand across territory, bumping up against each other in conflict and capture. The strategy of the game is the larger picture, the overall shape of the board that will ultimately determine the winner. The elegance of the design of Go lies in its ability to effortlessly link the micro and the macro, so that every move a player makes works simultaneously on both levels. Micro-interaction and macro-interaction are usually intertwined and there are, of course, numerous shades of gray in-between.

Keep in mind that “choice” does not necessarily imply obvious or rational choice, as in the selection of an action from a menu. Choice can take many forms, from an intuitive physical action (such as the “twitch” firing of a Time Crisis pistol) to the random throw of a die. Following are a few more examples of designed choices in games.

The choice of whether or not to take a hit in Blackjack. A Blackjack player always has a clear set of choices: the micro-choice of taking or not taking a hit will have the eventual outcome of a win or a loss against the house. On the macro-level, each round affects the total amount of money the player gains or loses over the course of the game. Playing each hand separately, according to its probability of beating the house is like tactics in Go. Counting cards, which links all of a players’ hands between rounds, is a more long-term, strategic kind of choice-making.

The choice of what to type into the flashing cursor of a text adventure. This is a more open-ended choice context than the simple hit or pass of Blackjack. The micro-choice of typing in a command gives the player feedback about
how the player moves through or changes the world. The choice to type the words “Move North” takes the player to another location in the game where different actions are possible—perhaps actions that will eventually solve the multi-part puzzles that exist on the macro-level of game play. Even when a player tries to take an action that the program cannot parse (such as typing “grab rock” instead of “get rock”), it is meaningful: the outcome of bumping up against the limits of the program’s parsing ability serves to further delineate the boundaries of play.

The choice of what play to call in a Football game. This moment of game-choice is often produced collaboratively among a coaching staff, a quarterback, and the rest of the offensive players. There are a large number of possible plays to call, each with variations, and the choice is always made against the backdrop of the larger game: the score, the clock, the field position, the down, the strengths and weaknesses of both teams. The most macro-level of choices address the long-term movement of the ball across the field and the two teams’ overall scores. The most micro-level of choices occur once the play is called and the ball is hiked: every offensive player has the moment-to-moment challenge of executing the play as the defensive team does its best to put a stop to it.

As these examples demonstrate, choice-making is a complex, multi-layered process. There is a smooth transition between the micro- and macro-levels of choice-making, which play out in an integrated way for the player. When the outcome of every action is discernable and integrated, choice-making leads to meaningful play. Game designer Doug Church, in his influential online essay “Formal Abstract Design Tools,” outlines the way that these levels of choice transition into a complete game experience.

In a fighting game, every controller action is completely consistent and visually represented by the character on-screen. In Tekken, when Eddy Gordo does a cartwheel kick, you know what you’re going to get. As the player learns moves, this consistency allows planning—intention—and the reliability of the world’s reactions makes for perceived consequence. If I watch someone play, I can see how and why he or she is better than I am, but all players begin the game on equal footing.

As Church points out, the macro-levels of choice-making include not only what to do over the course of a game, but also whether or not you want to play a game, and against whom. If you are beaten in a fighting game that doesn’t contain clear and meaningful play, you will never know why you lost and you will most likely not play again. On the other hand, if you know why your opponent is better than you are, your loss is meaningful, as it helps you assess your own abilities, gives you ideas for improvement, and spurs on your overall interaction with the game.

Choice Molecules
[The designers of Spacewar!, the first computer game] identified action as the key ingredient and conceived Spacewar! as a game that could provide a good balance between thinking and doing for its players. They regarded the computer as a machine naturally suited for representing things that you could see, control, and play with. Its interesting potential lay not in its ability to perform calculations but in its capacity to represent action in which humans could participate.—Brenda Laurel, Computers as Theater

The capacity for games to “represent action in which players participate“ forms the basis of our concept of "choice." If we consider that every choice has an outcome, then it follows that this action > outcome unit is the vehicle through which meaning in a game emerges. Although games can generate meaning in
many ways (such as through image, text, sound, etc.), to understand the interactive nature of meaningful play, we focus on the kinds of meaning that grow from player interaction. At the heart of interactive meaning is the action > outcome unit, the molecule out of which larger interactive structures are built.

In order to examine this concept more closely we look at the classic arcade game Asteroids, a direct descendent of Spacewar!. In Asteroids, a player uses buttons to maneuver a tiny spaceship on the screen, avoiding moving asteroids and UFOs and destroying them by shooting projectiles. The action > outcome interactive units of Asteroids are manipulated through a series of five player commands, each one of them a button on the arcade game’s control panel: rotate left, rotate right, thrust, fire, and hyperspace. Within the scope of an individual game, possible player actions map to the five buttons:

- Press rotate right button: spaceship rotates right
- Press rotate left button: spaceship rotates left
- Press thrust button: spaceship accelerates in the direction it is facing
- Press fire button: spaceship fires projectile (up to four on the screen at a time)
- Press hyperspace button: spaceship disappears and reappears in a different location (and occasionally perishes as a result)

Action on the screen is affected through the subtle (and not so subtle!) orchestration of these five controls. As the game progresses, each new moment of choice is a response to the situation onscreen, which is the result of a previous string of action > outcome units. The seamless flow that emerges is one of the reasons why Asteroids is so much fun to play. Rarely are players aware of the hundreds of choices they make each minute as they dodge space rocks and do battle with enemy ships—they perceive only their excitement and participation inside the game.

Anatomy of a Choice

Although the concept of choice may appear basic upon first glance, the way that a choice is actually constructed is surprisingly complex. To dissect our action > outcome molecule, we need to ask the following five questions. Together, they outline the anatomy of a choice:

1. What happened before the player was given the choice? What is the current state of the pieces on a game board, for example, or the level of a player’s health? What set of moves just finished playing out? What is the game status of the other players? This question relates to the both the micro and macro events of a game, and addresses the context in which a choice is made.

2. How is the possibility of choice conveyed to the player? On a game board, the presence of empty squares or a “draw pile” might indicate the possibility of choice, whereas choices in a digital game are often conveyed through the game’s controls. In Asteroids, for example, the five buttons on the control panel communicate the opportunity for choice-making to the player.

3. How did the player make the choice? Did the player make a choice by playing a card, pressing a button, moving a mouse, running in the opposite direction, or passing on a turn? The mechanisms a player uses to make a choice vary greatly, but all are forms through which players are given the opportunity to take action.
4. What is the result of the choice? How will it affect future choices? A player taking action within a system will affect the relationships present in that system. This element of the anatomy of a choice speaks to the outcome of a player action, identifying how a single choice impacts larger events within the game world. The outcome of taking a "hit" in Blackjack impacts whether or not the player wants to take another hit, as well as the outcome of the game.

5. How is the result of the choice conveyed to the player? The means by which the results of a choice are represented to a player can assume many guises, and forms of representation are often related to the materiality of the game itself. In a game of Twister, for example, the physical positioning of bodies in space conveys the results of choices; in Missile Command, the result of the choice to "fire" is conveyed by a slowly moving line of pixels, ending in an explosion; in Mousetrap, the mechanical workings (or non-workings) of the mousetrap convey the results of moving a mouse into the trap space. Note that step 5 leads seamlessly back to step 1, because the result of the choice provides the context for the next choice.

These are the five stages of a choice, the five events that transpire every time an action and outcome occur in a game. Each stage is an event that occurs internal or external to the game. Internal events are related to the systemic processing of the choice; external events are related to the representation of the choice to the player. These two categories make a distinction between the moment of action as handled by the internal game state and the manifestation of that action to the player.

The idea that a game can have an internal event represented externally implies that games are systems that store information. Jesper Juul, in a lecture titled "Play Time, Event Time, Themability," describes this idea by thinking of a game as a state machine:

A game is actually what computer science describes as a state machine. It is a system that can be in different states. It contains input and output functions, as well as definitions of what state and what input will lead to what following state. When you play a game, you are interacting with the state machine that is the game. In a board game, this state is stored in the position of the pieces on the board; in computer games the state is stored as variables, and then represented on the screen.8

In Juul's example of a board game, the "internal" state of the game is immediately evident to the players in the way that the pieces are arranged on the board. In the case of a computer game, as Juul points out, the internal variables have to be translated into a representation for the player. The distinction between internal and external events helps us to identify and distinguish the components of a choice. Within the action > outcome molecule, stages 1, 3, and 4 are internal events, and

Anatomy of a choice
1. What happened before the player was given the choice? (internal event)

2. How is the possibility of choice conveyed to the player? (external event)

3. How did the player make the choice? (internal event)

4. What is the result of the choice? How will it affect future choices? (internal event)

5. How is the result of the choice conveyed to the player? (external event)
Figure 1

<table>
<thead>
<tr>
<th>Anatomy of a Choice</th>
<th>Asteroids</th>
<th>Chess</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What happened before the player was given the choice? (internal event)</td>
<td>Represented by the current positions and trajectories of the game elements.</td>
<td>Represented by the current state of the pieces on the board.</td>
</tr>
<tr>
<td>2. How is the possibility of choice conveyed to the player? (external event)</td>
<td>The possible actions are conveyed through the persistent button controls as well as the state of the screen, as it displays the relationships of the game elements.</td>
<td>The possible actions are conveyed through the arrangement of pieces on the board, including the empty squares where they can move.</td>
</tr>
<tr>
<td>3. How did the player make the choice? (internal event)</td>
<td>The player makes a choice by pressing one of the 5 buttons.</td>
<td>The players make a choice by moving a piece.</td>
</tr>
<tr>
<td>4. What is the result of the choice? How will it affect future choices? (internal event)</td>
<td>Each button press affects the system in a different way, such as the position or orientation of the player's ship.</td>
<td>Each move affects the overall system, such as capturing a piece or shifting the strategic possibilities of the game.</td>
</tr>
<tr>
<td>5. How is the result of the choice conveyed to the player? (external event)</td>
<td>The result of the choice is then represented to player via screen graphics and audio.</td>
<td>The result of the choice is then represented to the player via the new arrangement of pieces on the board.</td>
</tr>
</tbody>
</table>

stages 2 and 5 are external events. These two layers of events form the framework within which the anatomy of a choice must be considered. To see how this all fits together, let us take an even closer look at the way choice is constructed in two of our example games, Asteroids and Chess. (Figure 1)

Although all five stages of the action > outcome choice event occurred in both games, there are some significant differences. In Asteroids, the available choices and the taking of an action both involve static physical controls. In Chess, the pieces on the board serve this function, even as they convey the current state of the game. The internal and external states of Chess are identical, but in Asteroids, what appears on the screen is only an outward extension of the internal state of the software. The "anatomy of a choice" structure occurs in every game, although each game will manifest choice in its own way.

This way of understanding choice in a game can be extremely useful in diagnosing game design problems. If your game is failing to deliver meaningful play, it is probably because there is a breakdown somewhere in the action > outcome chain. Here is a sample list of common "failure states" that can often be found in games and the way that they relate to the stages of a choice.

- **Feeling as if decisions are arbitrary.** If you need to play a card from your hand and it always feels like it doesn't matter which card you select, the game probably suffers in stage 4, the effect of the player's choice on the system of the game. The solution is to make sure that player actions have meaningful outcomes in the internal system of the game.
- *Not knowing what to do next.* This can be a common problem in large digital adventure games, where it is not clear how a player can take action to advance the game. The problem is in stage 2, representing choices to the player. These kinds of problems are often solved with additional information display, such as highlights on a map, or an arrow or indicator that helps direct the player.

- *Losing a game without knowing why.* You think that you're about to reach the top of the mountain, when your character dies unexpectedly from overexposure. This frustrating experience can come about because a player has not sufficiently been informed about the current state of the game. The problem might be in stage 5, where the new state of the game resulting from a choice is not represented clearly enough to the player.

- *Not knowing if an action had an outcome.* Although this sounds like something that would never happen, there are many examples of experimental interactivity (such as a gallery-based game with motion sensor inputs) in which the player never receives clear feedback on whether or not an action was taken. In this case, there is a breakdown at stages 3 and 4, when a player is taking an action and receiving feedback on the results.

These examples represent only a small sampling of the kinds of problems that a game's design can have. The anatomy of a choice is not a universal tool for fixing problems, but it can be especially useful in cases where the game is breaking down because of a glitch in the player's choice-making process.

**Space of Possibility**

We conclude this chapter with an excerpt from David Sudnow's book, *Pilgrim in a Microworld*, a wonderfully detailed personal account of one man's very real obsession with the video game Breakout. Sudnow brings readers into the space of designed interactivity through detailed descriptions of what he experienced—physically, psychologically, emotionally—as he played. There are remarkably few documents that offer such a sensitive and insightful analysis of designed interaction.

I'd catch myself turning my chair into a more en face position vis-à-vis the TV. An obvious delusion. Maybe I could rest one elbow on the set to help feel the angle of my look and deepen a sense for the scale of things. See it from this side and that, see the invisible backside of things through an imaginary bodily tour of the object. Nonsense, if only I could feel the impact of the ball on the paddle, that would certainly help, would give me a tactile marker, stamping the gesture's places into a palpable little signature so I'd feel each destination being achieved and not just witness the consequences of a correct shot. Nonsense.

Non-sense, just your eyes way up top, to be somehow fixed on things in ways that can't feel them fixing, then this silent smooth little plastic knob down there, neither near nor far away but in an untouchable world without dimensions. And in between all three nodes of the interface there's nothing but a theory of electricity. So fluid, to have to write your signature with precise consistency in size within the strict bounds of a two and three-sevenths of an inch of space, say, while the pen somehow never makes contact with the paper. There's nothing much to hold on to, not enough heft in this knob so your hands can feel the extent of very minor movements, no depth to things you can use to anchor a sense of your own solidity.

As game designers, what can we glean from Sudnow's observations? His analysis suggests that there is a wealth of information to be gained about a game's interactivity by looking at it from the player's point of view. One of our disappointments with current writing on games and interactivity is that much analysis occurs not from the point of view of the player, but from the point of view of an outside spectator. This style of over-the-shoulder journalism fails to recognize that interac-
tivity is something to be experienced, rather than observed. In writing a player-centric account of his encounter with the game, Sudnow calls attention to key concepts for designed interaction. Concepts such as directed choice, player control, amplification of input, system representation, and direct, visible feedback emerge in his poetic meditation on perception, attention, cognition, and the body.

Creating a game means designing a structure that will play out in complex and unpredictable ways, a space of possible action that players explore as they take part in your game. What possible actions might players take in the course of a game of Musical Chairs? They might push, shove, tickle, poke, or fight for their seat once the music stops and the mad scramble for chairs begins. The game designer must carefully craft a system of play in which these actions have meaning in support of the play of the game, and do not distract or interrupt its play.

But game designers do not directly design play. They only design the structures and contexts in which play takes place, indirectly shaping the actions of the players. We call the space of future action implied by a game design the space of possibility. It is the space of all possible actions that might take place in a game, the space of all possible meanings which can emerge from a game design. The concept of the space of possibility not only bridges the distance between the designed structure and the player experience, but it also combines the key concepts we have presented so far. The space of possibility is designed (it is a constructed space, a context), it generates meaning (it is the space of all possible meanings), it is a system (it is a space implied by the way elements of the system can relate to each other), and it is interactive (it is through the interactive functioning of the system that the space is navigated and explored).

The space of possibility springs forth out of the rules and structures created by the game designer. The space of possibility is the field of play where your players will explore and cavort, compete and cooperate, as they travel through the experience of playing your game. But like David Sudnow who wishes he could reach out and touch the electronic blip of his Breakout paddle, as a game designer you can never directly craft the possible space of your game. You only can indirectly construct the space of possibility, through the rules you design. Game design is an act of faith—in your rules, in your players, in your game itself. Will your game create meaningful play? You can never know for sure. But understanding key concepts like design, systems, and interactivity can help bring you closer to a meaningful outcome.

Further Reading

_Computers as Theater_, by Brenda Laurel

Although Laurel is not speaking about games directly, her discussion of a dramatic theory of human-computer activity has many connections to the interactivity of games. The most relevant discussions to game design focus on the mechanics of interaction and the way people interact with machine interfaces.

_Recommended:_

- Chapter 1: The Nature of the Beast
- Chapter 5: Design Principles for Human-Computer Activity

_The Design of Everyday Things_, by Donald Norman

Norman's book is a must read for any designer involved in the design of interactive systems. His approach has been formalized more recently within the catch-phrase "experience design," which places the user at the center of any designed activity. Although Norman is writing about everyday objects such as telephones and car doors, his observations have direct application to the design of games as interactive systems.

_Recommended:_

- Chapter 1: The Psychopathology of Everyday Things
- Chapter 2: The Psychology of Everyday Actions
- Chapter 3: Knowledge in the Head and in the World
“Designing Interactive Theme Park Rides: Lessons From Disney’s Battle for the Buccaneer Gold,” by Jesse Schell and Joe Shochet
In this design postmortem of one of Disney’s interactive theme park rides, Schell and Shochet discuss the reasons for the ride’s success. Their analysis is design-driven, and offers insights into the tools, techniques, and psychology used to create an effective and entertaining interactive experience. Available at <www.gamasutra.com>.

“Formal Abstract Design Tools,” by Doug Church
In making one of the most robust arguments for the development of a common vocabulary for games, Doug Church establishes a precedent for critical thinking within the emerging field of game design. “Formal Abstract Design Tools” is written from a game design perspective and explores concrete concepts of interactivity in the design of player experience. Available at <www.gamasutra.com>.

Pilgrim in the Microworld, by David Sudnow
This first-person account of one man’s genuine obsession with the Atari 2600 game Breakout offers a clear portrait of the aesthetics of interactive systems. Concepts related to the anatomy of a choice, discernability and integration of player action, pleasure, and core mechanics are discussed in terms of player experience, making it a valuable resource for those intent on understanding just what is happening from moment-to-moment during game play.

Recommended:
- Memory
- Interface
- Cathexis
- Eyeball
- Coin

The Art of Interactive Design: A Euphorious and Illuminating Guide to Building Successful Software, by Chris Crawford
The Art of Interactive Design is a non-technical book about the design of interactivity. Crawford uses his experience as a designer of games and interactive systems to discuss how interactivity works. For Crawford, interactivity is “a cyclic process in which two actors alternatively listen, think, and speak.” This conversational model of interaction is used throughout the text to good effect.

Recommended:
- Part I: Chapters 1–6

Notes
1. <dictionary.com>.
6. Ibid; p. 7.
- **Interactivity** is closely linked to the concepts of design, systems, and meaningful play. When a player interacts with the designed system of a game, meaningful play emerges.

- There are many valid definitions of interactivity. Cutting across all of them are **four modes of interactivity**:
  - Mode 1: Cognitive interactivity; or interpretive participation;
  - Mode 2: Functional interactivity; or utilitarian participation;
  - Mode 3: Explicit interactivity; or participation with designed choices and procedures;
  - Mode 4: Beyond-the-object-interactivity or cultural participation.

- These four modes are not distinct categories but are instead overlapping ways of understanding any moment of interactivity. They usually occur simultaneously in any experience of a designed system.

- Not all interaction is **designed interaction**. When an interaction is designed, it has an internal structure and a context that assign meaning to the actions taken.

- An interactive context presents participants with **choices**. Choices can be micro-choices of moment-to-moment interactivity or macro-choices, which concern the long-term progress of the game experience.

- The basic unit out of which interactive meaning is made is the **action > outcome** unit. These units are the molecules out of which interactive designers (including game designers) create larger structures of designed interaction.

- Within each action > outcome event is a series of five stages that help construct a **choice** in a game. These stages are expressed through the following questions:
  1. What happened before the player was given the choice?
  2. How is the possibility of choice conveyed to the player?
  3. How did the player make the choice?
  4. What is the result of the choice? How will it affect future choices?
  5. How is the result of the choice conveyed to the player?

- Each of these stages represents either an **internal event**, in which the system of the game processes and receives the choice, or an **external event**, in which the choice is represented to the player.

- The **space of possibility** of a game is the space of all possible actions and meanings that can emerge in the course of the game. This concept ties together meaning, design, systems, and interactivity.
It is true that every aspect of the role of dice may be suspect: the dice themselves, the form and texture of the surface, the person throwing them. If we push the analysis to its extreme, we may even wonder what chance has to do with it at all. Neither the course of the dice nor their rebounds rely on chance; they are governed by the strict determinism of rational mechanics. Billiards is based on the same principles, and it has never been considered a game of chance. So in the final analysis, chance lies in the clumsiness, the inexperience, or the naiveté of the thrower—or in the eye of the observer....

As for billiards, it can easily be transformed into a game of chance by simply tilting the table, outfitting it with studs that would cause the balls to rebound and swerve, and by placing the six pockets at the bottom of the table, or at other points, so that the ball would necessarily fall into one of them. Since we're not trying to favor skill, there would be a mechanical trigger and the ball would be shot up the slope by a spring that the player would pull with more or less force. This game of mechanical billiards is no less random than traditional dice.—Ivar Ekeland, The Broken Dice
Introducing Uncertainty

Imagine how incomplete you would feel if, before the game, you were already declared the winner. Imagine how purposeless the game would feel.—Bernard DeKoven, The Well-Played Game

Uncertainty is a central feature of every game. That’s right: every single game. As game designer and philosopher Bernard DeKoven points out, uncertainty about the outcome of a game is a necessary ingredient in giving a game a feeling of purpose. Uncertainty, in other words, is a key component of meaningful play.

In this chapter, we explore games as Systems of Uncertainty. Games express uncertainty on two levels: on a macro-level relating to the overall outcome of a game, and on a micro-level relating to specific operations of chance within the designed system. Although all games possess uncertainty on a macro-level, not all games formally possess elements of uncertainty on a micro-level. As we will see, a player’s experience of uncertainty is not always congruent with the actual amount of mathematical chance in a game. Exploring these relationships, linking macro- and micro-uncertainty to each other, and understanding how both of them impact the design of meaningful play, is our primary focus in this schema.

Does every game really possess uncertainty? The word uncertainty brings to mind ideas of chance and randomness. But a game does not have to have a die roll or random algorithm to contain an element of uncertainty. If you are playing a multiplayer session of Halo against players of roughly equivalent ability, the outcome of the game is uncertain, even though the game is a game of skill, not chance. When we say that uncertainty is a central feature of every game, we are echoing DeKoven in the quote above: it is crucial in a game that players don’t know exactly how it will play out. Think about it: if you knew who was going to win a game before it started, would you even bother to play? There is a reason why televised sports are almost always aired live: robbed of the drama of uncertain outcome, they fail to hold our interest.

One way to understand why games need uncertainty is that if the outcome of a game is predetermined, the experience cannot provide meaningful play. If a game has no uncertainty—if the outcome of the game is completely predetermined—then any choices a player makes are meaningless, because they do not impact the way that the game plays out. Meaningful play arises from meaningful choices. If a player’s choices have no meaning in the game, there really is no reason to play.

There is an intrinsic connection between uncertainty and meaningful play. Uncertainty is usually thought of as something that disempowers players by removing a sense of choice and agency, yet paradoxically, it is the uncertain outcome of a game that allows players to feel like their decisions have an impact on the game. Meaningful play, as we know, emerges from these kinds of decision-outcome relationships.

Throughout a game system, this larger notion of an uncertain outcome is linked to the micro-level of uncertainty within a game. The specific mechanisms of uncertainty that incorporate randomness and chance, whether through the spin of a Roulette wheel or the generation of a random number in a game program, are just as important as the larger feeling of uncertainty linked to a game’s outcome. From the interaction between these two levels, the meaningful play of uncertainty arises.

Certainty, Uncertainty, and Risk

The essence of the phenomenon of gambling is decision making. The act of making a decision consists of selecting one course of action, or strategy, from among the set of admissible strategies. A particular decision might indicate the card to be played, a horse to be backed, the fraction of a fortune to be hazarded over a given interval of play.... Decisions can be categorized according to the specific relationship between action and outcome.—Richard Epstein, The Theory of Gambling and Statistical Logic
In *The Theory of Gambling and Statistical Logic*, mathematician Richard Epstein investigates the mathematics of uncertainty in gambling. His research, however, can be applied to all kinds of games. In his emphasis on decision making and the relationship between action and outcome, Epstein echoes some of our own core ideas.

In his book, Epstein identifies three types of decision-outcome relationships, leading to three degrees of uncertainty: *uncertainty*, *risk*, and *certainty*. Each category corresponds to a different kind of decision-outcome relationship and game experience. A game that is completely *certain* is hardly a game at all, and certainly not much fun to play. It is like flipping a two-headed coin: there is no doubt what the end result will be. Sometimes, certainty is contextual. A game of Tic-Tac-Toe between two people that are completely familiar with the logic of the game play has a *certain* outcome: the game will always end in a draw. Although the specific decisions of the players aren't certain, the overall result of the game will be. In a game that is completely certain, meaningful play is impossible.

Epstein's other two categories describe what we normally think of as uncertainty in games. Risk refers to a situation in which there is some uncertainty but the game's players know the nature of the uncertainty in advance. For example, playing a game of Roulette involves placing bets on the possible outcome of a spin and then spinning the roulette wheel to get a random result. There is some uncertainty in the spin of the wheel, but the percentage chance for a particular result occurring and the resulting loss or gain on the bet can be calculated precisely. Of the thirty-one numbers on the Roulette wheel, 15 are red, 15 are black and one of them (the zero) is neither red nor black. If you bet on red, you have 15 out of 31 chances (or 48.39%) to win and double your bet. In other words, in a game of pure *risk*, you can be completely certain about the degree of uncertainty in the outcome of the game.

Epstein's category of *uncertainty* describes a situation in which players have no idea about the outcome of the game. For example, imagine that you are a moderately skilled Chess player and you go to an online game site to play a game of Chess with an opponent that you select at random. You have no idea who you are going to play against. It might be a Chess master, who will most likely beat you, or it might be someone learning to play for the first time, who you will most likely beat. There is no way for you to predict the outcome of the game. If, in contrast, you are playing a friend that you have played many times before and you know that you usually win three out of four games, you have a good sense of the outcome. But without knowing your opponent, you can't make that kind of guess.

Although games of pure certainty are extremely rare (and not much fun to boot), games of pure risk and games of pure uncertainty are also quite rare. Most games possess some combination of risk and uncertainty. Even though you know something about the general chances of winning against your friend, you certainly don't have absolute mathematical certainty about your chances of winning. And although you know the exact risk each time you make a bet on the Roulette wheel, your overall loss or winnings over an evening of play is much more uncertain.

**The Feeling of Randomness**

Roulette and Chess point to a very important aspect of uncertainty. Often, the degree of chance in a game has less to do with the actual mathematics of the game system and more to do with how the player's experience of the game is framed. When we look at only a single round of Roulette, the game is an experience of pure risk. But when we frame it as the gain and loss of money over many rounds, the overall outcome is more uncertain. Similarly, it is possible to produce a feeling of uncertainty in games that do not formally possess an element of chance. Below are two examples:

Chinese Checkers. When four, five, or six players play this game, it can feel quite random. As the game unfolds and players move their pieces, the center of the board becomes crowded with a seemingly random arrangement of pieces. This is true even though every single move on the board is the result of a player making a strategic choice about where to play next.
If you roll more than one die, determining outcomes becomes more complex. With two dice, Knizia notates the basic outcomes in the form of 3–4, where the first number is the number rolled on the first die and the second number is the number rolled on the second die. A chart of all of the possibilities of basic outcomes for two dice are shown in Table 1.

Note that symmetrical outcomes such as 2–5 and 5–2 both appear on the chart, as they represent different possible basic outcomes. There are 36 basic outcomes with two dice, so the chance of anyone outcome appearing is 1/36. When two dice are thrown in a game, instead of the individual results on each die, the game uses the combined total of both dice. We can determine the chance of rolling a combined outcome equal to a particular number in the same way as with a single die: by adding up the basic outcomes. To determine the chance of rolling a 5, count the basic outcomes that add up to 5: 1–4, 2–3, 3–2, 4–1. This is four basic outcomes, and 4/36 = 1/9 or 11.11 percent.

| Total Outcomes | 1–1 | 2–1 | 3–1 | 4–1 | 5–1 | 6–1 | 1–2 | 2–2 | 3–2 | 4–2 | 5–2 | 6–2 | 1–3 | 2–3 | 3–3 | 4–3 | 5–3 | 6–3 | 1–4 | 2–4 | 3–4 | 4–4 | 5–4 | 6–4 | 1–5 | 2–5 | 3–5 | 4–5 | 5–5 | 6–5 | 1–6 | 2–6 | 3–6 | 4–6 | 5–6 | 6–6 |
|----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Percentage | 2.78 | 5.56 | 8.33 | 11.11 | 13.89 | 16.67 | 13.89 | 11.11 | 8.33 | 5.56 | 2.78 |

What is the chance of rolling doubles? There are six basic outcomes that are doubles: 1–1, 2–2, 3–3, 4–4, 5–5, 6–6. This is six outcomes, and 6/36 = 1/6 or 16.67 percent.

Putting all of the two-die outcomes on one chart, we get the figures in Table 2. Notice the radically unequal distribution of probabilities for rolling the highest and lowest numbers, as opposed to rolling numbers in the middle of the range.
We can apply the same basic principles Knizia articulates to a wide variety of game design situations. For example, if your game requires players to flip a coin, you can determine the two basic outcomes as heads and tails, with a 50 percent chance of achieving each outcome. Or you might be designing a computer simulation game that uses lots of random numbers to determine the frequency of events, or a special deck of cards with a particular chance of certain cards appearing each turn. In any of these cases, the general principles remain the same. If players are flipping three coins, how likely is it that they will all come up heads? What is the chance of having a hand of five cards that are all the same?

If you are designing a game that involves dice-rolling, card-shuffling, or other forms of random number generation, it is important that you understand the basic principles of the probabilities involved. However, mathematical principles alone won't lead you to design meaningful play. The key, as with other aspects of games, is in understanding how probability relates to player decisions and outcomes. For example, in designing a board game such as Monopoly, in which players' pieces circle the board on a track, how will you determine the number of spaces on the board? In Monopoly, the board has forty spaces. Because the average combined outcome of a two-die roll is 7, it takes on average six throws to get around the board. This means that by about turn seven, some of the players will likely have already started their second loop, and will begin to land on each other's properties. If you are creating a game with a similar structure, design the board and the use of dice to achieve a pacing of events appropriate for your game.

**Chance and Game Play**

There is a curious relationship between chance and game play. One way of framing chance, especially a game of pure chance, is that players completely give up control and have to passively accept the results of the game as they occur. As anthropologist Roger Caillois writes in his book *Man, Play, and Games*, chance "signifies and reveals the favor of destiny. The player is entirely passive: he does not deploy his resources, skill, muscles, intelligence. All he need do is await, in hope and trembling, the cast of the die."5

With all due respect to Caillois, we wholeheartedly disagree. What Caillois describes may in fact be an accurate depiction of the emotions of some players during a game of chance, but there are plenty of chance-based games that do offer player decision and meaningful play. Even in a game of pure chance, a well-designed game continually offers players moments of choice. Meaningful play requires that at some level a player has an active and engaged relationship to the game and is making choices with meaningful outcomes. A player that does nothing but "await, in hope and trembling, the cast of the die" cannot be engaged in meaningful play.

Let us look again at a game with which we are all too familiar, Chutes and Ladders. Formally, it is a game of pure chance. On your turn, you roll the die, move your token appropriately, and then pass the die to the next player. Players do not make any strategic decisions in the course of play. However, Chutes and Ladders can be a fun game. Even without considering the social, narrative, and cultural forms of pleasure the game might provide, there are any number of ways that Chutes and Ladders provides pleasure through formal aspects of its game play:

- The mechanic pleasure of inhabiting a game system and helping that system move forward by rolling dice, counting spaces, and moving your token.
- The uncertainty of knowing who will win and the struggle to finish first.
- The chutes and ladders themselves, which reinforce both of the previous pleasures. On the one hand, with the erratic swoops of movement they produce, the chutes and ladders make the mechanistic system itself richer and more fun to inhabit. They also allow for unexpected reversals of fortune, increasing the dramatic potential of who will finish first.
Although it is true that almost any game will possess the first two qualities, it is always challenging to harness these two pleasures in the service of meaningful play. In Chutes and Ladders, it is the chutes and ladders themselves that serve as the central feature in the formal game structure to provide interest. Imagine the game without the chutes and ladders: rolling a die, moving a token, and getting to the last space first. As a game, it would be a completely flat experience. The chutes and ladders create a structure that results in more meaningful play, even without any real choices to make.

How does this happen? Consider the formal flow of the game play. Without the ups and downs of the ladders and the chutes, players’ scores would slowly accumulate at a roughly equivalent rate. A player might pull ahead or fall behind, but the pattern of the game would remain fairly flat, with each player moving ahead from one to six spaces each turn. In fact, if a player does move substantially ahead or behind of the rest of the players, he or she is on average more likely to stay there, further reducing dramatic uncertainty. The chutes and ladders provide changes of position greater than the die roll’s relatively modest adjustment of 1–6. These leaps disrupt the otherwise flattened chances of winning.

A second and very different example of pure chance games are lottery-based games. The basic game play of a lottery game is incredibly simple: pick a number or series of numbers and then wait to see if your number or numbers were picked. Again, at first glance, it is difficult to imagine how such a simple game could be so compelling. Yet even though a lottery is a game of pure chance, there are many moments at which players make choices: selecting the kind of lottery game to play, selecting a number or set of numbers, selecting the number of times to enter a given lottery, and even (for regular players) selecting a pattern of play over time. Many lottery games offer additional choices, such as a selection of “scratch-off” spaces on a lottery card. Each moment of choice is an event with the potential for meaningful play. Lottery players often use elaborate systems to help them select numbers, based on past winning numbers, their birthdays, random hunches, or other numerological speculations. The simple choice of what number to play becomes infused with meaning as players explore the space of possible options. Of course, the chance to win money is undeniably an essential part of the appeal of lottery games. However, it is not the only aspect of the game that makes its play meaningful. The opportunities for players to decide how to navigate the system of chance are the decisions that let players rail against pure fate, keep hope alive for winning, and help give the game its meaning.

The lesson learned from successful games of pure chance such as Chutes and Ladders or a lottery game is that meaningful play can occur in systems in which there are no actual strategic decisions to make. In games of pure chance, the players’ relation to the game system needs to be carefully designed. At every moment that they come into contact with the system, the possibilities for meaningful play should be teased out and emphasized.

Case Study One: Thunderstorm

Although Dice Games Properly Explained has an entire chapter on dice probability theory, most of the book offers descriptions and analyses of more than one hundred dice games. The games range from simple children’s games to complex betting and bluffing games. Some of them are of Knizia’s own design, but most of them are traditional games. Following are two simple game examples from Dice games Properly Explained that clearly illustrate the successful integration of chance into game play.
Thunderstorm

This is a popular family game in Germany, there called Gewitter. Hit the required target number, or watch the thunderstorm move close, until lightning finally strikes. Any numbers of players can participate, best with four to eight. You need six dice and a notepad.

Object: The aim of the game is to produce at least one 1 on each turn to become the last remaining player in the game.

Play: One player begins, then play progresses clockwise. The first player throws all six dice. Later players may have fewer dice available, even only one.

- If your throw contains at least one 1, you are fine. Set aside all 1s and pass the remaining dice to the next player. If you roll nothing but 1s, recover all six dice and pass them to the next player.
- If your throw does not contain any 1s, you fail and pass the dice to the next player.

In the course of the game, a six-line house is drawn for each player. Each time you fail, a line is added to your house. When your house is complete and you fail again, your house is struck by lightning and you are out of the game.

The game continues until only one player remains. This player wins.

Even though Thunderstorm is a game of pure chance, the kind of chance that a player faces changes from turn to turn. If you are rolling six dice, you have a relatively safe roll and are quite likely to roll a 1. On the other hand, if the previous player hands you just a single die, your chances of rolling a 1 are much lower. Initially, players are making relatively safe rolls, rolling many dice at once. Occasionally a player will get unlucky and miss rolling a 1, but chances are better than 50 percent that for the first few rolls with four, five, or six dice, they will roll a 1. As 1s appear and these dice are stripped away from the group of rolling dice, the tension mounts and the game accelerates as the chance for rolling a 1 decreases. A single die might be passed for quite some time without anyone rolling a 1. Then suddenly someone rolls a 1, avoids drawing a line on his or her house, and the next player begins the pattern again by rolling all six dice.

As a player, you feel two ways about this progression. It is great to see the other players rolling a single die, not rolling 1s, and adding a line to their houses. On the other hand, as the die approaches you around the circle of players, you would love for another player to roll a 1 because it means that the next player rolls all six dice, making it likely that you will have more dice to roll on your turn. This formal structure of uncertainty results in a game with a compelling dramatic rhythm, which takes place in a number of overlapping cycles:

- Every turn a player throws the dice, establishing a regular pace to the game.
- On top of this rhythm, the reduction of the number of dice from six to one and then back to six again sets up a cycle that lasts for many turns and repeats itself a number of times within a single game.
- Each player also sets up a linear progression of house-building. Although the elements of this construction occur in the same sequence for all players, it happens at a different pace for each player.
The fourth cycle happens near the end, as players begin to drop out of the game and the circle closes until there is only one house left standing, the house belonging to the winner.

The overall result is an exciting game with a sense of dramatic inevitability—the destruction of all of the houses but one becomes a dreadful certainty. What is striking about Thunderstorm (no pun intended) is that all of this complexity arises out of a simple game of pure chance—and no betting. Thunderstorm is an example of a game that provides players with a rich chance-based system that generates surprisingly meaningful play.

**Case Study Two: Pig**

The game of Pig differs from Thunderstorm in that it offers choice within the context of a game of chance. Like Thunderstorm, Pig demonstrates how meaningful play can be designed into a system with a great deal of uncertainty. The description from *Dice Games Properly Explained* is as follows:

**Pig**

This is an amusing family game based on a very simple idea. You throw one die and keep adding to your total. If you do not stop before you roll a 1, everything is lost.... Any number of players can play, best is for three to five. You need one die and a notepad.

**Object.** The aim of the game is to avoid rolling 1s and to be the first player who reaches 100 points or more.

**Play.** One player begins, then play progresses clockwise. On your turn, throw the die:

- If you roll a 1, you lose your turn and do not score.
- If you roll any other number, you receive the corresponding points.

As long as you receive points you can throw again, and again. Announce your accumulated points so that everybody can easily follow your turn. You may throw as often as you wish. Your turn ends in one of two ways:

- If you decide to finish your turn before you roll a 1, score your accumulated points on the notepad. These points are now safe for the rest of the game.
- If you roll a 1, you lose your turn and your accumulated points.

Record all scores on the notepad and keep running totals for each player. The first player to reach 100 points or more is the winner.

The first thing to note about Pig is how it creates interesting game choices from a very simple structure. A core component of the game—to avoid rolling 1s—is actually an inverse of the formal demands of Thunderstorm, where the players attempt to roll 1s. In Pig, the player has to balance the desire to keep on rolling and accumulate a higher score with the risk of rolling a 1, which becomes more and more likely each time the player chooses to roll again.

We can analyze the game mathematically. From a probability point of view, one out of the six basic outcomes spells disaster for the player: rolling a 1. This means there is a 5/6 or 83.33 percent chance of rolling safely each time you roll—or conversely, a 1/6 or 16.66 percent chance of rolling a 1. However, even though the chances are the same for every roll considered in isolation, the more that you decide to roll, the more likely it is that you will eventually roll a 1. If you decide in advance that you are going to roll twice, the chances of rolling a 1 on your turn is the combined outcome of a 2-die roll. In the 36 possible 2-die rolls, there are 11 ways you can roll a 1 (1-1, 1-2, 1-3, 1-4, 1-5, 1-6, 2-1, 3-1, 4-1, 5-1, 6-1). This adds up to 11/36 or 30.56 percent chance to roll a 1 in two rolls.

Each time you roll again, your overall chance of rolling a 1 increases, and as soon as you roll a 1, your entire accumulated points for that turn are erased. The drama of the decision to roll or not to roll is that each time you roll the die, you increase your chances of getting more points, as well as increasing your chances to fail. But because you have control over your decision to roll more than once, you know the degree of risk.
plots out the chances of rolling a 1 with successive rolls in the table above. He also calculates the average points you are likely to earn with a certain number of rolls, using 4 as the average number of points per roll (as your possible earnings are 2, 3, 4, 5 or 6). See Table 3.

Knizia suggests that the best Pig strategy is to stop rolling once you have 20 points or more. However, he also acknowledges that a good player takes into account the progress of the other players as well. There are two sides to the formal strategy of playing Pig. On the one hand, there is the aspect of a single player playing against chance, trying to maximize points earned each turn. But this moment-to-moment decision-making process also takes place within the larger context of the other players' scores. In other words, if you are falling behind, you may want to press your luck to try and catch up; but if you fail, you will have to risk even more to regain ground. If you are in the lead, perhaps you should play more conservatively. But then it might be easier for other players to catch up to you.

Pig is an elegant game design because the player's simple choice to roll or not to roll is a decision that sits at the nexus of many intersecting vectors of game play meaning. The result is a game that is astonishingly simple, strategically deep, and increasingly dramatic. Pig is a great example of how pure chance can be harnessed through simple choices and transformed into meaningful play. Can you say that the decisions in your game are as meaningful the decision to roll in Pig?

Breakdowns in Uncertainty

Luck is very much fate's last hope. It is the play of the last chance. It is the play of everyman.... In this sense it is useful to think of games of chance not only as models of the irrevocability of fate but also as fate fantasized. —Brian Sutton-Smith, The Ambiguity of Play

That ends our brief introductory investigation into the classical operations of probability. But before departing from this subject altogether, we would like to discuss a few ways that probability fails to operate in exactly the way that we think it should. As the "feeling of randomness" we discussed earlier demonstrates, the actual operation of probability does not always match up with the way that players experience or interpret it. For the design of meaningful play, understanding the player's point of view is paramount. The next few sections touch on three problematic contexts for probability: randomness on a computer, strategic manipulation of chance processes, and commonly held fallacies about uncertainty.

Breakdown 1: Computer Randomness

If you are designing digital games, it is important to have a sense of how computer programs generate random numbers. Digital games make extensive use of random algorithms, whether to determine which player goes first, to generate the background texture of a game level, or to randomize the behavior of an in-game agent. Ironically, computers cannot produce random numbers. They can execute algorithms that result in random-seeming results, but they are not capable of producing pure randomness. Why is this so? John Casti offers an explanation:
Back in the early days of computers, one of the more popular methods of generating a sequence of random numbers was to employ the following scheme:

1. Choose a starting number between 0 and 1.
2. Multiply the starting number by 4 ("stretch" it).
3. Subtract 4 times the square of the starting number from the quantity obtained in step 2 ("fold" the interval back on itself in order to keep the final result in the same range).

Given a starting number between 0 and 1, we can use the procedure—often termed the logistic rule—to generate a sequence of numbers that to all appearance is completely random. For example, in such a sequence each of the ten digits 0 through 9 appears with equal frequency and the statistical correlation between groups of digits is zero. Note, however, that the members of this sequence are specified in a completely deterministic way by the starting number. So the sequence is certainly not random in the everyday sense of being unpredictable; once we know the starting number and the rule for calculating an element of the sequence from its predecessor, we can predict with complete confidence what every element in the sequence will be.8

Although Casti uses a historical example, the ways that computer programs generate random numbers today are not fundamentally different. Computers can never compute purely random numbers, because the numbers they provide are always the result of algorithms. A computer program can "flip a coin" internally to determine whether a computer-controlled character will turn left or right with equal probability, but the program is iterating a deterministic formula that only superficially resembles the operation of a random coin flip. The generation of random numbers is a well-heeled problem in computer science. We won't go into detail about it here, except to point out that it remains a challenging dilemma. Still, for most game design purposes, the randomness that computers can generate is sufficiently random.

Usually, an intimate understanding of how computers compute random numbers is not part of what a game designer needs to know. But you should never forget that random functions are not infallible. Eric was once working on a game prototype about swarming microbe-like creatures in a fluid environment (the game was never published). The microbes would grow, give birth, and die, flocking together to seek out food in their environment. Although they exhibited complex behavior, it was more or less clear why the microbes were doing what they were doing. However, one aspect of the game was puzzling. The microbes always tended to seek out the upper left corner of their 2D environment. The designers first thought they had stumbled upon a genuinely emergent behavioral pattern, but couldn’t for the life of them figure out what was causing it. Did it have to do with the way the food multiplied? Or the way a player was handling the mouse? At the same time, the inevitable and universal drift of the microbes was ruining the game play by making the overall behavior too predictable.

Eventually, they discovered that the emergent behavior was coming from an error in a randomizing function. Each timed step, a microbe would move in one of sixteen directions. Even though they could sense their immediate surroundings and moved accordingly (towards food and away from danger), the program always weighted their decision with a random input. The problem was that because of a programming oversight, the program began counting in the upper left corner, and then counted that corner again at the end, giving the randomizer twice as much chance of picking the upper left than any other position. Even though the degree of additional chance this error added was very small, because of the complexity of the system, the emergent effects were quite strong. Once the randomizer was fixed, the corner drift ceased. The lesson? Even if you are not a computer programmer, understand how randomness operates in your game's program.
Breakdown 2: Strategizing Chance

The second example of the unexpected nature of chance is when the use of chance becomes strategic, when players manipulate uncertainty itself during a game. Will your players take randomness at face value, or will they scheme to turn chance into strategy? In "Strategies in Counting Out," an essay in The Study of Games, folklorist Kenneth Goldstein looks at the ways that children aged four to fourteen in northwest Philadelphia in 1967 secretly and expertly manipulated the operation of chance. His study focuses on "counting-out," operations such as "eenie-meenie-miny-moe" that kids use to determine who will be "It" in a traditional neighborhood game like Kick the Can.

Counting Out is not usually considered a game: it is a procedure that helps determine roles in a future game. However, by our definition, we can consider it as a simple game of chance. In Counting Out, a player appoints himself or herself the counter; the goal of the game is to avoid being selected as "It." The quantifiable outcome requires that one player is selected as the loser. The premise of counting-out procedures is that they are patterns of counting that randomly select a player. This is, in fact, the way that the children in the study described the act of counting-out to Goldstein when he interviewed them. However, his essay's conclusion is that despite the fact that the children described Counting Out as a purely chance operation, they used complex and subtle strategic methods to achieve the results they desired.

Following are the six general methods of manipulation that Goldstein observed in use. Many of them represent techniques that would require sophisticated mathematical skills to operate in a group with changing numbers of participants.\textsuperscript{9}

\textit{Specific Rhyme Repertory:} This straightforward strategy requires the counter to select a rhyme of a specific length that will achieve the desired result.

\textit{Extension of Rhyme:} The counting-out rhymes are modular and extendable, and if the rhyme is about to end on someone that the counter does not want to be selected, the counter can spontaneously add an additional phrase or rhyme of the proper length to achieve a different result.

\textit{Skipping Regular Counts:} The counter simply skips himself or herself when going around the circle, if the counter is about to be selected. Although this was the most popular technique employed, it was also the most obvious, and the one most frowned upon.

\textit{Stopping or Continuing:} Because most rhymes do not specify whether the selected player is "It" or whether the selected player is "counted out" and is safe from becoming "It," the counter can decide the significance of the selection after the first player has been picked.

\textit{Changing Positions:} This mathematically intensive strategy entails the counter subtly switching to a new spot in the circle in order to be selected as the next player counted "out."

\textit{Respite by Calling Out:} In this blatant strategy of avoidance, a player will simply call "safe" or "free" and be exempt from counting in the current round. The groups that allowed this technique did place restrictions on it, such as having only one player be able to call "safe" per counting round.

The paradox of Counting Out is that even though players describe it as a game of chance, it is a game with a rich strategic component, in which experienced players can achieve the results they desire. The point of this example, as with the operation of chance in software, is to demonstrate that sometimes the differences between randomness and non-randomness are more subtle than they appear. When you are designing a game, pay close attention to the procedures used to determine randomness and make sure that they operate in the manner that
you intend. Of course, the bigger issues to which this example points is that when you design a game, that game is always going to be used in a particular context by particular players. In our schemas on Breaking the Rules and Games as Social Play, and in many of the chapters within the primary schema CULTURE, we explore in more detail some of the experiential and contextual issues raised by this complex example.

**Breakdown 3: Probability Fallacies**

A third example regarding the problems of probability does not concern computer software or strategy and chance, but the ways that players conceptualize and understand randomness itself. You may have created a game that contains very specific kinds of probabilities, and you may even communicate these to players. But this in no way means that your players will accurately understand the way that chance operates in your game.

Game players will rarely have the same grasp of the random functions of your game system that you do. Game players and the public tend to suffer from a number of fallacies and misunderstandings when it comes to the operations of chance. The following list is a paraphrased sampling from Epstein’s longer list of fallacies in *The Theory of Gambling and Statistical Logic.*

- **Overvaluing the long shot.** Game players have a tendency to overvalue “long-shot” bets that have a low probability of achieving a high gain, in contrast to “safe” bets that have a higher probability of achieving a low gain.

- **The tendency to think of successive chance events as additive.** For example, the chance of rolling a 1 on one die is 1/6 or 16.67 percent. The chance of rolling a 1 with two dice is not 2/6 or 33.33 percent, as you might think at first glance. As we know from probability theory, the chances are 11/36 or 30.56 percent. This difference might seem small in this example, but with successive iterations, the differences between the actual probability and the presumed one can be quite large.

- **The Monte Carlo Syndrome.** This refers to the tendency to think that after a run of failures, a success is likely, and vice versa. In other words, if the Roulette wheel has just landed on a black number, it is not more likely that the next number will be red.

- **Overemphasis on good outcomes.** Given a very unlikely negative outcome and a very unlikely positive outcome, people tend to overemphasize the good one. Epstein uses the example of winning the lottery and being killed in a car accident in the next year. Both have about the same chance of occurring (1 in 10,000), even though most believe that the lottery win is more likely.

- **Lightning striking twice.** Related to the previous fallacy, people tend to believe that highly unlikely negative events will not repeat themselves (such as getting struck twice by lightning), but that highly unlikely positive events will happen again (such as winning the top jackpot on a slot machine). In fact, the chance of a random event occurring is not related to the frequency of past occurrences.

- **Luck.** From a purely mathematical point of view, there is no such thing as luck. People aren’t lucky, dice aren’t lucky, charms aren’t lucky, calendar dates aren’t lucky. However, widespread belief in luck persists, even among experienced game players.

Each of these fallacies has important implications for game design. For example, think about the long shot fallacy. If your game allows players a choice between a long shot and safe bet, you should expect most players to take the long shot and balance your formal system accordingly. Overemphasis on good outcomes and the lightning striking twice fallacies can help keep players optimistic in a game with a large chance element. Even if a player has seen a lot of bad luck, these fallacies keep hope for a turnaround alive.
The larger lesson is that when you design a game with a random element, it is important to understand not just the probabilistic mechanisms of chance, but also the way that players will interpret or misinterpret these mechanisms. All three “breakdowns” of chance highlight common pitfalls to avoid in game design. On the other hand, any of these hiccups in the strictly formal operation of chance could be used positively, as the starting point for a game design:

- It may be true that a computer cannot generate true randomness. Why not make a digital game in which the operation of randomness is intentionally out of balance? Perhaps what seems to be a randomly generated string of numbers is really a secret code that needs to be deciphered. Or in certain locations of the game-world, the player can shift the operation of chance to his or her advantage.

- It may be true that chance operations become strategic elements in the hands of competitive players. You might design a game in which players can legally construct or modify the “random” component of a game, such as spending game money to affect the outcome of a die roll or giving players the ability to strategically stack a deck of cards.

- It may be true that players suffer from probability fallacies. Design a game around one of them. Build a game around luck, in which players pick lucky and unlucky numbers for themselves, rolling dice and trying to avoid unlucky numbers and score the lucky ones.

Any “rule” of game design that you might think of can be broken, and as we will discover in *Breaking the Rules*, broken design rules can often lead to innovative game design ideas.

---

**Meaningful Chance**

*But (chance in a game) is never sure. That's what makes the game interesting. Not only is there a possibility that, despite the odds against us, the chance we take will pay off. There is also the further possibility that, despite the apparent confidence of the players, this hand, which seems to be of markedly unimpressive value, might be, in fact, the best of all. My two kings might win the game for me. That's what confidence games are all about. They provide the opportunity and reward for your display of self as well as for your ability to play well with chance—they call for control over yourself as well as control over the game.—Bernard DeKoven, The Well-Played Game*

In thinking about games as systems of uncertainty, we have looked at the micro-level of chance operation as well as the macro-level of the uncertainty of the game as a whole. One important insight we can apply to both of these levels is that the purely mathematical functioning of uncertainty is insufficient to understand the richness of chance within the mechanisms of a game.

As DeKoven observes, it is true that a player interacts with the system of a game, taking risks, placing bets, and calculating the odds. At the same time, that system is also playing with the player, making demands, rewarding and punishing, and asking for leaps of faith. In thinking about games as formal systems, we cannot ultimately divorce the formal system of a game from the ways that players manipulate and inhabit the system. This is as true for the operation of chance as it was for the emergence of complexity. Uncertainty is in the eye of the beholder, or perhaps, in the play of the player.
Further Reading

*The Broken Dice and Other Mathematical Games of Chance*, by Ivar Ekeland

Part philosophy, part mathematics, and part folklore, *The Broken Dice* is an idiosyncratic book that explores philosophical questions of chance and fate from multiple points of view. Particularly relevant is Ekeland's analysis of the impossibility of generating random numbers on a computer.

**Recommended:**
- Chapter 1: Chance
- Chapter 2: Fate
- Chapter 5: Risk

*Dice Games Properly Explained*, by Reiner Knizia

This book by board game designer Knizia packs in descriptions and analysis of more than a hundred dice games, some of which are original designs by the author. It also contains a chapter on probability theory applied to dice—a great non-technical introduction to the subject. The dice games range from the purely random to the intensely strategic and are a good source for classroom games.

**Recommended:**
- Chapter 3: The Theory of Dice

*The Jungles of Randomness: A Mathematical Safari*, by Ivars Peterson

Ivars Peterson writes popular books about mathematics, and in this playful work he tackles the thorny dilemmas of randomness and probability. He makes common references to games throughout, and the chapter recommended below includes a spectacular analysis of the layout of the Chutes and Ladders gameboard.

**Recommended:**
- Chapter 1: The Die is Cast

**Notes**

4. Ibid. p. 63.
- Uncertainty is a key component of every game. If a game is completely predetermined, the player’s actions will not have an impact on the outcome of the game and meaningful play will be impossible.

- There are two levels at which uncertainty operates in a game. On the micro-level are the actual operations of chance that occur at isolated moments in the system of a game. On the macro-level are larger questions of uncertainty, which relate to the ultimate outcome of the game.

- The relationship between a game decision and a game outcome can have three degrees of uncertainty. A certain outcome is completely predetermined. A risk is an outcome with a known probability of happening. An uncertain outcome is completely unknown to the player. It is rare to find a game of pure certainty, risk, or uncertainty. Most games combine some degree of risk and uncertainty.

- It is possible for a game to possess a “feeling of randomness” even if no actual random mechanisms are present in the game system. This feeling can stem from strategic or social complexities that cannot be predicted in advance.

- A game that has very little feeling of randomness can become too dry or competitive. A game that has too much of a feeling of randomness can become overly chaotic, leaving the players feeling powerless. There is no magic formula for how much randomness should be present in a game. In all cases, the key is to create meaningful play that takes unique advantage of the game structure.

- When designing a game with chance elements, it is vitally important to understand the basic mathematics of probability and how they will impact the system you are designing.

- Even games of pure chance can provide meaningful game play as long as players are given meaningful opportunities to take action within the game system.

- There are many surprising ways that the operation of uncertainty can “break down” in the system of a game:
  - Because computer programs cannot generate true randomness, game designers should be skeptical about the random number-generating algorithms in a game.
  - Players can sometimes take a random component of a game and turn it into a strategic activity.
  - There are many commonly held fallacies about chance.
Cybernetics enforces consistency. It permits change, but the change must be orderly and abide by the rules.—Jeremy Campbell, Grammatical Man
Introducing Cybernetic Systems

Cyberspace. Cyberpunk. A Cyborg Manifesto. The term cybernetic has been appropriated by science fiction and technoculture to mean anything associated with computer technology. In point of fact, the field of cybernetics precedes the advent of digital computers. Mathematician Norbert Weiner coined the term "cybernetics" in his 1948 book Cybernetics or Control and Communication in the Animal and the Machine. The word is derived from the Greek word for steersman or navigator, and appropriately enough, cybernetics studies the regulation and control of systems.

Cybernetics grew out of systems theory and information theory, and like these fields, cybernetics studies a range of subjects, from mechanical and electrical systems to social and biological systems. In looking at the basic principles of cybernetics, we are觸ching on a field filled with great debates and a rich history, a field that greatly influenced contemporary ideas about computer technology and society.

This chapter can only offer a brief introduction to cybernetics, focusing on the ways dynamic systems change over time and the formal structures that allow these changes to occur. What are the rule structures that monitor change within a game system? How does a game system adjust to change over time? What constitutes feedback within a game? How can positive and negative feedback loops be used in the design of meaningful play? Within this schema on Games as Cybernetic Systems, we bring cybernetics to bear on these important game design questions.

Elements of a Cybernetic System

Cybernetics deals with the ways a system gauges its effect and makes necessary adjustments. The simplest cybernetic device consists of a sensor, a comparator, and an activator. The sensor provides feedback to the comparator, which determines whether the machine is deviating from its established norm. The comparator then provides guidance to the activator, which produces an output that affects the environment in some way.

This fundamental process of output-feedback-adjustment is the basis of cybernetics.—Stephen Littlejohn, Theories of Human Communication

As communications theorist Stephen Littlejohn makes clear, cybernetics studies particular kinds of systems. The cybernetic conception of a system is based on the interaction of inputs and outputs with the internal mechanism of a system. Inputs are how the system monitors the environment— they allow the environment to influence the system. Outputs are the ways that the system takes action—they are how the system influences the environment. Through the back-and-forth exchange between the environment and the system, the system changes over time.

A cybernetic system contains three elements: a sensor, a comparator, and an activator. The sensor senses something about the environment or the internal state of a system. The comparator decides whether or not a change to the system needs to be made as a result of the sensor’s reading, and the activator activates that change. Together, these three elements regulate how a system operates and changes over time.

A common example of a cybernetic system is a thermostat. Imagine a hot summer day and a room with an air conditioner that is attached to a thermostat. The thermostat contains the system’s sensor, a thermometer. The thermostat also contains a comparator it can use to compare the temperature of the room to a user-set temperature. If the thermostat measures the air temperature above the set amount, it activates the air conditioner, the activator of the system, which cools down the room.

As the air begins to cool, the system continues to monitor the room temperature. When the room is sufficiently cooled so that the thermostat’s sensor doesn’t register the temperature as being above the set limit, the thermostat no longer sends a signal to activate the air conditioner, and so shuts off the cold air. However, the hot summer sun will begin to heat up the
room again. When the temperature rises above the thermostat's limit, the air conditioner will again be activated. This cyclic behavior of the system is the "process of output-feedback-adjustment" Littlejohn describes. The fact that the cybernetic system is running as a circuit, constantly monitoring itself to see whether or not conditions have been met, is the reason why cybernetic systems are sometimes called feedback systems, or feedback loops.

The thermostat example represents a negative feedback system. The system is negative because it seeks to sustain the same temperature. Instead of letting the room get hotter and hotter from the sun, the system acts to return the room to its normative state.

A positive feedback system works in the opposite fashion. Instead of bringing the system to a steady state, a positive cybernetic circuit encourages the system to exhibit more and more extreme behavior. For example, if the thermostat were reversed so that it only activated the air conditioner when the room was below a certain temperature, we would have a positive feedback system. If the room temperature ever went below the comparator's threshold, it would continue to run, making the room colder and colder, so that the temperature would steadily get lower and lower. Brrrr!

Positive feedback leads to divergent behavior: indefinite expansion or explosion (a running away toward infinity) or total blocking of activities (a running away toward zero). Each plus involves another plus; there is a snowball effect. The examples are numerous: chain reaction, population explosion, industrial expansion, capital invested at compound interest, inflation, proliferation of cancer cells. However, when minus leads to another minus, events come to a standstill. Typical examples are bankruptcy and economic depression.

In every feedback loop, information about the result of a transformation or an action is sent back to the input of the system in the form of input data. With the thermostat, the input data is information about air temperature. If this data causes the system to continue moving in the same direction (the temperature continues to rise), then it is positive feedback. This means that the effect is cumulative. If, on the other hand, the new data produces a result in opposition to the previous result (the temperature is rising, it will now be lowered), the feedback is negative. The effects of negative feedback stabilize the system.

Positive feedback loops create an exponential growth or decline; negative feedback loops maintain an equilibrium. As cyberneticist J. de Rosnay explains,
We could also construct negative and positive feedback loops with a heater. In a negative feedback loop, the heater would turn on when the temperature was below a certain level, raising the temperature until it reached its original state, at which point the heater would shut off. In a positive feedback loop, the heater would turn on when the temperature rose above a certain level, continuing to heat the room indefinitely. Hot hot hot!

Now imagine what would happen if we combined two simple cybernetic systems using an air conditioner and a heater. This dual system would have a sensor to detect the temperature, a double-comparator to compare the room temperature to a pre-established setting, and heating and cooling activators. Using a dual system allows us to control the room temperature in more subtle ways. If both sub-systems were negative feedback systems, the room temperature would be very stable, as both would seek to sustain a middle room temperature. The cooler or heater would turn on when the room became too hot or too cold, and the temperature would always be brought back to its normative position. The system would never let the temperature vary too greatly. This is, in fact, how central heating and cooling works in many homes.

Alternately, both the heating and cooling circuits could be made into positive feedback sub-systems. Whenever the temperature became too hot or too cold, one of the activators would turn on and keep pushing the temperature in that direction. If the temperature setting for the heater were above the temperature setting for the air conditioner, once the room temperature strayed from the middle range, it would never reach the center again. On the other hand, imagine that the heater’s activation temperature was below the air conditioner’s activation temperature. If the room started out in a middle temperature range somewhere between the two activation temperatures, when the two systems were turned on, both activators would begin battling with each other in a tug-of-war to either raise or lower the temperature.

The important thing to notice in all of the heating and cooling examples is that cybernetic systems affect phenomena like temperature in very specific ways. When more than one cybernetic system is operating together, things get complex quite quickly.
Hot and cold negative feedback system

Hot and cold positive feedback system
Feedback Systems in Games

How do feedback systems operate in games? As a cybernetic system, the rules of a game define the sensors, comparators, and activators of the game’s feedback loops. Within a game, there are many sub-systems that regulate the flow of play, dynamically changing and transforming game elements. Do you want your game to move toward a balanced, steady state? Or do you want it to spin wildly toward one extreme or another? Designing feedback loops within your game can help you shape these tendencies. Feedback loops can be tricky to grasp, but they offer a crucial way of understanding how formal game systems function.

Game designer Marc LeBlanc has done a great deal of thinking about the relationship between game design and feedback systems, and this schema is indebted to LeBlanc’s important work on the subject. In 1999, LeBlanc gave a presentation at the Game Developer’s Conference, titled “Feedback Systems and the Dramatic Structure of Competition.” In this lecture, LeBlanc proposed a way of thinking about games as feedback systems, summarized in the following chart:

In this model, the game state represents the current condition of the game at any given moment. In a Chess game, for example, the game state is represented by the arrangement of the pieces on the board, the captured pieces, and which player is about to move next. In a console fighting game such as Virtua Fighter 4, the game state includes which two combatants were chosen, the health and other fixed and variable stats of the two fighters, their relative spatial positions, and the arena in which they are fighting. The game state is a formal way of understanding the current status of the game, and does not take into account the skills, emotions, and experience of the players. Of course, these player-based factors will definitely affect the game state. If you are a masterful Virtua Fighter 4 player and your opponent is not, this will be evident in the play of the game. However, the game state itself refers only to the formal, internal condition of the game.

The other elements of LeBlanc’s model correspond directly to the components of a cybernetic system as we have discussed them. The scoring function is the system’s sensor that measures some aspect of the game state. The controller is the comparator, which looks at the sensor’s reading and makes the decision whether or not to take action. The game mechanical bias is the activator, a game event or set of events that can be turned on or off depending on the decision of the comparator.

When looking at games as cybernetic systems, it is important to note that we are not necessarily considering the entire game as a single feedback system. Instead, our emphasis is on the ways that cybernetic systems are embedded in games. Embedded cybernetic systems affect a single aspect of a larger game, such as determining which player goes first next round or the relative speed of players in a race. We know from our study of systems that all parts of a game are interrelated in some way. A cybernetic system within a game that directly affects just one component of a game will indirectly affect the game as a whole.

Positive and Negative Basketball

To bring this abstract discussion closer to game design, let’s look at several game examples. In his talk “Feedback Systems and the Dramatic Structure of Competition,” LeBlanc invented two variations on the formal structure of Basketball: Positive Feedback Basketball and Negative Feedback Basketball. Each variation adds just a single rule on top of the existing formal structure of the game:
Negative Feedback Basketball: For every \( N \) points of difference in the two teams' scores, the losing team may have an extra player in play.

Positive Feedback Basketball: For every \( N \) points of difference in the two teams' scores, the winning team may have an extra player in play.

How do the addition of these rules change the game? Say, for example, that \( N \) is 5. In a game of Negative Feedback Basketball, when Team A fell behind by 5 points, it would gain a player on the court and begin to play with a team of 6. As soon as Team A scored points that put it behind by less than 5, it would drop its extra player. On the other hand, if Team A continued to do poorly, when its score was 10 points behind the other team, it would gain a second extra player. Why is this an example of negative feedback? Because the adjustments in the system (gaining and losing players) encourage the system to move toward a stable, steady state. A losing team gets extra players, which helps it catch up to the winning team; when it moves to within 5 points, the two teams are evenly matched. The steady state of this system is not that the total points tend towards zero, but that the difference between the two teams' scores stays near zero. The end result is that Negative Feedback Basketball games would tend to be very close games.

Positive Feedback Basketball creates the opposite situation. As soon as one team increased its lead, it would gain additional players. These new players would help the team do even better against the opposing team, which would increase the winning team's lead even more, which would result in yet more players for that team. Eventually, the court would be absurdly crowded with members of one team, who would completely overwhelm and defeat the team with only five players. Positive Feedback Basketball encourages a large difference between the two teams’ scores, so that there is a runaway, devastating victory instead of a closely matched game.

As in the examples of heating and cooling, there are many ways to transform the game system. We could, for example, change the rules to remove players instead of adding them. In this case, in Negative Feedback Basketball, when one team pulls ahead by \( N \) points, it would lose a player, making it easier for the other team to catch up. In Positive Feedback Basketball, the team that was behind by \( N \) points would lose a player, encouraging them to fall further behind, which would result in the loss of even more players. Eventually, one team would fall so far behind that none of its players would be left on the court. In both games, even though players are removed rather than added, the end results remain the same: Negative Feedback Basketball tends toward stable, close matches and Positive Feedback Basketball tends toward unstable, unbalanced matches. Each variation on the game of Basketball would result in vastly different player and spectator experiences. Yet all we did was add one rule that affected the behavior of the system. Feedback systems offer game designers a powerful tool to affect a game’s formal structure and the way that structure manifests in play.

Racing Loops
Positive Feedback Basketball and Negative Feedback Basketball were variations on the game of Basketball. But many existing games already make use of feedback systems in their designs. Here, we look at the use of cybernetic systems in two digital racing games.

Wipeout is a science-fiction racing game originally released for the Playstation, in which the player pilots a fast-moving hover vehicle around a track, trying to beat the computer-controlled vehicles and come in first place. It is common in racing games such as Wipeout for the program to employ feedback mechanisms. Obviously, a computer program can drive a vehicle as poorly or as skillfully as the game designer wants. It would be simple to program the computer-driven cars so that they drove in a mathematically optimal fashion and always beat the player. However, that would simply not be fun. Instead, in racing games the computer vehicles are programmed to drive in a
less than "perfect" manner, sometimes not steering or accelerating efficiently, in order to provide a challenge that a human player can overcome.

One way to create a scaled challenge for the player would be to program different skill levels for the computer-controlled vehicles. Some vehicles would be easy for a beginner to beat whereas others could only be bested by experienced players. Programming a static skill level for each opponent vehicle, however, is not yet a cybernetic feedback loop. Why would we want to add a feedback system to a racing game? In order to keep the flow of play exciting, of course. Part of the fun of a racing game such as Wipeout is jockeying for position among a dense cluster of hover vehicles, battling for first place with another racer who is hot on your tail or dead ahead in your sights.

Without a feedback loop, these moments are unlikely to occur. What if a player crashes early in a race—will she ever catch up to the computer-controlled vehicles? Or what if a player's skill far outmatches the pre-programmed computer opponents? Once she gains a lead early in the race, she might as well be racing alone, because the computer opponents will never catch up to her.

This is precisely why Wipeout (and many other digital racing games) make use of cybernetic feedback systems to control the speed of the computer opponents. There are two general rules we can abstract from the behavior of the computer-controlled vehicles in the game. Although these are not the only factors determining their speed, they do have a clear impact on the experience of the game:

- If the human player is in first place, the vehicle in second place will accelerate and catch up to the human player's vehicle.
- If the human player is in last place, the last few vehicles will slow down to let the player catch up to them.

The result of these two rules is a negative feedback system. Like Negative Feedback Basketball, together these two rules operate to reduce the distance between vehicles in the game, eliminating the "extremes" of the player being very far ahead or very far behind the computer opponents.

In this system, there are three states that the comparator needs to monitor: when the player is in first place, when the player is in last place, or when the player is in neither first nor last place. If the player is somewhere in the middle of the pack, then no special activator event comes into play. But if the player is in first or last place, vehicle behavior adjusts accordingly. The outcome of this feedback system is that racing in Wipeout tends to offer exciting and satisfying play. Significantly, Wipeout only affects the computer-opponent vehicles, not the hovercraft that the player is driving. In essence, the program carefully adjusts the competitive backdrop, rather than boosting or handicapping the player directly. However, there are games that apply a negative feedback system more directly to a player's abilities.

One example of such a game is Super Monkey Ball for the Nintendo GameCube. Super Monkey Ball contains several different game modes; one of them is a racing game in which up to four players simultaneously race monkey characters through
a series of tracks. When players drive through a power-up object on the track, they gain a special power that can be used one time. These powers range from forward-firing attacks (shoot a bomb at another player ahead of you) to rear-based attacks (drop a banana peel, hoping a player behind you will run over it and slip) to non-attack powers (a speed-up that temporarily boosts a player’s velocity).

Whereas many racing games use this power-up convention (including Wipeout), Super Monkey Ball uses a feedback system to determine which power-up a player will receive, depending on whether the player is ahead or behind other players. If a player is in last place, the player is much more likely to receive the speed-up power, which will help that player catch up to the other competitors. On the other hand, a player in first place is more likely to get forward-firing attacks, rather than speed-ups or rear-based attacks. The lead player thus receives the least useful kind of power-up: a player in first place can’t use a forward-firing attack to better his position, because no one is ahead of him. These rules add up to a negative feedback system. As with Wipeout, Super Monkey Ball’s feedback loops encourage a close race, in which no player is too far ahead of or behind the others.

In Super Monkey Ball and Wipeout, negative feedback loops are used to engender meaningful play. As we know from Games as Systems of Uncertainty, the outcome of a game needs to be uncertain for meaningful play to occur. If, as a player, you fall so far behind or ahead of the other players that the outcome is a foregone conclusion, meaningful play is diminished, because decisions you make won’t have an impact on the outcome of the game. This does not mean that feedback systems guarantee a close race every time: skill plays an important role in racing games, and it is possible for a player in Super Monkey Ball to fall so far behind that there is very little chance of victory. There is no universal strategy for crafting meaningful play. But in Wipeout and Super Monkey Ball, feedback systems support meaningful play by making the game responsive to the ongoing state of the game.

Positive Feedback in a Game

Not all games use negative feedback systems. Some make good use of positive feedback systems as well. Powerstone is a console fighting game for up to four players that features cartoonish, fast-paced brawling action. In Powerstone, when a player is successfully hit by a powerful attack, the target of the attack will be stunned for a short time, during which the target player cannot move his or her character, and the attacker can continue
to strike the stunned character. A stunned character cannot move out of the way, defend, or counterattack, and is much easier to hit. This is, in effect, a positive feedback system: because a player successfully launches a powerful attack, the player has a better chance of launching yet more attacks as the target remains stunned and easier to hit. The new attacks continue to stun the target, increasing the ability of the attacker to deliver more damage. Positive feedback creates dramatic results, in which a player can be devastated by a rapidly delivered series of attacks. This kind of slapstick action makes sense in a lighthearted and humorous game such as Powerstone.

However, if the positive feedback system were permitted to play itself out until the end of a match, then the game as a whole wouldn’t work. Once your character was hit for the first time, you would remain stunned while an opponent continued to attack you. In effect, the first blow landed would determine your fate and the game would lose the back-and-forth struggle that is an important ingredient of fighting games. Powerstone gets out of this feedback trap by adding a different behavior to the game system. After receiving a certain number of attacks, a stunned character will be hurled across the playfield, making it impossible for the attacker to indefinitely continue rapid-fire strikes. The attacker can pursue the character that flew across the playfield, but the far-flung character is usually no longer stunned by the time the attacker gets there.

Because positive feedback systems are inherently unstable and push a game system toward an inevitable outcome, they are usually dampened by other game factors that limit the acceleration of the feedback loop. In real-time multiplayer strategy games such as Warcraft II, players gather resources, which allow them to build more units that can gather yet more resources, increasing the acceleration of resource-gathering. In this way, all of the players are building their own positive feedback loops, joined together in an arms race to see who will gather enough resources and be the first to front an army capable of winning the game.

In Warcraft II, potentially unstable positive feedback loops are balanced by the fact that each player is creating his or her own feedback loop in parallel. Furthermore, these feedback loops help bring the game to conclusion. Because of their complexity, real-time multiplayer strategy games can sometimes drag on interminably, with players evenly matched and unable to get an upper hand. Because of the way that positive feedback systems can quickly grow out of control, a player that can gain a slight advantage (such as capturing a resource-rich gold mine from another player) can use the advantage to overwhelm an opponent. Obviously, there are many strategic factors other than the resource-feedback loops that determine the outcome of Warcraft II. (For example, skillful battle tactics can help defeat a more resource-powerful opponent.) However, positive feedback systems are clearly a key element of the game design and contribute to the successful play of the game.

**Dynamic Difficulty Adjustment**

Increasingly, digital game designers are incorporating more sophisticated feedback techniques into their game designs. The game developer Naughty Dog Entertainment is known in the game industry for what it calls “Dynamic Difficulty Adjustment,” a technique it has used in the Crash Bandicoot series of games, as well as the more recent Jak and Daxter.

Dynamic Difficult Adjustment, or DDA, uses feedback loops to adjust the difficulty of play. For example, in the original Crash Bandicoot game, the player is generally maneuvering the character Crash through a series of jumping and dodging obstacles, trying to overcome damaging hazards and reach objectives to finish the level. When a player dies, the game restarts at the beginning of the level or at the most recent “save point” reached in the level.

The danger in designing this kind of game is that players possess widely varying skill levels. An experienced gamer might breeze through a level, whereas a beginner might become frustrated after dying several times without making any progress.
The DDA operations in Crash Bandicoot evaluate the number of times that a player is dying at a particular location in a level, and make the game easier as a result. A player having trouble might suddenly find that there are more helpful objects nearby, or fewer enemies to avoid. This kind of attention to the balancing of player experience is evident in the play of Crash Bandicoot games, and it helps explain the fact that a wide audience of both hardcore and less experienced players enjoys them.

Using DDA and other feedback mechanisms in games raises some fascinating game design issues. If we consider the millennia-old tradition of pre-computer play, games are traditionally about a player or players competing within a formal system that does not adjust itself automatically to player performance. As you play a game such as Baseball or Othello, your fluency with the system and your ability to manipulate it grows. The game itself and the other players provide the challenge for you. As your play deepens, you find new forms of play, new ways of expressing yourself within the system of the game.

DDA points to a different kind of game, a game that constantly anticipates the abilities of the player, reads the player’s behavior, and makes adjustments accordingly. Playing a game becomes less like learning an expressive language and more like being the sole audience member for a participatory, improvisational performance, where the performers adjust their actions according to how you interact with them. Are you then playing the game, or is it playing you? Is a game “cheating” if it constantly adjusts its own rules? Could such a scheme be designed into a multiplayer experience and still feel “fair” for everyone involved? These questions have no definitive answers, as there are always many solutions for any given game design problem. Dynamic Difficulty Adjustment could be considered a heavy-handed design tool that takes agency away from the player, or it could be considered an elegant way of invisibly shaping game play so that every player has an optimal experience. Regardless of your opinion on the matter, DDA is an important tool, and as digital games rely more and more on their ability to automate complex processes, this kind of design strategy will become more common.

A Simple Die Roll
Because most of the examples used so far have come from complex digital games, we wanted to finish by looking at a cybernetic feedback system within a more minimal game context. Sometimes in games, there is no game AI or referee to sense and activate the changes in the game state. However, elegant feedback systems can still emerge directly from the game rules.

Let us take a look at one of our favorite examples, Chutes and Ladders. Chutes and Ladders is an extremely simple game of pure chance. But can you spot the feedback system in it? It is not the actual chutes and ladders. Yes, those seem like they regulate the positions of the players, but they do not act in a cybernetic way. They merely randomly shift the position of the players on the board. The chutes and ladders do not constitute a dynamic feedback loop.

The feedback loop in Chutes and Ladders occurs at the very end of the game, when players must land exactly on the final square in order to win (rather than being able to overshoot the final space and land there anyway). This rule creates a kind of negative feedback system. The exact landing rule serves as negative feedback on the distance between players. In a game of Chutes and Ladders, the player that is farthest ahead will eventually be within six spaces of the finish square and will usually end up spending a few more turns trying to make the exact roll, or possibly inch ahead by rolling small numbers. During this time, the other players often catch up. The overall effect is to level out the playing field by reducing the difference between the positions of the players. The result of stretching out the end of the game in this way is a closer and more dramatic finish.
Think about the game without this rule. If players can overshot the final space and still win, imagine that you are playing against someone who is just three spaces away from the last square. Even if that player has very bad luck (rolling three 1s in a row), that player is no more than three turns from winning the game. If you are more than 18 spaces away (the total of rolling three 6s in a row), there is no way you can win. On the other hand, if your opponent has to make an exact roll, then he has a 50 percent chance of rolling too high so that he has to stay put, as you keep getting closer. The game is prolonged, the outcome remains uncertain, and in general, the game is more satisfying to play. Those last few die rolls become dramatic, nail-biting game events.

We should point out that this is not a true example of a cybernetic feedback system. An orthodox systems theorist would point out that there is no sensor, comparator, and activator in actual operation. As a counter-example, if there were a rule requiring that the player in first place subtract 1 from his die roll, we would have a true feedback loop, in which a procedural change is enacted when certain conditions are met. Here the player is the sensor, the rule itself the comparator, and the activator is the action of subtracting one from the die roll.

Coming back to our exact landing rule, if we frame the rule in the following fashion, we might consider it to have a feedback loop: “If a player is fewer than 6 spaces from the final space, then rolling higher than N, where N is the number of spaces between the player and the final space, has no effect.”

The rule now feels more like a feedback loop, where the player senses proximity to the finish and the rule acts to limit the effectiveness of the die roll. Ultimately, it does not really matter whether an orthodox systems theorist would approve of this example or not. As designers, the value of a schema is its ability to solve design problems. The rule that requires players to land by exact count on the final space does create more meaningful play. Understanding the rule as a cybernetic feedback loop, or even a pseudo-cybernetic feedback loop, can only enhance our appreciation for the game’s design.

Putting Feedback to Use
As a game design schema, *Games as Cybernetic Systems* is one of the most practically applicable frameworks presented in this book. Cybernetic feedback systems can be wonderful ways of balancing your game to arrive at a particular result. What is wrong with your game: Is it ending too soon? Running on for too long? Is it too uncertain? Not uncertain enough? Is it too easy or too difficult for players to gain an advantage? You can address all these fundamental questions by looking for feedback loops existing within the formal structure of your game’s design, or by adding additional loops of your own.

In his lecture, Marc LeBlanc boiled down the relationship between game design and feedback systems to a set of design “rules.” These rules offer a useful set of guidelines for integrating feedback systems into your design. Here are a number of LeBlanc’s “rules” and some of our comments on each of them:

- **Negative feedback stabilizes the game.**
- **Positive feedback destabilizes the game.**

These two observations form perhaps the most fundamental cybernetics insight for game design. As a designer, you should be aware of the ways that your game creates stabilities and instabilities. If your two-player card game lets the most powerful player take cards from the weaker player, then you have created a positive feedback system where the most powerful player will quickly dominate. The game is unstable, and will rapidly fall out of balance. Perhaps the solution is to add more players to the game and allow them to team-up on the player that is ahead. This would be adding a negative feedback system to re-balance the game and make it less likely that a player who gains a small advantage will end up winning.
Although our examples have emphasized negative feedback as a useful game design tool, too much negative feedback can make a game too stable. Imagine a variation on Chutes and Ladders in which, whenever a player is ahead of another player, all of the players go back to the start. Although this rule would certainly add negative feedback to the game, ensuring that no player would get ahead of the others, it stabilizes the game to the point of stasis, so that the game doesn’t move forward at all. Finding a balance of negative and positive factors for your game is crucial in designing meaningful play.

- Negative feedback can prolong the game.
- Positive feedback can end it.

LeBlanc’s next two “rules” should follow intuitively from our many examples of feedback systems. Positive feedback can rush a game to conclusion, rewarding a player that is already ahead, as in Warcraft II. Negative feedback, as in the Chutes and Ladders exact landing rule, makes it easier for a losing player to catch up, prolonging the game by reducing the winning player’s lead.

- Positive feedback magnifies early successes.
- Negative feedback magnifies late ones.

These two “rules” follow closely from the last pair. In Warcraft II, an early advantage in establishing positive feedback resource loops can put a player too far ahead of the other players. In Chutes and Ladders, on the other hand, negative feedback at the very end can allow a player that has been behind the whole game to catch up.

How can you apply these ideas to your own game design? It depends on the kind of game experience you want to create. There are no universal guidelines for the proper length of a game. Wargamers might play a game for weeks, whereas less hardcore gamers might think an hour is a long time to be playing a single game. Similarly, there are no fixed rules that tell you to make the opening moves or the ending moves the most important ones in the game.

Your guide to making these kinds of decisions should be the core principles of meaningful play. Regardless of the length of your particular game, you should strive to create meaningful play at all moments, where the game outcome is uncertain until the end and every action a player takes can help determine that outcome in an integrated way. In general, players that play well should be rewarded with victory. But perhaps there is always a chance for a dramatic turn of events at the end, where the first becomes the last and the last becomes the first.

- Feedback systems can emerge from your game systems “by accident.” Be sure to identify them.
- Feedback systems can take control away from players.

LeBlanc’s final few “rules” are crucial. Game systems are complex and unpredictable and you can never be sure what feedback systems might be hiding out in the space of possibility you are constructing. Feedback systems can be great ways of shaping player experience, but as LeBlanc warns, as you incorporate systems into your game that actively reshape the experience, you run the danger of removing player agency, leaving your players feeling powerless. Some feedback systems, such as the last space rule of Chutes and Ladders, are relatively innocuous. But many game players will feel “cheated” if they can detect a game adjusting itself to their play. If that second place car is always on your tail, does it really matter how well you perform? Perhaps there should be limits on the speed of the second place car, so that a truly masterful player can have the satisfaction of driving far ahead of the rest of the pack.

As this last example clearly demonstrates, the most important thing about players and control is not their actual control in a game, but their feeling of control in the experience of play. We
explored this phenomenon in the schema on Uncertainty, and it is just as valid here. Meaningful play is, after all, measured by what a player experiences, not by the underlying rules of a game.

**Afterword: Don’t Forget the Participant**

Before departing a discussion of cybernetic systems entirely, we would like to make a few critical comments on the field. Cybernetics is clearly a formal way of understanding systems, which is why the schema of Games as Cybernetic Systems belongs within our RULES primary schema. However, as with all formal schemas, there are many things that cybernetics fails to address.

As a field, cybernetics initially considered a system as a completely self-contained entity. Cybernetics played into the classical scientific idea that the observer of a system had no effect on the operation of the system. This initial model of a cybernetic system was rocked by the introduction of second-order cybernetics into the field. Second-order cybernetics took the observer into account as a part of the system itself, undermining the “objective” stance of classical cybernetics. The insight of second-order cybernetics is that to observe a system in operation is to be part of that system. Although many thinkers, such as Katherine Hayles in her book How We Became Post-Human, have since criticized second-order cybernetics for falling into many of the same objectivist traps as its predecessor, second-order cybernetics went far in attempting to understand systems within a larger context.

What does all of this mean for game design? For the purposes of this schema, we made use of the more “classical” first-order cybernetics. We looked at games as self-contained systems, ensconced entirely within the magic circle demarcated by the rules. Occasionally we peeked a bit at the way formal changes play out in the experience of a game, but by and large we kept to the formal mechanics of game systems. This formal emphasis, of course, is what the RULES schemas are all about. The concept of looking at games as formal systems is to leave out all of the emotional, psychological, social, cultural, and contextual factors that influence the experience of the game for the players. In the PLAY and CULTURE sections of this book, we do in fact look at games as much more than self-contained systems. For the time being, however, we continue our rules-based investigations. Even considered as purely formal structures, there are still many layers to the complex phenomena of games for us to uncover.

**Further Reading**

*How We Become Post-Human*, by Katherine Hayles

Hayles’s book is less an explication of cybernetic theory and more an ideological critique of the field. However, her detailed research on the development and evolution of cybernetics and second-order cybernetics provides a great deal of insight to how these movements intersect with cultural beliefs about technology. In this complex book, Hayles also relates these subjects to literary theory and contemporary ideas about computer technology, virtuality, and identity.

**Recommended:**

- Chapter 3: Contesting for the Body of Information: The Macy Conferences on Cybernetics
- Chapter 4: Liberal Subjectivity Imperiled: Norbert Wiener and Cybernetic Anxiety
- Chapter 6: The Second Wave of Cybernetics: From Reflexivity to Self-Organization

*Theories of Human Communication*, by Stephen W. Littlejohn (see page 200)

**Recommended:**

- Chapter 3: System Theory

**Notes**

1. <pespmc1.vub.ac.be>.
Cybernetics studies the behavior of self-regulating systems. A cybernetic system consists of three elements:

- A sensor that measures some aspect of the system or its environment
- A comparator that compares this measure to a set value and decides whether or not to take action
- An activator that creates a change in the state of the system

For example, in an air conditioner, the sensor and comparator are in the thermostat, which activates the air conditioner activator to cool down a room when the temperature gets too high.

Cybernetic feedback systems can be positive or negative:

A negative feedback system is stabilizing and brings a system to a fixed, steady state. The air conditioner example, which keeps a room from getting too hot, but shuts off when the room cools down, is a negative feedback system. The temperature remains within a narrow range.

A positive feedback system is cumulative and makes a system unstable. If the air conditioner turned on when the temperature was below a certain number, then the room would become colder and colder, moving away from a stable state.

- A game can contain many feedback systems that interact with each other within the larger system of the game.
- Many game feedback systems are negative, reducing the advantage or disadvantage of a player or a team. This phenomenon is common in digital racing games.
- Games also make use of positive feedback systems for dramatic effect or to bring a game to conclusion. Often, a positive feedback system is countered by a negative feedback system in a game. Powerstone's stunning and hurling features demonstrate positive and negative feedback systems working together.
- Dynamic Difficulty Adjustment, or DDA, is the modification of a game's challenge according to player performance. It is most often used in complex single-player digital games.
Game Designer Marc LeBlanc outlines a number of design "rules" that apply cybernetics to game design. These "rules" include the following:

- Negative feedback stabilizes the game.
- Positive feedback destabilizes the game.
- Negative feedback can prolong the game.
- Positive feedback can end it.
- Positive feedback magnifies early successes.
- Negative feedback magnifies late ones.
- Feedback systems can emerge from your game systems "by accident." Be sure to identify them.
- Feedback systems can take control away from the players.

In the field of cybernetics, the more classical first-order cybernetics, which considers a system as a self-contained entity, was challenged by second-order cybernetics, which includes the observer of a system as an element of the system. Within this formal schema, we have not made use of second-order cybernetic thinking.
These are the main problems: How does each player plan his course—i.e., how does one formulate an exact concept of a strategy? What information is available to each player at every stage of the game? What is the role of a player being informed about the other player's strategy? About the entire theory of the game?—Oscar Morganstern and John Von Neumann, Theory of Games and Economic Behavior
Introducing Game Theory?
Perhaps you thought that this entire book was about game theory. If that were the case, what does “Games as Game Theory Systems” mean? Actually, game theory is not what it may appear to be. It is not a general term that means theoretical approaches to games. Game theory means something quite specific: it is a branch of economics that can be traced back to the work of two mathematicians, Oscar Morganstern and John Von Neumann. The classic text in the field is Theory of Games and Economic Behavior, published in 1942.

Game theory is the mathematical study of decision making. It looks at how people behave in specific circumstances that resemble very simple kinds of games. The founders of game theory intended to create a new kind of mathematical approach to the study of economics. Morganstern and Von Neumann were writing during a time when Marxism was very much in vogue in the field of economics, and Theory of Games and Economic Behavior was, in many ways, an attempt to replace the ideological approach of Marxism with a more rational and scientific set of techniques. Although it caused quite a sensation when it was introduced, the promises of game theory were never quite fulfilled, and it has largely fallen out of favor as a methodology within economics. But game theory can still be quite useful for game designers.

The theory of games is a theory of decision making. It concerns how one should make decisions and, to a lesser extent, how one does make them. You make a number of decisions every day. Some involve deep thought, while others are almost automatic. Your decisions are linked to your goals—if you know the consequences of each of your options, the solution is easy. Decide where you want to be and choose the path that takes you there. When you enter an elevator with a particular floor in mind (your goal), you push the button (one of your choices) that corresponds to your floor.¹

As a formal game design schema, Games as Game Theory Systems looks at games as systems of rational choice. It is potentially useful to game designers for two chief reasons. First, it analyzes situations that resemble simple games in a very detailed way. Even more importantly, as game theorist Morton D. Davis points out in the previous quotation, game theory specifically focuses on relationships between decisions and outcomes. We know from our earlier discussion of interactivity that actions and outcomes are the building blocks of meaningful play. Within this schema, we explore questions of how players plan their course of action within a game and how they formulate strategies and make decisions. From decision trees to degenerate strategies, we will look closely at the application of game theory concepts to the design of meaningful play.

Decision Trees
As a formal approach to understanding games, game theory looks at games as a series of strategic decisions made by the players of a game. What does it mean to reduce a game to its strategic decisions? One common game theory method is to create a decision tree for a game. A decision tree is a branching tree-style diagram that outlines all of the possible moves a player can make in a game. Decision trees are a common way of flow-charting interactive experiences. For example, if you are programming an interactive story that has a hypertext structure, you might draw a diagram that shows all of the links between the different parts of your story. This kind of diagram would be a decision tree.
Creating a decision tree for a game is more complicated than creating a decision tree of a hypertext structure. The difference is that in a typical hypertext, the links and the actions that a player can perform at any location in the larger structure do not change as the participant moves through the structure. A reader's first choice in a hypertext structure does not change the way that the other hypertext links function. The only thing that changes is where the reader is positioned in the structure.

A game is more complex. In a game, what you can do at any given moment depends on what has already happened in the game. At the beginning of a game of Chess, for example, you can't move either of your rooks because they are both blocked by pawns. Later in the game, you might be able to move your rooks if your pawns have been maneuvered out of the way. The complexity of games leads to a system of many possible actions. Which actions can happen at a given moment is contingent on the current state of the game.

Because Chess is a complicated game to diagram as a decision tree, let's start with a simpler example: Tic-Tac-Toe. In Prisoner's Dilemma, a book about game theory and its historical context, writer William Poundstone leads us through the process of making a decision tree of the game of Tic-Tac-Toe.

Tic-Tac-Toe starts with the first player ("X") putting a mark in any of nine cells. There are consequently nine possible first moves. The nine choices open to Player X on the first move can be diagrammed as nine lines radiating up from a point. The point represents the move, the moment of decision, and the lines represent the possible choices.

Next it's Player O's move. There are eight cells still open—which eight depending on where the X is. So draw eight secondary branches at the top of each of the nine primary branches. That leaves seven open cells for X on his second move. As the diagram of possible moves is continued upward, it branches like a very bushy tree.

As you continue the process, you will eventually diagram moves that put three markers in a row. That's a win for the player who moves. It's also the termination of that particular branch in the diagram, for the game ends when someone gets three in a row. Mark that point (call it a "leaf" of the diagram) as a win for X or O as the case may be.

Other branches of the diagram will terminate in a tie. Mark them as ties. Obviously, the game of ticktacktoe cannot go on forever. Nine moves is the maximum. So eventually, you will have a complete diagram of the game of ticktacktoe. Every possible ticktacktoe game—every game that ever has been played or ever will be played—must appear in the diagram as a branch starting at the "root" (X's first move) and continuing up to a "leaf" marked as a win for X, a win for O, or a tie. The longest complete branches/games are nine moves long. The shortest are five moves (this is the minimum for a win by the first player).

Creating a decision tree can be a powerful way of understanding the formal structure of a game. It is in essence a way of mapping out a game's formal space of possibility. For a simple game such as Tic-Tac-Toe, the complete space of possibility can in fact be diagrammed. However, not all games can be mapped out in this way.

Being able to make a decision tree of a game or other interactive structure implies that the decisions participants make are discrete decisions that lead to knowable outcomes. For example, a game that involves physical skill, such as American Football, does not have self-contained moments of decision making that can be diagrammed like the alternate turn-taking of Tic-Tac-Toe. Instead, the game exists as a continuous flow of action. When the ball is hiked, a quarterback does not take a single discrete action. Instead, the game flows forward in a complex web of activity. Perception, movement, and the granularity of the real world creates a non-discrete game space.

Although the moment-to-moment play of Football is continuous, the game can be broken down into a system of separate
plays. Does that mean it is possible to create a decision tree of Football by widening the frame of analysis, so that each decision point on the chart represents the choice of a play by one team's coach? The answer is no. The problem with this proposal is that to create a decision tree, the result of a decision needs to be a knowable outcome or set of outcomes. Think about a game of Tic-Tac-Toe. When a player makes a decision to place an X or an O in a particular square, there isn't any doubt that the player will finish the action and make the mark. On the other hand, just because a Football team picks a certain play does not mean that they will be able to successfully complete it, or complete it in a way that can be predicted with any accuracy. The outcome of picking a particular play in a Football game could result in a yardage loss or gain, a penalty, a fumble, a reversal, or a touchdown, making it impossible to diagram the outcome in the same way we could Tic-Tac-Toe. (Note also that "X and O" play diagrams that show where players will run on certain plays can be used to schematize the play of Football. But these play diagrams do not qualify as decision trees.)

What kinds of games can we turn into decision trees? Decision trees work for any game that has the following qualities:

- Time in the game takes place in turns or other discrete units.
- Players make a finite number of clear decisions that have knowable outcomes.
- The game is finite (it can't go on forever).

Although this disqualifies many games (including Football), it does include a wide variety of games, such as turn-based strategy games like Tic-Tac-Toe, which clearly fulfills all three criteria listed above. What about Chess? Chess takes place in turns and decisions have clear outcomes, but is it finite? Chess might seem like an infinite game (imagine an endgame with two kings shuffling back and forth between the same squares forever), but in fact there are rules that resolve the game in a stalemate when a certain number of moves have elapsed without a capture. How about a game such as Chutes and Ladders? It seems to fit the three criteria, but it does have a random die roll. Could we map it out with a decision tree? Surprisingly, yes we could. The first point or "root" of the decision tree would have six branches coming out, depending on what the first player rolled. Each of those six branches would have six more, depending on what number the next player rolled. And so on. Of course, there would have to be a different tree for a two-player game, a three-player game, and a four-player game.

Although the decision tree for games as simple as Tic-Tac-Toe might seem large, a decision tree for a game such as Chess or Chutes and Ladders would be extraordinarily vast and complex. Remember that the decision tree contains all of the possible games that have ever or will ever be played. The decision tree for Chutes and Ladders would have to contain every possible die roll at every possible moment in the game with every possible arrangement of players on the board, in every possible sequence that could logically occur. The decision tree for Chess would have to contain every possible move and every possible response to every possible response to every possible move. The decision trees for these games would be immense. According to Poundstone, if a decision tree for Chess were graphed out on paper at a legible size, the diagram would span the solar system.

If decision trees for games are so unwieldy in the real world, how are they possibly useful for game designers? Decision trees are more theoretical constructs than engineering tools. At the same time, the ability to understand what a decision tree is and how it works is crucial to game design. Why? Because a decision tree is also a diagram of the formal space of possibility of a game. Being able to conceptualize the space of possibility you are designing is an important game design skill.
Even though true decision trees are usually impossible to create, often you can create very useful decision trees for sections or aspects of a game. For example, say you are designing a mission-based strategy game that contains many level "missions" that the player has to complete. A player can succeed or fail at a mission, and the next mission depends on the outcome of the most recent mission. While it might be impossible to draw a decision tree of the battle that takes place within an individual mission, it would be extremely useful to chart out the relationships between missions.

![Decision Tree Diagram]

Making a decision tree of the game's missions will tell you, for example, how many missions a single player will play through in an average game. Or it will help you eliminate game designs that loop back on themselves. When you can make use of them, decision trees are a straightforward and useful way of understanding the structure of a game. Perhaps more importantly, however, decision trees are an important part of understanding game theory.

**Strategies in Game Theory**

Decision trees help us understand how players move through the space of possibility of a game. To see how this works, think back to the Tic-Tac-Toe decision tree. The tree contains every conceivable move, in every possible iteration of the game. This is actually more information than we need. Most players will not randomly pick their next square, but will actively try and score three in a row while keeping an opponent from doing the same. With this in mind, we can start to trim all of the "stupid move" branches from our tree. Poundstone describes what this process of "trimming" would be like:

Go through the diagram and carefully backtrack from every leaf. Each leaf is someone's last move, a move that creates a victory or a tie. For instance, at Point A, it is X's move, and there is only one empty cell. X has no choice but to fill it in and create a tie.

Now look at Point B, a move earlier in the game. It is O's turn, and he has two choices. Putting an O in one of the two open cells leads to the aforementioned Point A and a sure tie. Putting an O in the other cell, however, leads to a win for X. A rational O player prefers a tie to an X victory. Consequently, the right branch leading upward from Point B can never occur in rational play. Snip this branch from the diagram. Once the play gets to Point B, a tie is a forgone conclusion.

But look: X could have won earlier, at Point C. A rational X would have chosen an immediate win at Point C. So actually, we can snip off the entire left branch of the diagram.

Keep pruning the tree down to the root, and you will discover that ties are the only possible outcomes of rational play. (There is more than one rational way of playing, though.) The second player can and will veto any attempt at an X victory, and vice-versa.
From this pruned-down version of Tic-Tac-Toe, it is possible to create what game theory calls a strategy. A strategy in game theory parlance offers a more precise meaning than what is commonly meant by "strategy." A common understanding of a strategy in Starcraft might be: "If you're playing the Zergs, create a lot of Zerglings at the beginning of the game and rush your opponent's central structures before they have time to build power." A strategy in this casual sense is a set of general heuristics or rules of thumb that will help guide you as you play. However, a strategy in game theory means a complete description of how you should act at every moment of the game. Once you select a strategy in the game theory sense of the word, you do not make any other choices, because the strategy already dictates how you should act for the rest of the game, regardless of what the other player does. This can make game theory strategies quite intricate. Poundstone lists a sample strategy for Tic-Tac-Toe for the first player X.

Put X in the center square. O can respond two ways:

1. If O goes in a non-corner square, put X in a corner cell adjacent to O. This gives you two-in-a-row. If O fails to block on the next move, make three-in-a-row for a win. If O blocks, put X in the empty corner cell that is not adjacent to the first (non-corner) O. This gives you two-in-a-row two ways. No matter what O does on the next move, you can make three-in-a-row after that and win.

2. If instead O's first move is a corner cell, put X in one of the adjacent non-corner cells. This gives you two-in-a-row. If O fails to block on the next move, make three-in-a-row for a win. If O blocks, put X in the corner cell that is adjacent to the second O and on the same side of the grid as the first O. This gives you two-in-a-row. If O fails to block on the next move, make three-in-a-row for a win. If O blocks, put X in the empty cell adjacent to the third O. This gives you two-in-a-row. If O fails to block on the next move, make three-in-a-row for a win. If O blocks, fill in the remaining cell for a tie.4

As you can see, the strategy for even a simple game such as Tic-Tac-Toe is somewhat complex. A complete strategy is ultimately a methodology for navigating the branches of a decision tree. A strategy prescribes exact actions for the player utilizing the strategy, but it also has to take into account all of the possible branches that an opposing player could select. In Poundstone's example, the strategy dictates the way that the first player would move from the root of the tree to the first of nine possible points. From there the opposing player could move to any of the other eight points, a move that the strategy has to take into account.

A complete strategy for a game such as Chess would be mind-bogglingly huge. However, game theory does not study games as strategically complicated as Chess. In fact, the games that game theory studies are remarkably simple. But as we already know, even very simple games can play out in quite complex ways.

Game Theory Games

Game theory demands a sacred character for rules of behavior which may not be observed in reality. The real world, with all its emotional, ethical, and social suasions, is a far more muddled skein than the Hobbesian universe of the game theorist.—Richard Epstein, The Theory of Gambling and Statistical Logic

Now that we have outlined decision trees and strategies, we are ready to take a look at what it is that game theory calls a game. As the mathematician Richard Epstein points out, game theory games are not about real-world situations or about all kinds of games. Game theorists look at very particular kinds of situations in a very narrow way. What kind of situations? We can summarize a game theory game in the following way: a game theory game consists of rational players who simultaneously reveal a strategy to arrive at an outcome that can be defined in a strict measure of utility. Usually, game theory limits itself to games with only two players.
Rational play, simultaneity, strategy, outcome, utility, and two players. Let us look at each of these elements separately. First, game theory focuses its attention on rational players. Rational players are perfectly logical players that know everything there is to know about a game situation. Furthermore, rational players play to win. As Poundstone puts it, “Perfectly rational players would never miss a jump in checkers or ‘fall into a trap’ in Chess. All legal sequences of moves are implicit in the rules of these games, and a perfectly logical player gives due consideration to every possibility.” As we know from our detailed investigation of Tic-Tac-Toe, if two rational players played the game, the outcome will always end in a draw, because both players would select strategies that would stalemate the other player. Rational players are a fiction, of course, as Epstein makes clear. Real-world players are not like game theory players, as rational as Mr. Spock, completely immune from “emotional, ethical, and social” liabilities. But rational players are still a useful theoretical construct, for they allow us to look at games in a very isolated and controlled way.

The fact that rational players follow a strategy is an important aspect of a game theory game as well. As we mentioned previously, a strategy is comprehensive. It is a complete plan for playing an entire game, from start to finish. A strategy includes explicit instructions for playing against any other strategy an opponent selects. In a game theory game, both rational players simultaneously choose and reveal their strategies to each other. In other words, instead of the “I take my turn, you take your turn” pattern of many games, in a game theory game, players only make one decision, at the same time, without knowing what the other player will do. In making a simultaneous decision, a player has to take into account not just the current state of the game, but also what the opponent is thinking at that very moment. A classic example of a simultaneous decision game is Rock-Paper-Scissors, in which both players have to decide what they are going to do based on the anticipated action of the other player.

So although game theory does not study psychology directly, there is a psychological element in game theory games, where players might consider “bluffing” or using other indirect strategies against each other. Though they might take these kinds of actions, rational players are still psychologically predictable. Players in a game theory scenario are never going to be vindictive, forgetful, self-destructive, or lazy, as this would change their status as rational players. In game theory games one can always assume that both rational players are acting in their own best interest and are developing strategies accordingly.

Why would game theory choose blind, simultaneous decision making as the game play process that it studies? Remember that game theory is not a form of game design: it is a school of economic theory. Within an economic situation, decisions have to be made without knowledge of how the other “players” are going to act. Should you sell your stock in Disney, or buy more shares? Should you purchase two gallons of milk this week, or buy one and wait to see if the price goes down? Should a nation increase or decrease import taxes? All of these micro- and macro-economic scenarios involve making decisions. But the outcome of the decision is based on factors outside the decision maker’s direct control. Simultaneous, blind decisions offer a way of simulating this decision making context, a context that lies at the intersection of mathematics and psychology. As Morganstern and Von Neumann explain,

> It is possible to describe and discuss mathematically human actions in which the main emphasis lies on the psychological side. In the present case, the psychological element was brought in by the necessity of analyzing decisions, the information on the basis of which they are taken, and the interrelatedness of such sets of information (at the various moves) with each other.

Another important component of a game theory game is utility, which is a mathematical measure of player satisfaction. In order to make a formal theory of decision making, it was necessary that Von Neumann and Morganstern numerically quantify the
desire of a player to achieve a certain outcome. In a game theory game, for every kind of outcome that a decision might have, a utility is assigned to that decision.

A utility function is simply a "quantification" of a person's preferences with respect to certain objects. Suppose I am concerned with three pieces of fruit: an orange, an apple, and a pear. The utility function first associates with each piece of fruit a number that reflects its attractiveness. If the pear was desired most and the apple least, the utility of the pear would be greatest and the apple's utility would be least.

Utility can become more complex when multiple factors come into play. For example, if you were building a house for yourself on beachfront property, thinking in game theory terms, you could measure different locations of your house in terms of utility. You might be able to get the highest utility, say +10, if you built right on the beach. There might be a lower utility, such as +5 or +2, if you had to build it several meters away from the shoreline.

On the other hand, if you had to build the house so far away from the beach that the ocean was no longer in view, your utility might go into the negative numbers, indicating an outcome that you would find unpleasant. Of course, you might not have the money to afford the situation with the highest utility. For example, you might require a house of a certain size and if it were directly on the beach it couldn't have a basement and would have to be smaller. Or the cost of the house might be higher on the beach because of the extra architectural complexity required to build in the sand. Cost, size, and location would all be assigned different values. In making your decision, you would try and maximize the total utility given your available options.

These examples touch on the ways that game theory employs the concept of utility. It might seem silly to turn something like human satisfaction into a numerical value, given the innumerable complexities that go into our feelings of pleasure, but Morganstern and Von Neumann felt very strongly that a scientific theory of economics necessitated such an approach. In their book, they use an analogy to physical properties such as heat. Before scientists developed a way of conceptualizing and measuring heat, it was an unknown, fuzzy property that seemed impossible to measure: a sensation that occurred as one approached a flame. But the precise measurement of heat is now an important part of contemporary physics. The aim of Von Neumann and Morganstern was to begin a similar revolution in economics, by quantifying pleasure as a measure of utility.

Utility may well be an oversimplification of human desire, but it does make a good fit with the formal qualities of games. As we know from our definition of games, all games have a quantifiable outcome: someone wins, or loses, everyone wins or loses, or player performance is measured in points, time, or some other numerical value. The concept of assigning a numerical utility to decision outcomes is really just another way of creating a quantifiable outcome. When looking at games through a formal frame, we do not have the luxury of being non-numerical. The formal systems of both digital and non-digital games require an exactness that does in fact come down to numbers. How many kills did you earn that round? What qualifying time do you need on the next heat in order to continue the race? Which team won the game? These very simple game results are all quantifiable outcomes, and are all examples of utility as well.

The last component of most game theory games is that they are usually played by only two players. This was not part of the original formulation of game theory as proposed in Theory of Games and Economic Behavior. The original idea was that the theory could cover n-player games, where n was a number of any size that indicated the number of players. But Von Neumann and Morganstern found that, as with the problem of three planetary bodies discussed in Games as Emergent Systems, their theory became vastly more complex when it took three or more players into account. As a result, most game theory work has focused on two player games. We follow suit in the material to follow.
Cake Division

It is finally time to take a look at a real game theory game. The following description is taken from *Prisoner's Dilemma* and is the classic "cake division" game theory problem:

Most people have heard of the reputed best way to let two bratty children split a piece of cake. No matter how carefully a parent divides it, one child (or both!) feels he has been slighted with the smaller piece. The solution is to let one child divide the cake and let the other choose which piece he wants. Greed ensures fair division. The first child can't object that the cake was divided unevenly because he did it himself. The second child can't complain since he has his choice of pieces...

The cake problem is a conflict of interests. Both children want the same thing—as much of the cake as possible. The ultimate decision of the cake depends both on how one child cuts the cake and which piece the other child chooses. It is important that each child anticipates what the other will do. This is what makes the situation a game in Von Neumann's sense.

Game theory searches for solutions—rational outcomes—of games. Dividing the cake evenly is the best strategy for the first child, since he anticipates that the other child's strategy will be to take the biggest piece. Equal division of the cake is therefore the solution to this game. The solution does not depend on a child's generosity or sense of fair play. It is enforced by both children's self interest. Game theory seeks solutions precisely of this sort.

The cake division problem contains all of the elements of a game theory game listed earlier. There are two rational players (the children motivated by self interest). These two players choose a strategy about how to behave (how to cut or select the pieces). These strategies result in some kind of utility for the two players, measured in how much cake they get. Note that even though the "play" of this very simple game consists of a two-part action (first slice the cake and then choose a slice), the two players can still reveal and enact their strategies simultaneously. For example, the strategy of the player that chooses from the two pieces is always going to be "take the bigger piece." A rational piece-choosing player is going to choose this strategy regardless of the strategy that the cake-cutting player takes. (Note that although these "strategies" may seem like forgone conclusions rather than choices, this is because this game theory game has a saddle point, a concept explained in detail later on.)

A powerful analytical tool provided by game theory is to map this decision making process into a grid. One axis of the grid represents one player's decision. The other axis represents the other player's decision. The cells in the grid represent the outcomes reached depending on which decisions were made. A game theory table of this sort is called a payoff matrix (payoff being another term for utility). Figure 1 shows a payoff matrix for the cake division problem, taken from *Prisoner's Dilemma*. Note that William Poundstone makes the assumption that the cake slicing is going to happen in an imperfect world, so that even if the child that cuts the cake tries to slice it evenly, the two resulting slices will still differ a tiny bit, say by one crumb.

Along the left side of the matrix are strategies that the cutter can take: either cut the cake evenly or cut it unevenly. Although there are any number of ways to cut the cake, these are the two essential strategies from which the cutter can choose. Across the top of the matrix are strategies the chooser can take: choose the bigger piece or choose the smaller piece. The cells show the utility or payoff for only one of the players (the cutter), but it can be assumed that the inverse payoff would happen for the chooser. If the payoff matrix indicates that the cutter receives the "small piece," the chooser would therefore receive the "big piece." This is also true for half of the cake plus or minus a crumb.
The cake division problem illustrates two important game theory concepts. The first is the concept of a zero-sum game. In a zero-sum game, the utilities of the two players for each game outcome are the inverse of each other. In other words, for every gain by one player, the other player suffers an equal loss. For example, playing a version of Poker in which everyone puts money into a pot is a zero-sum game. At the end of the game, every dollar won by one player is a dollar lost by another player. A group of gamblers playing Roulette is not a zero-sum game between the players, because they are not playing directly against each other. On the other hand, if we frame Roulette so that one player is playing against the casino, then it is a zero-sum situation: if a player wins a dollar, it is taken from the house, and vice-versa.

Many games are zero-sum games, even those that do not involve money. When one player wins a game of Checkers and the other player loses, the loss by one player equals a gain for the other player. In this case, game theory would assign a utility of −1 for the loss and +1 for the gain. The utilities add up to zero, which is exactly why it is called a “zero-sum” game. Some games, such as the cooperative board game Lord of the Rings, are not zero sum games. In the basic version of Lord of the Rings, players cooperate against the game system itself. Players either all lose, or they all win. Because the players are not competing against each other, they either all get a −1 for losing or all get a +1 for winning. The losses and wins among the players do not add up to zero; it is therefore not a zero-sum game.

Not all game theory games are zero-sum games, but many are. Cake division is clearly a zero-sum game. Consider the problem intuitively: there is only so much cake, which is all going to be divided into two slices and eaten. The more cake that one player eats, the less the other player eats. We could assign the following utilities to the four cells of the diagram:

```
<table>
<thead>
<tr>
<th></th>
<th>-1</th>
<th>+1</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10</td>
<td></td>
<td>+10</td>
</tr>
</tbody>
</table>
```

We know that the two player outcomes are inverses of each other. If one player receives half of the cake minus a crumb (−1) the other player will receive half of the cake plus a crumb (+1). The total is zero. Cake division is a zero-sum game.

Why is this important? Because, according to game theory, every finite, zero-sum, two-player game has a solution (a proper way to play the game), the strategy that any rational player would take. What is the solution to the cake division problem? The game will always end in the upper left corner. The cutter will get half of the cake minus a crumb and the chooser will get half of the cake plus a crumb. Why is this so? Look at the cutter’s strategies. The cutter would love to end up with the lower right cell, where he gets the big piece. So perhaps he should choose the strategy of cutting the pieces unequally. But the cutter also knows that if the chooser is given the chance to choose, the chooser will always choose the bigger piece. As a result, the cutter has to minimize the bigger piece that the
chooser will select by cutting the cake as evenly as possible. The game resolves to the upper left corner.

This situation clearly illustrates another key game theory concept: the saddle point property of payoff grids. In cake division, each player is trying to maximize his own gains while minimizing the gains of the other player. When the choices of both players lead to the same cell, the result is what Von Neumann and Morganstern call a saddle point. A saddle point refers to a saddle-shaped mountain pass, the intersection of a valley that goes between two adjacent mountains. The height of the pass is both the minimum elevation that a traveler going across the two mountains will reach, as well as the maximum elevation that a valley traveler crossing the mountain pass will achieve. The mathematical proof of saddle points in games is called the minimax theorem, which Von Neumann first published in 1928, many years before the 1944 publication of Theory of Games and Economic Behavior.

The concept of saddle points is extremely important in game design. In general, you want to avoid them like the plague. Remember, a saddle point is an optimal solution to a game. Once a player finds it, there is no other reason to do anything else. Think about the cake division saddle point: if either player deviates, that player will lose even more cake. If you think of the space of possibility that you are crafting as a large 3D structure carefully crafted to give a certain shape to the experience of your players, saddle points are short-circuits in the structure that allow players to make the same decision over and over. That kind of play experience does not usually provide very meaningful play. Why? Because if there is always a knowable saddle point solution to a game, a best action regardless of what other players do or what state the system is in, the game loses the uncertainty of possible action. Meaningful play then goes out the window.

Saddle points do not just occur in game theory games. Many fighting games are ruined, for example, because despite all of the special moves and combinations that are designed into the game, the best strategy to use against opponents is simply to use the same powerful attack again and again and again. Saddle point! Another common occurrence of saddle points involves the programming of computer opponents. In many real-time strategy games there are “holes” or weaknesses in the AI that allow for saddle points. If a player discovers that the computer opponent does not know how to defend well against a certain type of unit, he is likely to abandon all other game strategies and simply hammer on the AI’s weakness over and over, regardless of how much care went into carefully designing missions that require different kinds of problem-solving. Saddle point!

This style of play, based on exploiting a strategic saddle point, is called an exploit or degenerate strategy. A degenerate strategy is a way of playing a game that ensures victory every time. The negative connotation of the terms “exploit” and “degenerate” imply that players are consciously eschewing the designed experience in favor of the shortest route to victory. There are some players that will refuse to make use of degenerate strategies, even after they find out about them, because they wish to play the game in a “proper” manner. On the other hand, many players will not hesitate to employ a degenerate strategy, especially if their winnings are displayed in a larger social space outside the game, such as an online high score list.

Degenerate strategies can be painful for game designers, as players shortcut all of the attention lavished on a game’s rich set of possibilities. Try to find degenerate strategies and get rid of them! We learned in the previous schema that positive and negative feedback systems can emerge unexpectedly from within a game’s structure and can ruin a game experience for players. The same is true of degenerate strategies. A close analysis of your game design can sometimes reveal them but the
only real way to root them out is through rigorous playtesting. If you see players drawn to a particular set of strategies again and again, they may be exploiting a weakness in your design.

**Playing for Pennies**

Not all game theory games have a saddle point. Consider a simple game that requires a more complex playing strategy: Matching Pennies, another classic game theory problem. Here is how the game works: two players each have a penny. Hiding their penny from view, both players pick a side, heads up or heads down, and then simultaneously reveal their pennies. If they match, Player 1 gets both pennies. If they don't match, Player 2 gets them. We can graph this game on a payoff grid:

<table>
<thead>
<tr>
<th>Player 1's choices</th>
<th>Heads</th>
<th>Tails</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heads</td>
<td>-1 cent</td>
<td>1 cent</td>
</tr>
<tr>
<td>Tails</td>
<td>1 cent</td>
<td>-1 cent</td>
</tr>
</tbody>
</table>

This table shows the outcomes for Player 1, the player that wins if the pennies match. Because this is another zero-sum game, the utility for Player 2 is the inverse of Player 1's payoff. What is the proper way to play this game? What strategy should a rational player choose: heads or tails? There does not seem to be a single best answer to the question. If one player decided to pick heads or tails as a permanent strategy, the other player could take advantage of this strategy and win every time. But Matching Pennies is a finite, zero-sum, two-player game, and game theory should be able to solve this game and provide the proper strategy for two rational players. The solution turns out to be more complex than the cake division problem: players do not choose a single, fixed strategy, but select a mixed strategy. In a mixed strategy, players choose one of their options according to a certain probability ratio. For Matching Pennies, the mixed strategy requires rational players to randomly pick heads or tails, with a 50/50 chance of selecting either one.

Remember that rational players will attempt to maximize their own gains in utility while minimizing the gains of their opponents. If rational players play many, many games of Matching Pennies, they will end up with an average utility of zero. This means that neither player will ever come out ahead, but that is the best that they can hope for in this "game."

**The Prisoner's Dilemma**

Of course, it is possible to construct payoff grids in many different ways, and they do not always have to be zero-sum. In fact, constructing game theory problems that are intentionally less symmetrical than Mixed Pennies and Cake Division can lead to some very perplexing "games." One famous game theory problem is called the Prisoner's Dilemma. It is from this problem that William Poundstone takes the title of his book. He describes the "story" behind this game as follows:

Two members of a criminal gang are arrested and imprisoned. Each prisoner is in solitary confinement with no means of speaking to or exchanging messages with the other. The police admit they don't have enough evidence to convict the pair on the principal charge. They plan to sentence both to a year in prison on a lesser charge. Simultaneously, the police offer each prisoner a Faustian bargain. If he testifies against his partner, he will go free while the partner will get three years in prison on the main charge. Oh yes, there is a catch... If *both* prisoners testify against each other, both will be sentenced to two years in jail.

The prisoners are given a little time to think this over, but in no case may either learn what the other has decided until he has irrevocably made his decision. Each is informed that the other prisoner is being offered the very same deal. Each prisoner is only concerned with his own welfare—with minimizing his own prison sentence.9
Game theorists do not agree on the proper solution to the Prisoner's Dilemma. There are two ways of thinking about this problem. Using a minimax approach, it is clear that it is always better to defect, no matter what the other prisoner does. If you defect and the other prisoner does not, you get the best possible outcome. But if the other prisoner decides to defect, then it is a good thing you did too, because you saved yourself from the worst possible outcome. According to this logic, both players will defect and the rational outcome is the lower right cell of the payoff grid. The other approach is to say that because both players are rational and because the payoff grid is symmetrical, both players will make the same choice. This means that the two players are choosing between the upper left and the lower right cells. Given this choice, two rational players will end up choosing the better of their two options, the upper left, where they receive only one year of jail time.

The Prisoner's Dilemma remains an unsolved game theory problem. It clearly demonstrates that even very simple sets of rules can provide incredibly complex decision-making contexts, which raise questions not just about mathematics and game design, but about society and ethics as well.

**Game Theory and Game Design**

Game theory is a curious thing. It promises to be a detailed theory of decision making in a game context. At the same time, its relationship to real-world games seems incidental: the "games" that game theory studies are far removed from the kinds of games that most game designers would like to create.

Does that mean that game theory is irrelevant to game design? Absolutely not. This schema on Games as Game Theory Systems, like most of our other RULES-based schema, borrows concepts and theories from disciplines that make a formal study of systems. Like systems theory, complexity theory, information theory, and cybernetics, game theory was not created in order to assist in the game design process. But that doesn't mean that it isn't relevant to designers.
Decision trees that mark out a game’s formal space of possibility; utility that measures the desire of a player for a given game outcome; saddle points that erase meaningful play—game theory is rife with connections to some of our core design concepts. Game theory games are microcosms for game design problems, an opportunity to plot out a simple decision in great detail and appreciate the complexity that even elementary moments of choice can generate. Game theory, as a formal approach to understanding decisions, is an extremely useful game design tool.

The rules of games constitute systems of incredible subtlety and complexity. As a design discipline with a very young history, game design must turn to these more established ways of thinking in order to try and make sense of the phenomena of games. Perhaps as the field matures, the theoretical borrowings that take place in this book will be replaced by more game-centric schools of thought. At least, we certainly hope so.

Further Reading

*Emergence: From Chaos to Order*, by John Holland (see page 169)

Recommended:
- Chapter 2: Games and Numbers
- Chapter 3: Maps, Game Theory, and Computer-Based Modeling

*Prisoner’s Dilemma*, by William Poundstone

Prisoner’s Dilemma combines a biography of John Von Neumann with an analysis of Cold War politics and a detailed explanation of game theory. It is the clearest non-technical book on game theory we have found, with a range of detailed examples. Taken as a whole, Prisoner’s Dilemma helps put game theory in its proper historical and cultural context.

Recommended:
- Chapter 3: Game Theory
- Chapter 6: Prisoner’s Dilemma
- Chapter 12: Survival of the Fittest

Notes

3. Ibid., p. 46.
4. Ibid., p. 48.
9. Ibid., p. 118.
Game theory is a branch of economics that studies rational decision making. It often looks at game-like situations, but it is not a general theory of games or game design.

A decision tree is a diagram that maps out all of the possible decisions and outcomes that a player can take in a game. A completed decision tree is equivalent to the formal space of possibility of a game. A game must have the following characteristics to be reducible to a decision tree:

- Time in the game takes place in turns or other discrete units.
- Players make a finite number of clear decisions that have knowable outcomes.
- The game is finite (it can’t go on forever).

Even if a game meets these criteria, most games are too complex to be diagrammed as a decision tree. Decision trees are most useful for mapping aspects of games, or as conceptual tools for thinking about the formal structure of a game.

A game theory game is limited to rational players who simultaneously reveal a strategy to arrive at an outcome, which can be defined in a strict measure of utility. Usually, game theory limits itself to games with only two players.

- A rational player doesn’t exist in the real world. A rational player is a completely logical player that plays only to maximize winnings, regardless of emotions, ethics, and social attachments.
- A game theory strategy is a complete plan for playing a game. A strategy explicitly and comprehensively covers every possible situation that a player might encounter in the course of playing a game, including every possible strategy that an opponent might select.
- In a game theory game, rational players make a simultaneous decision about what strategy to take. They know the complete rules of the game and the possible outcomes of their decisions, but they do not know the strategy that the other player will take.
- The results of a game theory game are measured in utility, which is a numerical representation of the players’ desire for a certain outcome. Attractive outcomes are assigned higher positive numbers, and less attractive outcomes are assigned lower numbers. Negative numbers represent an unpleasant utility.
- A payoff matrix is a grid of cells used to diagram the possible outcomes of a game theory problem.

- In a zero-sum game, the winnings of the victor are equal to the losses of the loser. Games such as Chess with a single winner and a single loser are zero-sum games.

- Every two-player, zero-sum game theory game has a solution, a proper way to play the game that will maximize winnings for the player every time. When there is a single best solution to a game for both players, the solution is known as a saddle point.

- Saddle points in any game can lead to degenerate strategies, also called exploits. A degenerate strategy is a way to play a game that leads to victory every time. Generally, degenerate strategies are to be avoided in games because they diminish uncertainty and meaningful play.

- Some game theory solutions consist of mixed strategies, where players select among different strategies with a weighted percentage.
A video game usually mimics some real-life situation: rockets accelerating and moving in space, bouncing Ping-Pong balls, a kayak in river currents, the food-chain in an ecology. The game of Chess is an abstraction based on a battle between two small groups of warriors: similarly, video games imitate life. A video game is a simulation, a model, a metaphor. —Warren Robinett, Inventing the Adventure Game
Introducing Simulation

*Games as the Play of Simulation* is our third and final schema exploring the play of representation. In *Games as the Play of Meaning*, we examined how games become meaningful through the process of signification. In *Games as Narrative Play*, we unearthed the wealth of techniques by which games tell stories. For the purposes of this schema, we hone in tightly on the mechanics of play itself, and the way representations are constructed dynamically, through interaction with a game. How, for example, does the board game Diplomacy simulate the art of negotiation? How does the paper-based game Ace of Aces dynamically represent World War I air combat? How does the digital game Deus Ex depict action and intrigue through designed algorithms and rules? The concept of simulation lies at the intersection of representation and dynamic systems. As simulations, games create representations, but they do so in a very particular way: through the process of play itself.

We look for answers to questions regarding games and simulations by focusing on the representational mechanics of game systems. A game creates representations in many ways, from its instruction manual text and imaginative fictive world to the visual design of its spaces and the audio design of its soundtrack. At the center of all of these depictions is the game system itself. This system generates representations from a player’s interaction with the game, out of the experience and logic of play. This special class of representations, experienced as procedures, sets of behaviors, or forms of interaction, is the raw material from which simulations are constructed. We call this form of depiction *procedural representation*. A simulation-based approach to representation in games, it is the central concept of this schema.

However, procedural representation is only part of what we study in this chapter. In addition to exploring the mechanisms of procedural representation, we also investigate the relationship of those representations to the world outside the game. We know something is a simulation, in part, because we are familiar with the thing that it is simulating. Diplomacy is a political simulation because it mimics processes of negotiation that are known and familiar in the real world. Yet even though Diplomacy faithfully models the art of negotiation, its representation is still in some measure artificial, contained within the game, separate from the real world. The relationship between a game and the "reality" that it depicts is a fundamental aspect of considering games as simulations.

This is not our first mention of representation and "reality." Back in *Games as the Play of Meaning* we introduced the concept of the cognitive frame. A cognitive frame is a way of organizing or understanding the world, a framework that shapes interpretation and therefore what we take things to mean. Considering not only how a game simulates, but also what it simulates raises questions regarding the relationship between the artificial world of a game and the "real life" contexts it intersects. These questions will play an important role in our understanding of games as simulations, and will become increasingly important as we move into our primary schema on CULTURE.

Defining "Simulation"

A video game is an imaginary world: its inhabitants are nonexistent creatures that nevertheless the eye can see, and the hand can move. It is imaginary in the sense that there is no solid reality behind the picture. A bouncing ball may be faithfully simulated, but that moving blip of light has no real mass or elasticity. The ball’s position, velocity, mass, and elasticity are just numbers stored in the computer that controls the video game; and the laws of physics that govern the ball’s trajectory and its bounce are just mathematical equations stored in the computer’s program.—Warren Robinett, *Inventing the Adventure Game*

In *Inventing the Adventure Game*, Warren Robinett, the game designer and programmer best known for the Atari 2600 game Adventure, looks at games through the lens of simulation. He is particularly interested in the way that digital games are "imag-
inary worlds," as he puts it, in which players experience blips of light and sound as a representation of some other real-life situation. His description of a simulated bouncing ball, in which the sensory components of position, velocity, mass, and elasticity are peeled back to reveal the underlying mechanisms of the programmed software, reminds us of the often hidden relationship between the formal structure of a game and the experience of that structure through play.

Robinett specifically addresses the way that representations in video games "mimic" real-world phenomena as diverse as bouncing balls and warring soldiers."A video game is a simulation, a model, a metaphor," writes Robinett. What exactly does he mean? What is a simulation? How are games simulations? Is every game a simulation? What is the relationship between a simulation, a model, a metaphor, and the real-world? We tackle these thorny questions in the pages to follow. But first, let us take a moment to define the concept of simulation. The educational game reference A Handbook of Game Design provides a good starting point:

A *simulation* can be defined as "an operating representation of central features of reality." This definition again identifies two central features that must both exist before an exercise can reasonably be described as a simulation. First, it must represent an actual situation of some sort — either a situation drawn directly from real life, or an imaginary situation that conceivably could be drawn from real life (invasion by extraterrestrial beings, for example). Second, it must be operational, i.e., must constitute an on-going process — a criterion that effectively excludes from the class of simulations static analogues such as photographs, maps, graphs, and circuit diagrams, but includes working models of all types.!

The authors Eddington, Addinall, and Percival identify two components that make a representation a simulation. First, a simulation represents something: an "actual situation," which is either a circumstance from real life or an imaginary situation that is conceivably real. This component points out the referential qualities of a simulation: a simulation refers to something in the real world. It is significant that the authors use the phrase "central features of reality" rather than just "reality" when describing what a simulation represents. As we will see, a simulation cannot depict every aspect of something; it has to choose a very small subset of characteristics around which to build its representation.

The second component of the definition identifies the fact that a simulation is a very particular type of representation, what the authors call "operational." According to them, a simulation is a representation in the form of "an on-going process" instead of a static representation such as a diagram or flowchart. This component of the definition describes the systemic character of simulations. A simulation is a dynamic system: a set of parts that interrelate to form a whole. A simulation is therefore a procedural representation, one achieved through an ongoing process. In the case of games, the ongoing process is play.

Eddington, Addinall, and Percival's statement, "A simulation is an operating representation of central features of reality," offers quite an efficient little definition. In proposing our own definition, we would, however, like to make three small adjustments: first, in keeping with our system-based terminology, we replace the word "operational" with "procedural." Second, we generalize the idea of "central features" to "aspects" of reality. Third, we add quotation marks around the word reality. The result is the following definition of simulation:

A *simulation* is a procedural representation of aspects of reality.

Both components of this definition, the fact that simulations represent procedurally and that they depict elements of reality, represent surprisingly complex concepts. In the pages that follow, we look closely at these two aspects of simulations, considering each in turn.
Game and Non-Game Simulations

The general concept of a simulation is certainly not restricted to games. For example, economists and sociologists use simulations to study mathematical relationships among variables, often as a set of equations that process data. The data might be information from the U.S. Census—demographic information about income, housing, and voting patterns, for example—and the equations might spell out sets of relationships among data. Using this kind of simulation, a researcher would be able to speculate on changes in some of the variables, such as income and housing, input these changes into the simulation, and see how voting patterns would change as a result. This kind of simulation doesn’t seem like much fun to “play” (it certainly is not a game), but it does fulfill the requirements of a simulation. There are real-world referents represented in the simulation (economic and political realities) and the simulation functions by processing data through a set of procedures. The process might be merely a mathematical equation, but it is a process just the same.

Economic simulations are rarely explicitly interactive. Usually, a researcher sets up data and then “runs” the simulation to process the data. However, some simulations are designed to be highly participatory, such as training simulations. These include computer-based simulations allowing airplane pilots to practice flying without leaving the ground, live role-playing simulations that allow salespeople to refine their social skills on difficult clients, underwater simulations where astronauts practice zero-G maneuvers in a swimming pool, and emergency simulations in which the residents of an apartment building hold fire drills. Each simulation takes its identity from a real-world situation: flying a real plane, pitching a real sale, attempting a real space walk, or escaping from a real fire. In every case, the representation the simulation creates is a process: the complex machinery and interactivity of a flight simulator, the social and conversational process of role-playing, the physical simulation of being in outer space, and the flow of bodies down stairwells and along fire escapes.

Clearly there are many simulations that are not games. But what about game simulations? In the digital game industry, there is a genre of games called simulations, or “sims” for short. Sim City, for example, is a complex depiction of the process of urban planning, city economics, and the evolution of a human community; it is a simulation game. Other game simulations depict historical processes, natural ecosystems, or military vehicles. Although sims, perhaps more than other games, explicitly fulfill both components of our definition (a procedural representation of aspects of “reality”), all games can in fact be considered simulations. Remember that a game design schema must be applicable to all games. Therefore in proposing the schema Games as the Play of Simulation, we are arguing that any game can be considered a simulation.

As abstract or fantastical as games may be, it is possible to see them as simulations of one kind or another. Chess and Tic-Tac-Toe, for example, can be framed as representations of territorial conflict, in which simulated units war for control of a stylized battlefield. Games that involve fantastic elements, such as Dungeons & Dragons, also simulate through their play. Detailed rules, for example, simulate the way that different weapons impact different kinds of armor. Even the spell-casting system in D&D is a simulation of sorts: it simulates an imaginative “reality,”
one rooted in myth, religion, and popular culture. As the example
of extraterrestrial invasion in the quote from Eddington,
Addinall, and Percival illustrates, aspects of “reality” can refer to
things outside our lived experience.

Some games, such as Tetris, present more ambiguous referents,
but that does not mean that they are not a kind of simulation.
Tetris simulates the way objects can fall down, stack up, and
even make noises when they slide into place next to each other.
In this way, Tetris is a simple simulation of the forces of gravity.
Then there is the fact that falling Tetris objects are called
“bricks,” and these bricks form an interlocking brick wall grid.
These aspects of the game point to a different kind of depic-
tion, perhaps a simulation of construction. Tetris may not be a
particularly accurate or instructive simulation of gravity or con-
struction, but accuracy and instructiveness are not necessarily
what a game simulation is about. A game simulation, as any
kind of game representation, can be geometrically minimal,
outrageously whimsical, or even intentionally misleading.
Unlike a simulation designed for scientific research purposes,
a game simulation is not beholden to a notion of representing
empirical truth. Pong is not meaningful to players because it is
a scientifically accurate representation of Table Tennis; it is
meaningful because as a simulation it provides a context for
deep and engaging play.

Meaningful Play and Simulation
How does framing a game as a simulation assist in designing
meaningful play? In considering the play of simulation, we are
simply re-working fundamental game design concepts estab-
lished in previous chapters. Although the emphasis here is on
how games create representations, fundamental principles of
the design of meaningful play remain the same. Whether a sim-
ulation allows players to experience the representation of
something known and familiar or fantastically imaginative, it
does so through the design of meaningful play. In order to see
this principle at work, we can take a close look at Ace of Aces, a
game designed by history teacher Alfred Leonardi in 1980.

Ace of Aces simulates a dogfight between two World War I air-
craft, using a complex formal system to represent the speed,
maneuverability, visibility, weapons fire, and other aspects of
two-plane air combat. The striking thing about Ace of Aces is
that the game takes place not on a computer or even on a
paper wargame map, but instead inside two paperback game
books. Each player has his own book, and each of the more than
200 pages has an illustration of what the player sees from his or
her airplane cockpit. The point-of-view illustration shows the
enemy plane at a certain distance, location, and angle relative
to the player’s own plane. For example, if an illustration shows
your opponent’s plane coming towards you over your own tail,
it means the other player is directly behind you!

Players interact within the simulation by navigating through
their book (players cannot look into each other’s books) and
selecting maneuvers. At the bottom of each page is a list of the
possible maneuvers a player can take, with a number assigned
to each. Both players select a maneuver in secret and call out
the corresponding number, which determines the next page
each player turns to in their book. The elegant formal system of
the game is amazingly effective at simulating a dogfight
between two World War I airplanes.
Does this seem hard to believe? Consider the scenario we describe: your opponent positioned directly on your tail. You choose a maneuver to slow down and perform a weaving turn to the right, in which you shift your position to the side, ending up parallel to your previous position—something like a car changing lanes. Let's say your tailing opponent thought you were going to make a run for it and made a decision to move forward at top speed—this choice would cause your opponent's plane to zoom right by your decelerating plane. When you turn to the appropriate page in your book and see the outcome of last round's maneuvers, the illustration would show your opponent's plane ahead of you and to the left; and in your opponent's book your plane would be visible behind and to the right.

Through the use of a clever spatial model, Ace of Aces simulates aspects of World War I air combat. It does not simulate every facet of the experience (there are no rules to handle different kinds of weather and their effect on flying, for example), but it does represent important aspects of its referent. Spatial logic, tactical maneuvers, weapon jams, and even an increase in skill over several combats are all aspects of World War I air combat the game depicts. Furthermore, these representations are made possible through a dynamic system—a process based on a multifaceted mathematical model of air combat. It is through this process that Ace of Aces simulates a World War I dogfight.

Ace of Aces constructs this simulation by combining emergent and embedded elements. The drawings and pages themselves are fixed in print, and do not change as the game is played. In this sense, the book pages might be considered embedded narrative elements, pre-scripted narrative descriptors experienced by the player during play. However, the complexity of the underlying game rules incorporates these pages as elements within a truly emergent system. The Ace of Aces book pages are less like the pages of a Choose-Your-Own Adventure book and more like video frames from a real-time simulation display, snapshots of an ongoing battle. In creating a simulation, both emergent and embedded elements can be incorporated into the overall game. However, because of the way that simulations rely on dynamic systems, framing a game as a simulation tends to emphasize the emergent components of the game, the more purely systemic elements that interact in complex ways to generate unexpected results.

Ace of Aces not only provides a rich and coherent simulation of air-to-air combat, but also facilitates meaningful play. Because the pages of the two books contain all of the possible spatial relationships between planes (made possible by a set of rules), the players are literally navigating through the game's space of possibility, experimenting with maneuvers, taking daring risks, and psyching each other out. Each decision they make is both
discernable and integrated into the larger game experience, an experience made possible by the simulation. The representational mechanics of the simulation solidly support player decisions, establishing a taut and meaningful domain of interaction. The simulation creates a space of play that exists somewhere between the two printed books, the social interaction of the players, and the battle playing itself out in their overlapping imaginations.

Ace of Aces is a fascinating example of a game simulation, not just because it provides meaningful play, but because it does so through such unexpected means. Since the game was first published, real-time flying simulations on computers have become commonplace, used both for training and entertainment purposes. But Ace of Aces manages to engage players without illusionistic 3D graphics and sophisticated force-feedback pilot controls. Playing Ace of Aces is radically different than flying a plane, yet it somehow still manages to function as a successful simulation. Simulations do not need to literally embody the material and sensual forms of the phenomena they are simulating. This is what Robinett means when he calls a game "a simulation, a model, a metaphor." As representations, simulations often represent metaphorically, meaning they can create representations in non-literal ways. Sometimes, game simulations try and replicate the actual experience of the thing they are simulating, as with VR displays that take over a player's entire field of vision. More often, however, simulations take on modes of representation that are not so literal. There is an underlying mathematical model that connects Ace of Aces to planes moving through space. But the activity of playing the game—turning pages and calling out numbers—is nothing like sitting in an actual cockpit. In fact, this metaphorical difference between the core mechanic of Ace of Aces and its simulated referent is one source of the game's pleasure.

Procedural Representation

Seen as simulations, games are dynamic systems that construct representation through play. Asteroids, for example, represents the feeling of vast space through the inertial drift of the player's ship. The game designers could have created any navigational scheme they wanted, such as a space ship that could start and stop instantly and turn on a dime even when in motion. Instead, the player must maneuver the ship retro rocket-style, taking into account acceleration and momentum. Through this designed activity, the game expressively depicts deep space. In the Lord of the Rings Board Game, the dark force of Sauron is represented as a figure on a track that moves steadily toward the players; his evil nature manifest in his terrifyingly inevitable advance as well as in the deadly ramifications of an encounter with him on the board. The board game Up the River playfully recreates the experience of a flowing river through the unusual format of its board, which is made out of horizontal strips. Each turn, a player takes the strip from the rear and places it in the front. In this way, the sailboats belonging to the players must battle the steadily flowing water current as they race to be the first to reach the dock.

In these three examples, formal and experiential elements of the game work to create a representation that emerges out of the procedures of game play. We call this form of depiction procedural representation. The term "procedural" is shorthand for all of the process-based ways that games can signify. A procedural representation might arise from the functioning of a computer program's AI; it might be an emergent result of players following the rules of a game; or it could be an expressive core mechanic that references a particular action outside the game.

A miniatures-based wargame is a representation of war partly because the pieces themselves resemble miniature soldiers and because the battlefield can be painted to look like a contested landscape. But these visual signs make up only one part of the game's larger representation. Wargame representation is also procedural, created through the rules of the game and player choices that the rules engender. For example, units in a
wargame generally have a movement rating, representing the number of hexes or spaces through which the unit can move in a turn. In a typical wargame, a cavalry unit will have a higher movement rating than an infantry unit. This statistical difference between types of units is not only a formal distinction; it is a form of representation.

The fact that a cavalry unit moves more quickly in the game than an infantry unit is an act of signification that is fundamentally different than the visual aesthetics of the game token. Its representation is procedural, based on the unit’s formal identity and its interactive capabilities within the game system. Of course, the units in wargames have many other kinds of formal statistics as well, from offensive and defensive abilities to movement strengths and weaknesses on different types of terrain. All of these formal designations are based on the simulated characteristics of various battle units. As the game is played, these formal identities become systemic relationships that constitute a dynamic, procedural representation of war.

We can consider games as procedural representations on two levels. Borrowing concepts from Games as the Play of Meaning, we know that games are representations and also that they can represent. The idea that games are representations means that an entire game can serve to depict something. Considering procedural representation on this macro-level, a game represents as a whole. Pong is a procedural representation of Table Tennis; Tony Hawk’s Pro Skater 3 is a procedural representation of skateboarding. The notion that games can represent means that games contain smaller, internal depictions. This also holds true when considering the procedural aspect of game representation. We can look at the structure of a game on a micro-level and identify the procedural representations that make up the whole, such as the movement rules for cavalry and infantry wargame units.

Macro- and micro-level procedural representations can be embedded in each other. A card game that depicts social life in an eighteenth century royal court is a game that as a whole simulates a particular historical moment. But within the game, we find several micro-procedural representations. There might be one set of rules to simulate swordfight dueling and another to simulate the current political climate of the court. Of course, these micro-procedural representations are linked together within the complete system of the game that constitutes the overall simulation. The same is true of digital games, where the code that simulates light falling on 3D objects is generally separate from the code that simulates the behavior of computer opponents. Although they simulate different aspects of the game’s subject matter, these components are all contained within the larger macro-system of the game simulation.

Because procedural representations emerge out of the play of a game, the player’s participation is crucial in bringing the signifying procedures to life. As with all game representations, however, procedural representations also grow directly from the rules of the game, gaining meaning as players interact with them through play. Following are three examples of games that make very different use of procedural representation. In each case, the representation is brought to life through both the formal rule structure and the experience of play.

**Diplomacy**

The board game Diplomacy is a complex representation of World War I political negotiation. The game takes place on a map of Europe and (depending on the particular edition of the game) the tokens used by players might resemble land-based and naval military units, or they might be abstract shapes. Each player assumes the role of a European military power, vying to occupy a number of key cities and conquer the continent.

The game is played in turns. Each turn, players negotiate in public and private for a limited amount of time (usually about 10 or 15 minutes). At the end of the negotiation period, players write down and simultaneously reveal their selected actions. The outcomes of their actions that turn
are contingent on the decisions of other players. For example, during a turn one player moves an army into a territory occupied by another player. Support that the invading player has garnered from other players determines the success of the invasion. If the invasion pits one attacking unit against one occupying unit, the action is unsuccessful. However, if the invading player received support from another player with a unit in an adjacent region, the attack is successful: the strength of the invasion becomes two units acting against one. These rules make advancement of your armies on the board difficult, requiring players to make alliances and coordinate their actions.

The formal game mechanics of Diplomacy result in a dramatic procedural representation of negotiated alliances, uneasy agreements, and broken peace treaties: a representation, in other words, of diplomacy itself. Only one player can ultimately emerge as victor, and it is usually just a matter of time before deceit festers and player alliances are broken, reshuffled, and reformed. Which of your allies are going to betray you—and how? In the game of Diplomacy, as in the diplomatic processes it depicts, social skills are at least as important as strategic thinking.

Diplomacy as a whole procedurally represents the subject matter of diplomacy, and it does so through a number of internal representations that combine to form the overall simulation. For example, although Diplomacy could take place on an abstract map and still maintain the same sense of diplomatic intrigue, the map is also used as a means of procedural representation. Switzerland, for instance, is a central but impassable neutral territory, mirroring its isolationist role in World War I. Each country’s starting forces are a representation in miniature as well, appropriate to the time period: England has the strongest naval force but a weak army.

Diplomacy is a wonderfully engaging procedural representation of World War I political negotiation, but it only achieves status as a simulation because of careful design decisions. Procedures embodied in private player negotiations, simultaneous player action, contingent action > outcomes, the ability to support other players’ actions, and procedural use of the map and tokens combine to create a complex representation of diplomacy. This representation is a product of the process of play; a representation that only gains meaning when it is experienced as a system of dynamic relationships driven by player interaction.

Vampire
Vampire is a game that relies on procedural representation as well. The game comes from the New Games Book, which lists the following rules:

To start, everyone closes their eyes (Vampires only roam at night) and begins to mill around. You can trust the Referee to keep you from colliding with anything but warm living flesh. However, you can’t trust him to protect you from the consequences, for he is going to surreptitiously notify one of you that you are the vampire.

Like everyone else, the vampire keeps her eyes closed, but when she bumps into someone else, there’s a difference. She snatches him and lets out a blood-curdling scream. He, no doubt, does the same... 

If you are a victim of the vampire, you become a vampire as well. Once you’ve regained your composure, you too are on the prowl, seeking new victims. Now perhaps you are thinking that the game too quickly degenerates into an all-monster convention? Ah, but then you didn’t know that when two vampires feast on each other, they transform back into bread-and-butter mortals.

Will the vampires neutralize each other before all mortals are tainted by the blood-sucking scourge? Why don’t you try a little experiment and see? There’s always hope, even in the midst of a blood-curdled crowd?
The referent that Vampire simulates is quite different than the referent of Diplomacy. Whereas Diplomacy procedurally represents a real-world historical situation, Vampire comically evokes images of vampires ripped straight from the pages of pulp fiction. The way the game design achieves its procedural representation, however, is no less sophisticated.

Because Vampire requires no game materials (no map or game pieces), it relies entirely on the activity of the players' bodies to generate play: game representations emerge solely from the interaction between players. The initial limitation on the game, the fact that players must keep their eyes closed for the duration of play, orchestrates a certain kind of representational experience. Enclosed in darkness, the player is taken out of the ordinary world and placed in the imagined world of the vampire night, a setting whose drama is amplified by the fact that players spend the entire game stumbling through an unfamiliar space, feeling around for each other.

This tension-filled core game mechanic makes every meeting between players a surprise. There are three different ways that an encounter can play out, from two non-vampires exchanging thankful sighs of relief to a screaming vampire attack, (or double-attack, if two vampires collide). As players wander aimlessly, the sound of shrill yells map the darkness surrounding the players, transforming their invisible game world into a screamingly theatrical sonic landscape. In addition to supporting the goofy-horror flavor of the game, the sound component allows players to "see" the larger playfield, signaling areas of action. Are those bloodcurdling yells coming from somewhere safely far away? Or perhaps from RIGHT BEHIND YOU! These simple procedures (wandering around with eyes closed, meaningful chance encounters, screams erupting in the darkness) together create a coherent and distinctive vampire experience for every player in the game.

Furthermore, as with many New Games games, Vampire plays with the conventions of winning and losing. Does an individual player win at Vampire? How? Do players try to become bitten? Or do they try and remain safe? As vampires, are players seeking new victims? Or a cure for their condition, in the form of another vampire? How will the game end? With a field full of vampires or with the group purged of their vampiric tendencies? The emergent representational system of the game, in which players collide like charged particles, takes on unpredictable patterns. The fact that the game can end in one of two states heightens the drama of the experience and gives the overall game play the suspense of a thriller.

Vampire creates a representation through a number of procedures, from the game's immediate core mechanics to the long-term trajectory of play. Vampire is successful as a representation because it is squarely focused on the player experience. From the outside, Vampire looks like a silly group activity. It is only inside the game that the procedural representation works its magic.

Illuminati
We first mentioned the board game Illuminati in Breaking the Rules. The game design is remarkable, among other reasons, because it offers a play variant that encourages rule-breaking within certain boundaries. Although Illuminati uses many means to simulate its subject matter, we focus here on how rule-breaking itself acts to create a procedural representation.

Illuminati is based on the Illuminatus books of Robert Anton Wilson, and the subculture of conspiracy theory associated with them. Players in Illuminati assume the roles of shadowy power brokers that manipulate world events and political structures to their own devious ends. There are groups that players can control and link in the
Represented Conflict

We are beginning to understand how procedural representations work to simulate phenomena through dynamic depictions. But there is a question that precedes a discussion of how games create such representations. It is the question of what phenomena a dynamic system can depict. Can a game designer pick anything to simulate, or are there inherent limitations? Are there certain things that games are predisposed to simulate, certain subjects that lend themselves naturally to games? Game designer Warren Robinett seems to think that just about anything might be simulated:

Many provocatively complex phenomena await interpretation...trains and other vehicles which move cargo through spaces, kayaks in swirling river currents, planets orbiting their stars, competing creatures in evolving ecologies, visible melodies smeared upon harmonic wallpaper, looping programs in throbbing execution, and human thought darting across a tangled network of knowledge...

The real world offers an vast set of phenomena to simulate—animals behaving, plants growing, structures buckling, traffic jamming, snowflakes forming. Any process is a candidate. Every verb in the dictionary suggests an idea.3

Since Robinett originally penned this challenge, games have been designed to simulate some of the phenomena he describes: Sim Life attempted to simulate evolving ecologies of creatures; the shareware game Bridge Builder simulates structures buckling under the weight of a train. However, many of the phenomena on his list are still waiting to find themselves in games. As Robinett points out, "every verb in the dictionary suggests an idea" for a simulation. Why then, do games seem to focus on a narrow range of processes to simulate? Why do we see the same genres of games over and over: fighting, racing, war, sports, and so on? Of course, economic and business concerns greatly influence game content. But is there something else, something deeper about the underlying structure of games that determines the kinds of processes they can and cannot depict?
Our definition of a game describes them as systems in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome. The part of the definition relevant to our present discussion of simulation and representation is conflict. Games are contests of power: they are systems of conflict. Conflict is not only a product of the game's rules, but of its system of representation as well. Every game, on some level, dynamically represents conflict. The elements of a game—the players, the pieces, the rules— all have a role in generating the representation. The insight that games represent conflict through a dynamic process might help to explain the prevalence of certain content in games: perhaps some forms of conflict are simply easier to model than others. At the same time, understanding the kinds of conflict that games most often depict also helps us to strategize new kinds of subjects for games to simulate.

What are the forms of conflict we find dynamically represented in games? If the game has a strong narrative component, the conflict is easy to spot. The Lord of the Rings Board Game clearly simulates the struggle of the players, as the Fellowship hobbits, to reach Mount Doom at Mordor and destroy the One Ring. But in many games, it is more difficult to pin down the simulated conflict. What is the conflict in Baseball, Checkers, or Jeopardy? The key to comprehending the form of conflict simulated by a game is to figure out what is being contested. In what kind of arena is the conflict being held? Over what is the conflict being waged? How is the progress of the conflict measured? What aspects of the conflict are dynamically represented?

In order to answer these questions we distill the range of game conflict into three general categories: territorial conflict, economic conflict, and conflict over knowledge. These three categories are neither discrete nor mutually exclusive: many games incorporate two or all three of them at once within their design. Rather than a strict typology, they are instead conceptual frames for looking at the kinds of conflict that games can dynamically represent. Next, we explore each of these three categories in more detail.

Conflict Over Territory

Conflict over territory is perhaps the most intuitive of the three categories. Board games such as Chess, in which pieces are moved on a limited playing field, are a common game of this sort. In games of territorial conflict, players strategically position their units to capture enemies and gain ground. Conflicts of this kind are abstract representations of war: the pieces depict military units, and the play area dynamically represents the territory over which the battle is waged.

Go is another good example of a game focused on capturing territory. As players lay down their stones, their primary goal is to surround areas of the playfield to secure the captured space. At the end of the game, each player receives a point for each grid intersection secured (plus a point for each captured enemy piece). The game originated as a military simulation—in feudal Japan, Go was considered a martial art. As a territorial conflict, Go is a strikingly elegant representation.

There are many other games that simulate the process of territorial conflict. Tic-Tac-Toe is a simple territorial conflict where players attempt to strategically occupy territory in a pattern that will lead to victory. Ball-based sports such as Football and Soccer entail moving a team or a special marker across a stretch of terrain into the opponent’s end zone or goal: the enemy invaded. Tabletop games such as Warhammer offer incredibly complex representations of warfare dynamically enacted, with dozens of different kinds of units, large detailed maps, and thick rulebooks controlling the particulars of interaction. The U.S. military uses even more complex war games as training exercises, in which hundreds or even thousands of troops play vast games of laser tag in real and simulated environments.
Economic Conflict

Economic conflict is another common form of conflict in games. Within simulations of economic conflict, it is not terrain that is contested, but a unit of value. The word "economic" does not necessarily refer to money, but to any collection of pieces, parts, points, cards, or other items that form a system through which the conflict takes place. In a pinball game, you are trying to rack up a high score. In Magic: The Gathering, you are trying to reduce your opponent's life to zero. In these game economies, the rules give each unit a value, and progress through the game is measured according to the values assigned by this economy.

An economy in a game is generally a limited economy. This means that the units that make up the economy are finite, and usually the players know the composition of the economy. In Poker, it is crucial that all players understand the limited economy of a deck of playing cards. Knowledge about which cards appear in the deck allows them to understand which hands are more difficult to build. Four-of-a-kind is harder to build than a pair; a straight flush harder still. The other economy of Poker—the betting money—might or might not be a limited economy. Each player might start with the same amount of chips, in which case all players know the parameters of the chip economy. If players can use money in their pockets or other valuables for betting, the players don't know the full extent of the economy—although the economy is ultimately limited by the capital each player possesses outside the magic circle. On the other hand, if players are not betting "real money" but are instead playing for fun using an endless supply of chips, the normally limited betting economy becomes unlimited.

Because economic conflict is generally reducible to numbers and points, and games are intrinsically mathematical, we can frame almost any game in this way. For example, a race game, in which players roll a die and move a marker down a track, might at first seem to be a territorial conflict. However, the same game could also be played by throwing dice and adding up the points that players receive each turn, making the game more of an economic conflict. Since the two games would have similar constitutive rules, the operational rules would help determine what kind of conflict the game represents. Yet some games combine categories: Is Quake a territorial conflict or an economic one? It is clearly a hybrid. The play takes place within the representation of a space, in which relative position at each moment is quite important. However, much of the game consists of managing economies of resources such as health, armor, ammo, weapons, and kills.

Even the strongly territorial games of Chess and Go can be seen as procedural representations of economic conflict. In Chess, the pieces represent an economy, and the use-value of each piece is derived from the total set of relationships on the board. Of course, one unit—the King—has a special value, which determines the winner of the game. Similarly, at the end of a game of Go, territory is translated into points, and as with the race game example, Go could be interpreted as an economy—of contested points. Remember that the three kinds of conflict are not hard and fast categories; they are merely frames we use to understand the kinds of conflict that games traditionally simulate.

Conflict Over Knowledge

Conflict over knowledge offers a different model for understanding the way games simulate conflict. In Trivial Pursuit, for example, it is true that pieces move about on the spatial territory of a board. It is also true that the players acquire a set of colored plastic pieces within an economy of parts in order to win the game. But these ways of framing Trivial Pursuit seem to leave out the key component of the game conflict: the process of asking and answering trivia questions.
In Trivial Pursuit, as with many other games in which information itself forms the arena of conflict, the contested "terrain" of the game is knowledge. Game shows such as Hollywood Squares, computer trivia games such as You Don't Know Jack, and even games about translation of information from one form to another such as Charades, can all be understood as games in which the conflict is one of knowledge. Conflicts over knowledge are inherently cultural, because the game conflict itself engages with a cultural space that lies outside the game. In a game of conflict over knowledge, the outcome of a game action is dependent on whether or not the player knows the right answer to a question of some kind. This is quite different than representation of territorial or economic conflict: the process being simulated is the conflict of acquiring and sharing cultural knowledge. Games designed with factual knowledge as part of the system of conflict cross over the border of the magic circle, creating a game contingent on information brought into the game from external sources.

Games represent conflict as acquisition of and contestation over territory, economy, and knowledge. These three rather abstract categories don't tell us exactly what games are capable of simulating, but describe the general sorts of processes that games most often simulate. Identifying these three categories also helps explain why we see the same kinds of conflict being modeled over and over in games. For example, why is it that video games often seem to focus on simulating military conflict: fighting, shooting, and conquering? Or that so many games overflow with collectable item economies: magic coins, money, or other precious objects? Like it or not, the tendency toward military and economic representation in games has a long history, directly linked to the processes of territorial and economic conflict intrinsic to most games.

There is a relatively clear line of descent, for example, from Go and Chess to Kriegspiel, wargaming miniatures, and role-playing games, and from these non-digital games to today's RPGs, FPSs, and RTSs (role-playing games, first-person shooters, and real-time strategy games). A tremendous amount of design thinking regarding wargaming, military simulation, and other forms of territorial conflict has accumulated over the centuries. Simulating the difference between mounted units and infantry units; between melee and ranged weapons; between attacks that spread damage and attacks that penetrate; between size and maneuverability, strength and speed, and so on, have become well-worn design problems of game representation over the years. In this sense, today's highly detailed military games are the inheritors of millennia of design thinking.

Happily, this long history in no way limits what it is possible to simulate in games, even when it comes to forms of conflict. An important question for today's game designer is: What other kinds of conflict can games simulate? For example, what about Robinett's wish list? How could a game be designed to simulate social conflict, psychological conflict, or interpersonal conflict? These are truly tough design challenges. As we will see in the following pages, part of the challenge lies in the fact that simulations require radical simplification and stylization. Sid Meier's Civilization series is wonderful strategy games that tackle the Herculean task of simulating cultural development. But because cultural knowledge in the game is necessarily stylized into abstract units ("Do you trade Monotheism for Iron Working?") the game never comes close to representing the subtlety of its subject matter.

The history of games contains many robust examples for simulating military and economic conflict. A design lexicon for simulating social or cultural conflict may take generations to develop. Of course, these unsolved challenges are part of what makes game design as a field so remarkable. Despite the fact that games are a truly ancient phenomenon, there are still countless avenues for representational innovation—as long as you are ready to question long-standing assumptions about what games are and what they can be.
Procedural Characters

Conflict is an abstract, elemental way of thinking about the kinds of processes that games simulate. But it is not the only way to frame games as simulations. What about simulation and storytelling? Any of the game narratives discussed in the previous chapter could be thought of as a simulation, providing it was represented through a dynamic process. Combining narrative and simulation is a powerful way of thinking about games as a representational medium, because it forces a truly experiential approach to participating with a story. As Reiner Knizia wrote in his earlier essay for this book, his hope for the Lord of the Rings Board Game was that it would “not just re-tell Tolkien’s plot, but more importantly it would make the players feel the emotional circumstances of the story.”

Following are four examples of just one part of the storytelling equation, the element of character. The four examples each examine a very different strategy for procedurally creating narrative experience via character (here we are using a general concept of “character” that refers to a fictional persona contained within a game representation). Some of the examples are characters under the direct control of a player, whereas others remain outside of player control. In all cases, rules and interaction are used to procedurally construct a character, while also weaving the character into the larger fabric of the game representation.

Zelda: Link’s Awakening

In this adventure game for the Game Boy, players control a character named Link, moving him about the fictional world of the game, exploring new spaces, acquiring objects, and defeating enemies. The game is rich with characters—in addition to Link, many personalities populate the world of the game, including a witch character that players encounter early in the game.

The character of the witch signifies the idea of “witch” in many different ways. The character looks like a witch (she wears a tall pointed hat), sits next to a cauldron, cackles as she talks, and possesses other stereotypical trappings of a cartoon witch. In addition to these non-procedural representations, depiction of the witch occurs procedurally as well: the witch character has a number of systemic qualities that allow her to signify in ways that a non-procedural character could not.

For example, like many characters encountered in Zelda: Link’s Awakening, the witch’s character lives in her own house. Most houses are found in villages, where many of them are clumped together. As a result, they are easy to locate. But the witch’s house lies deep in the heart of a dangerous wood. In order to reach it, the player must overcome obstacles as he or she searches through the maze-like forest. Not only is the witch represented in the space of the game as living a life isolated from the villagers, but her very separation from society makes her character more difficult to find. Both of these attributes (isolated and dangerous to visit) are very “witchy” characteristics. In the game, these characteristics are a function of the position of the witch’s house relative to other elements on the game-grid of the imaginary world: they are procedural, growing from the formal characteristics of the game and the player’s own game interactions.

When a player reaches the witch, she informs Link (via on-screen text) to bring her a mushroom. If the player goes back out into the forest, collects a mushroom, and then brings it back to the witch, she will take the mushroom, stir her cauldron, and produce a magical powder for the player. The witch enacts a procedure that transforms the mushroom, changing its properties within the game. As a liminal character living outside the bounds of society, the witch has the ability to convert natural objects to useful technologies. This kind of procedural characteristic is common in adventure games (take object A to character X and receive object B). Yet in the case of the witch, it works very well to depict her character. The procedural elements constituting
her representation (her location in the world, her ability to convert objects) are truly witch-like and serve to create an effective and memorable character.

One last note of clarification: imagine a witch character in a book similar to the witch in Zelda: Link’s Awakening. The book’s story would describe how the witch lives in the woods, makes magical potions, and so on. What would be the difference between the two witches? Aren’t they basically the same character? Perhaps, but the form of their representation is radically different. Although both witches might have the same “literary” characteristics, the witch in the book would not possess these traits as procedural qualities, which are triggered as part of an activity of meaningful play. In fact, the witch in the story is exactly the kind of witch that the game witch references. Actively exploiting the witch’s witch-like qualities, not just by reading about them but also by playing with them, is what makes her representation so powerful to experience in the context of the game.

**Virtua Fighter 4**

The fighting game Virtua Fighter 4 integrates procedural representation into the narrative play of the game by tying the formal characteristics of the fighting characters to their appearance, personality, and embedded backstory. Every character is designed with explicit strengths and weaknesses, which are procedurally represented through character attributes. For example, Pai Chan has incredible speed but lacks power and hard-hitting moves. Kage-Maru has complex combinations and special moves, but they take time to execute, leaving him vulnerable to attacks from his opponent. Jeffry McWild has great power, but is large and heavy, and therefore less agile. Each character’s strengths are countered by logical weaknesses, adding up to a fighting “personality” that plays itself out during each match.

In a well-designed fighting game these procedural representations have strong ties to the fictive world and narrative of the game. How a character fights is usually an external representation of internal qualities, and fighting styles are mirrored in the narrative histories provided for each character, as well as more mundane information such as height and weight, profession, gender, and country of birth. All of these non-procedural narrative descriptors may seem superficial to the game play, but they help create an integrated character in which procedural and non-procedural elements are brought together in a character representation that players experience on many levels.

For example, Virtua Fighter 4’s Pai Chan is an action film star whose hobby is dancing. She is small and light and favors a fighting style that uses combos and quick reverses. Her father, Lau, bested her in a previous tournament, and in VF4 Pai is determined to defeat him and prove herself a worthy successor to her father’s legacy. Each component of this narrative offers players insight into the strengths and weaknesses of the character. Pai Chan’s background as an action hero and dancer make her quick and agile. Her preferred style of fighting favors rapid combinations that often leave her defenseless if her attacks are blocked. Because of her intense desire to defeat her father, Pai Chan is driven by emotion, not logic. As a result, her fighting style is visceral and immediate. She reacts quickly, but rarely plans for the long fight.

This character sketch not only describes Pai Chan herself, but also the playing style that her procedural characteristics engender. By “playing” Pai Chan in a match, the player participates in her representation, bringing her procedural characteristics to life. The play becomes a meaningful representation because of the well-designed synergy between her formal characteristics, appearance, backstory, and emergent personality. This kind of richly layered simula-
tion is a fantastic example of the unique ways that games signify through an integrated suite of representational mechanisms.

Deus Ex
Deus Ex is a role-playing game designed for both computer and console platforms. A strong feature of the game is the computer-generated characters, which respond to game events in surprisingly subtle and expressive ways. These characters follow a series of AI algorithms that determine how they behave in any given situation. In *Swords and Circuitry*, Hallford and Halford reprint a section of the Deus Ex design document that outlines how different character types act in certain circumstances. This fascinating document reveals the designed mechanisms by which the game’s characters are procedurally represented. Here is a brief excerpt from the document:

Civilian:
- Does not harm civilians
- Ignores unidentified sounds
- Aware of alarms
- Issues warning before attacking
- Flees when wounded below X% (where X is high)
- Tends to protect self
- Ground-based movement, normal

Thug:
- No concern for safety of civilians
- Investigates unidentified sounds
- Aware of alarms
- Attacks without warning
- Flees when wounded below X% (where X is low)
- Ground-based movement, normal

Military:
- Does not harm civilians
- Investigates unidentified sounds (if possible without abandoning post)
- Aware of alarms
- Issues warning before attacking
- Never flees when wounded
- Ground-based movement, fast

These character descriptions are quite high level—they do not formally specify exactly how each character will behave. For example, the speed of each character is only defined as “normal” and “fast” rather than numerically. However, as an abstract character sketch, these descriptions offer a snapshot of the sort of design decisions the game’s designers made as they developed the characters’ behaviors. Compare the thug character and the military character. Whereas the more cowardly thug will flee when his health is low, the brave military character will never abandon his post and will fight it out to the bitter end. The unscrupulous thug will attack without warning, but the honorable military character will issue a warning before attacking. In this way, the programmed characteristics of the characters take on a simulated personality, becoming expressive by virtue of their procedural differences. By adding different types of characters defined along a number of parameters, an entire cast of procedurally generated actors could be developed.

To appreciate the sophistication of these strategies for representation, compare the Deus Ex characters to the witch in Zelda. The witch behavior is fairly simple and wholly predictable (if: receive mushroom, then: create powder). In contrast, several Deus Ex characters encountering each other in the game create a scene rife with emergent drama.
Imagine, for example, the following situation: A thug approaches a civilian. The civilian (who ignores unidentified sounds) pays no heed to the footsteps of the approaching thug, who begins to mercilessly beat the civilian (attacks without warning; no concern for safety of civilians). After the thug strikes a single blow, the civilian starts to flee (flees when wounded below X%, where X is high), pursued by the thug (both characters move at normal speed). The pair of them run by a military character who takes notice of them (investigates unidentified sounds) and takes up the chase herself, quickly catching up to them (moving at fast speed). The military character ignores the civilian but catches up to the thug and begins to issue a warning... How will the scene play out? Does the thug pause when warned, giving the civilian time to flee? Do the military character and the thug battle to the finish? Or does the entire scene disperse as characters each go their separate ways?

This sequence of actions is not pre-scripted, but instead emerges from the simulated behaviors of the characters. Each character has a very distinct personality. Of course each character also has a different visual appearance, style of movement, tone of voice, and so on. These non-procedural traits are certainly important to their overall representation. But it is through the procedural representation of the characters, representation emerging from behavioral characteristics, that they take on active roles in the dramatic events of the game.

Blob

Procedurally represented characters are not exclusive to the domain of digital games. Our final example of a procedural character comes from a non-digital game, a New Games tag variant called Blob.

If you're addicted to late-night TV monster movies, here's a sure way to kick the habit and break out into the light of day...

The Blob begins innocently enough as a mere individual playing a game of tag. As soon as she catches someone, she joins hands with him. Now he's part of the Blob too, and they both set out, hand-in-hand, in search of victims. Everyone the Blob catches (only the outside hand on either end of the Blob can snatch at players) joins hands with it and becomes part of the lengthening protoplasmic chain. And thus the insidious Blob keeps growing.

Unlike your run-of-the-mill, mad scientist-created Blobs, this one is not content to merely ooze along, seeking its prey. It gallops around the field, cornering stray runners and forcing them to join up...

Moreover (horrors), the Blob can split itself into parts and, with its superior communal intelligence, organize raiding parties on the lone few who have managed to escape. The thrilling climax occurs when there's only one player left to put up a heroic last-ditch stand on behalf of humanity. But alas, there is no defense against the Blob, and humanity succumbs. (If that seems unfair, well, that's the plot.)

The moral of our story could well be: "You become what you fear." If you have the heart to destroy humanity again, you can have the last person caught start the Blob for the next game.

The game of Blob is centrally focused on the procedural representation of a single character: the Blob. The form that the character takes, a mass of moving bodies, is quite different than the characters in Zelda: Link's Awakening, Virtua Fighter 4, and Deus Ex, which are experienced as visual images and audio. Yet, like these characters, the Blob is generated out of a set of representational procedures.

The Blob parodies a B-grade movie monster: a humongous, horrifying, pudding-like creature. The rules of the game cleverly bring this character to life through a set of behavioral procedures for representation. The fact that the rest of the players try to avoid the Blob immediately creates an environment of fear. The slow-moving Blob scatters players before it, lumbering through the playfield. The
touch of the Blob is deadly, and when a player is brushed by one of the edges of the Blob, that player is ingested and incorporated into the body of the character.

As the Blob grows, it tends to move more slowly while covering a wider area; it can also fragment and recombine in a very protoplasmic way. The size of the Blob is an inverse function of the number of players running loose around the playfield. The dwindling non-Blob players become a community of hard-nosed survivors. Oh no! Don't tell me it got Sharyn too! At the game's climax, the Blob symbolically ingests the entire world, becoming synonymous with the group of players as they reach a competitively and cooperatively achieved endpoint. The game's narrative may only have one ending, but as the rules point out, there is a moral to the story.

Just as the witch in Zelda: Link's Awakening was represented through procedures that created distinctively "witchy" characteristics, the character of the Blob possesses procedural character traits (fearful, steadily growing, ingests players, fragments and recombines, inevitably wins) that are exceedingly Blob-like. An elegantly designed game persona, the Blob generates a character through exceptionally effective procedural means—and completely without digital technology!

Designing Simulations
Procedural representations clearly can provide meaningful play for game players. But designing simulations is challenging. Once you decide what it is you want your game to simulate, how do you put the pieces together to arrive at the kind of simulation that you want? A simulation, as a representational process, is more than a series of independent procedures producing a result. A simulation arises from the operation of a system in which every element contributes in an integrated way to the larger representation.

Designing the mechanisms of that system presents many challenging decisions for a game designer. In Inventing the Adventure Game, Warren Robinett outlines some of the key design issues involved in creating a simulation:

Given a phenomena to simulate, the problem is to decide what are its parts, how these parts can be represented with numerical values, and what the relationships are that let these parts affect one another....

Making a simulation is a process of abstracting—of selecting which entities and which properties from a complex real phenomena to use in the simulation program. For example, to simulate a bouncing ball, the ball's position is important but its melting point probably isn't. Any model has limitations, and is not a complete representation of reality.

In these few sentences, Robinett makes a number of very important points about the design of simulations, including:

- **Simulations are abstractions.** The real or imagined phenomenon you want to depict in your game is most likely overflowing with layers of detail. But as with all forms of representational media, you will never be able to fully represent every facet of your subject. Thus your simulated representation, as Robinett points out, is an abstraction. Chess is a highly abstracted representation of war. Sim City is a very stylized version of government and city planning. D&D even abstracts people—into the characteristics of Strength, Intelligence, Wisdom, Dexterity, Constitution, and Charisma. A simulation does not attempt to simulate every aspect of its referent, but instead focuses on those elements necessary to the game. Virtua Fighter 4 simulates the fighting capabilities of its characters: it does not simulate their biological immune systems or taste in classical music, since these are not relevant to the play of the game.

Being able to select which components of your subject to ignore and which to retain and abstract is an absolutely critical game design skill.
Simulations are systems. A simulation is a whole made up of smaller, interrelated parts. As with any complex system, meaning emerges from the interaction of the parts. Brainstorming a list of attributes or effects that you want to include in a simulation is not enough. You must conceive of a system that incorporates them all. You might want a fighting game that can simulate the difference between a fast but weak character and a slow but strong character. But what does “fast,” “slow,” “strong,” and “weak” mean in your system? How do they interrelate? How do these attributes affect the decisions and outcomes a player makes? All games are systems, but when we frame them as simulations, the systemic aspect of the game is harnessed for representational effect.

Simulations are numerical. Not only are simulations abstracted, systemic representations, but they are also reducible to a formal, numerical structure. We know this already about games: at some level games are composed of rules, and at their most formal level, all rules are logical, mathematical, constitutive rules. The six D&D statistics listed previously are represented in the game as numerical digits between three and eighteen. Complex physics simulations in computer games are based on mathematical modeling. The behavior of artificial characters, whether on a Magic: The Gathering card or inside Deus Ex, can be reduced to their formal identity. The fact that simulations must reduce their subjects to formal, numerical values is exactly why it is so challenging to procedurally depict social, psychological, and other experientially complex phenomena in a game.

Simulations are limited. Because simulations are numerical abstractions, they are intrinsically limited. As Robinett points out, “every model has its limitations and is not a complete representation of reality.” We emphasize this aspect of simulations because of prevalent ideas in the computer game industry that more complex simulations automatically guarantee meaningful play. In fact, on a digital platform even a supposedly “realistic” simulation only depicts a tiny slice of any real world or imagined phenomenon. But this doesn’t mean that simulations can’t provide meaningful play. The inventive shareware game Bridge Builder simulates its subject, but it chooses a very narrow aspect of bridge building—the engineering challenges of the support structure. This design leaves out thousands of other possible characteristics that might be simulated, from the aesthetics of the building materials to the effect of the bridge construction on the surrounding ecosystem. But that’s OK. The game turns the intrinsically limited nature of simulation into an asset, by focusing player activity on a fun and educational aspect of building bridges.

Design involves choice: to create a simulation, you need to decide what to simulate and how. Every choice you make as a game designer opens some possibilities and closes others. What is meaningful in the context of a particular simulation? Is it meaningful to blow wind into the face of the player as she is piloting a hang glider? Is it meaningful to provide a full-body harness in which the player can lie as she interacts with the simulation? Is it meaningful to simulate the insurance and legal procedures by which a player purchases or rents a hang glider? Pilotwings for the Nintendo Entertainment System, a popular game simulation of hang gliding, includes none of these features.

In digital games, much of this decision-making process involves the scope and depth of a simulation. If a racing game is composed of a single car on a single track, it can be extremely detailed. It might include a complex set of physics models, simulations of the internal suspension of the car, or wear on tire treads as they are used over time. Given the same design resources, the addition of more cars and more tracks means that fewer characteristics can be simulated in an equally detailed way. If even more elements are added—such as cars that can transform into jets and fly around the track—the focus
of the simulation shifts once more. If a character can get out of the vehicle, walk around, and interact with other people, that casts the net even wider and the "depth" of the simulation decreases accordingly.

In designing a simulation, you must decide exactly what kinds of procedural representations you want to provide for players. In a fighting game series such as Tekken, a detailed (if fancifully cinematic) fighting simulation, the characters don't also get into cars and drive around a track. On the other hand, in an ambitiously open-ended game such as Shenmue, a player's character can have simple conversations with many other characters in the environment, examine, carry, and utilize a wide array of objects, and explore a large detailed space. As a result of the range of activities simulated, the fighting system in Shenmue is much more stylized and limited in scope than that of Tekken.

Why is it that games can't simulate everything with a high degree of detail? Why can't a game simulation be both wide and deep? There are several reasons. Limited development resources require that game designers decide where those resources will be spent. But the limitations of time and budget are not the only things affecting the scope of simulations. There are game design factors as well. Meaningful play stems from the ability of players to make meaningful choices from a limited set of knowable options. If a player has trouble recognizing everything that is being simulated, an understanding of knowable options decreases.

In an essay by game designer Harvey Smith on simulation and games, he uses the fictional example of a vehicle-based game with terrain simulation so exacting that the geometry of a player-driven truck can get stuck on a tiny bump on the ground. In this case, the designer chose the wrong element of the game to simulate in detail. Smith's example also points out the problem of thinking that a simulation is anything but an abstraction. The "reality" of a game is determined by the mean-

nings it creates within the magic circle. The terrain simulation in Smith's fictional example might be based on scientifically accurate mathematical models, but the only thing the player will experience is the frustrating, "unrealistic" experience of being unable to drive on what looks like relatively smooth terrain.

The proper scope of detail for a simulation is largely determined by expectations set by the broader context of the game. In Gran Turismo, a game that deeply simulates real-world car physics, players come to expect a finely grained driving experience. However, driving is clearly the focus of the game. No player expects to exit their car, wander up into the stands, and interact with spectators. A game such as Shenmue, on the other hand, has been criticized for disappointing players. If players can interact with many different kinds of objects in the game, why can they enter only some of the buildings and not others? Player expectations are raised to unrealistic levels: the implied breadth of the simulation is far greater than what the game actually delivers.

Given that simulations are abstract and limited, as a game designer you must choose your battles wisely. The elements you select to depict through the procedural representation of a simulation determine the experiential focus of your game. Another of Harvey Smith's thought experiments is to take a typical driving game and give it additional depth by having computer opponents take note of their fuel consumption and try to drive directly behind other cars, strategically using wind shear to conserve fuel. He contrasts this to a driving game with emotional simulation, in which one of the computer opponents might drive recklessly during a race because he had just ended a relationship with a girlfriend.

Whereas Smith prefers the fuel consumption adjustment to the driving game, we find both of the game ideas equally interesting. Each of the two scenarios points to a very different game experience. Both driving games—one a hardcore strategic sim and the other a romance set on a race course—would require
simulating different kinds of phenomena. These procedural representations would be part of the larger game system and would determine what the game could represent and what the player would experience. You most likely wouldn't want both the fuel consumption simulation and the romance simulation in the same game. Why? Because each pull the space of possibility and the focus of the play in opposite directions. The fuel consumption feature implies an entire system of detailed car simulation mechanics that would be the central focus of the game. The broken hearted driver implies a game system that would focus on simulating dating, emotions, and stylized romance. As you craft representations in your game system, you simultaneously create the meanings that players will experience.

The key is to remember that just because a simulation is limited in scope doesn't mean that it is impoverished in what it can provide players. The abstract play of Go contains an infinite number of strategic options. The fine-grained driving focus of Gran Turismo supports meaningful exploration because the simulation rewards players for learning about and taking advantage of its subtleties. And the broad-but-shallow world simulation of Shenmue lets players focus on the rich narrative surface of the world without getting tripped up in interactive complexities that would not be appropriate to the game. The creation of a simulation is the creation of a space of possibility. By defining the exact nature of your game's simulation, you are sculpting the shape of your game's meaningful play.

**Learning from Wargames**

Working within the intrinsic limitations of simulations is one of the key challenges of game design. What are you going to simulate in your game? How are you going to abstract it? Which features of the phenomena will you include and which will you ignore? How deep and how broad can your simulation be? How do you tie each aspect of your simulation to the larger player experience? To understand these kinds of design decisions more fully, we look in detail at a particular genre of game, the historical wargame.

Historical wargames are complex strategy games that use cardboard chits or metal figures on a map to simulate a battle. We have already noted that game simulations are not universally beholden to "realism" or "accuracy." But historical wargaming is a genre of game design where both realism and accuracy are important. Historical wargame designers base their troop composition, map layouts, and game rules on historical research, a numerical approach to military history that wargame designer James Dunnigan calls "analytic history." In the game design sub-discipline of historical wargaming, part of the design ethos is that a game accurately simulates historical circumstances.

History, in a very general sense, represents a fixed series of events. But a historical wargame is a game, which means that uncertainty, risk, and unpredictable outcomes play a role. What a historical wargame really simulates are the starting conditions of a conflict. The way that the conflict plays out is what makes the game interesting as a game experience. Will history repeat itself? Was the historical outcome inevitable? How much can masterful strategy affect the outcome? These are all questions that wargame designers and players seek to answer through the creation and play of their games. The meaningful play of a historical wargame derives not only from the strategic complexities of military decision making, but also from the fidelity of the game to its historical referent.

As we know, a simulation can never contain every possible aspect of the phenomena being simulated. Historical wargaming has been wrestling with this challenge for at least a century, making it a wonderful case study for the design of simulations. We have already touched on one aspect of wargame design, the abstraction of unit characteristics. The pieces in a war-game are far more complex than in a game like Chess. When a wargame unit "attacks" an enemy piece, the outcome of the simulated combat is not simply to remove the attacked unit (as in Chess); instead, a variety of factors determine the outcome. Resolving an attack might involve some or all of the following:
- offensive strength of the attacking unit
- defensive strength of the defending unit
- whether or not either unit has already been wounded in battle
- terrain that might give advantages and disadvantages to either unit
- nearby units that can lend support
- the nearby presence or absence of a General or other leader
- the morale of a unit or of its team
- a random dice roll (to simulate the uncertainty of actual combat)

Generally, players tally these factors and consult an appropriate table that lists the outcome of the encounter. The complexity of the simulation doesn’t end there, however, because the result of an attack can also take a variety of forms, including:

- one or both units is eliminated
- one or both units is reduced in strength
- one or both units is forced to retreat
- one unit displaces the other’s position

A rich procedural representation emerges out of the factors going into and coming out of an encounter between wargame units. By tweaking the formal characteristics of units and the overall resolution system of a game, game designers can arrive at highly specific procedural representations of a historical battle. Depending on the particular game, the simulation might be a World War II tank division encountering enemy infantry, or a troop of horse-mounted archers fighting a phalanx of spearmen in the ancient Middle East.

Wargames are incredibly detailed. At the same time, everything about the representation of units in a wargame is highly stylized: a simple cardboard chit or metal figure “stands in” for a military unit, with each piece representing a whole group of soldiers or vehicles that move as a single block; the straightline or grid-based movement of the units; the reduction of combat to simple numerical factors; a single human player directing the entire battle from a birds-eye view. These are only a few examples of the many ways historical wargames radically abstract their subject matter. However, if you accept these limitations, if you take on the conventions of the game genre, within them there is room for endless play, both for the player exploring permutations of history and for the game designer constructing the systems that make the historical simulation possible.

The Field of Battle

In addition to characteristics of the units and the rules for their interaction, the design of the field of play presents its own challenges. A historical wargame has to function as both a playable game and an accurate simulation of history, two concerns that can often be at odds with each other. Wargame designer James Dunnigan writes about some of the design concerns in creating a game map:

There are two primary things to keep in mind when examining a geographical game map. First, it often has a grid, most often a hexagonal grid, superimposed over it. The second point is that in most historical situations, only very large (“gross”) terrain features have any significant effect on operations. Thus, a great deal of detail on a map will often get in the way of providing an accurate simulation. The designer usually feels obliged to justify all of this detail. Often the gamer will be equally expectant that all of this detail be put to some use or otherwise why bother him with it. There is an unspoken assumption that only that which is essential is displayed. It is normally considered a bad design if information is included in the game that does not contribute to one’s understanding of what is going on.
The real world is infinitely rich, and cartographers—including game map-makers—are faced with the representational challenge of simplifying geography in a way that is meaningful for the intended use of the map. In a map for a historical wargame, designers must decide what to include and what to leave out, how to abstract and structure the information to fit in the larger game system. As Dunnigan puts it, too much detail in the terrain can get in the way of a player’s understanding; only “gross” terrain features have a real impact on military operations. Abstraction emphasizes the features critical to understanding the terrain, while minimizing the “noise” created by less important elements.

The grid of the map is one important consideration. Units in a grid-based wargame move only within the hexagons or squares superimposed on the map. Because of this, there is a very specific relationship between the grid and the terrain. The game designer not only needs to select the relevant terrain features, but also decides how those features fit into the grid. Because terrain can affect movement, simply laying a grid over topographically correct terrain creates formal ambiguity. For example, in a particular game, units might not be able to cross rivers. If a river flows through the middle of a map hex, does it mean that players can enter the hex but not exit out the other side, or that they cannot enter the hex in the first place? A common solution in wargame map design is to stylize the shape of rivers so that they are located only on the edges of hexes. This solution makes the movement-blocking role of rivers in the game completely clear. Designing the terrain to accommodate the game grid lessens the geographic accuracy of the map: there are no naturally occurring rivers that can be plotted exactly along a hexagonal grid. But for game design purposes, abstraction that eliminates formal ambiguity is essential.

Questions about rivers and wargame map design do not end there. If one purpose of a historical wargame is fidelity to the real battlefield, which rivers should be included? When does a tributary or stream become too small to be indicated on the map? Which rivers should have an impact on the play of the game? As Dunnigan points out, if something is prominent on the map, a player will expect it to impact game play: a visible feature that does not contribute to the functioning of the rules is bad design.

In designing a wargame map, in deciding which features to include and how to represent them within the larger simulation, you are doing more than just creating a map. You are constructing a space of meaning. If your game simulates combat between individual soldiers, the terrain elements you include make a representational statement about which type of terrain affects a certain kind of combat. The meaning of a wargame map arises not just from its geographic or pictorial features: the meaning derives from the role the map plays in the larger game experience. The formal qualities of the map make certain player actions possible, actions that constitute the ongoing moments of game play.

Not all wargames use a grid. Some miniatures games measure unit movement in inches, and in some digital wargames, unit movement is free-form and highly granular. In these cases, the principles of abstraction and meaningful play still hold. Are the map elements communicating their meaning to the player? Do they affect game play in ways that make sense within the
larger system of the game? Is the simulation creating a coherent representation? Are the outcomes of the player’s choices meaningful? As play unfolds moment by moment, the total experience of the game emerges. In this way, the play of simulation brings us back to the most fundamental questions of game design.

**Simulation in Context**

Historical wargaming provides a terrific context for better understanding issues of representation and simulation because the premise of the genre is that a game accurately depicts a real-world historical referent. Most kinds of games do not have such an orthodox view of how accurately they simulate “reality.” However, a related set of issues has increasingly come to prominence within digital games. Exactly how does a computer or video game procedurally represents its subject? At stake are the same core concerns we explored in historical wargames: how it is that a game simulation can create meaningful play.

The steadily increasing power of computer technology to simulate and manage complex systems has opened up new possibilities for game design in the digital realm. Incredibly detailed simulations of light, sound, physics, agent behavior, and other phenomena are becoming commonplace within games. Many recent writings from digital game designers focus on this feature of digital media and suggest strategies for game designers to use in their work. Let’s compare three such examples:

- In *Swords & Circuitry*, Halford and Hallford discuss simulation design in games using the example of a grenade destroying a door. One approach to simulating this effect would be to specify the relationship between each grenade and each door in the game. In this case, every possible instance of a grenade effect in the game would have to be explicitly spelled out in the program. In contrast, in a more flexible system, grenades would belong to a general category of objects that cause damage, and doors would belong to a general category of objects that break when they receive damage. Hallford and Halford strongly prefer this latter approach to simulation design.10

- In *Game Design: Theory and Practice*, Richard Rouse III discusses a related set of ideas through the example of a dungeon puzzle, in which players open a secret door by dropping objects on a pressure plate trigger. One design approach would be to hard-code relationships between every object and the pressure plate, so that objects defined as “heavy” trigger the plate. Rouse advocates the creation of a generalized weight system instead, in which every object in the game has a numeric weight rating; if the weight value of the objects on the pressure plate reaches a certain number, the plate is triggered.11

- In “The Future of Game Design,” Harvey Smith shows a similar preference in an example involving bird behavior. Bird behavior could be modeled so that when a player moves within a certain radius of a simulated bird, the bird flies into the air. However, Smith would rather see a more detailed simulation in which the bird’s behavior could be triggered by the perception of light, sound, motion, or other modeled stimuli that would be more tightly integrated into the system of the game as a whole.12
All three of these examples make a similar point: there is a difference between a simple, case-based structure for a simulation and a more complex generalized structure that relies on integrated, systemic relationships. Although both approaches create procedural representations, the authors show a clear preference for one approach over the other. In their work, they cite a number of reasons why a generalized strategy is better for designing simulations:

- **It decreases work time:** When a game system is large, generalized systems allow for much more flexible design. In the grenade example, specifying every possible interaction between every possible weapon and every possible object in the game would be a major programming task. Making adjustments to these relationships once they are established (such as reverse-engineering metal doors to be immune to grenade damage) means going back into the code and modifying every affected instance. By creating classes of objects as Hallford and Halford suggest, categories of objects can be moved in and out of different effect classes, so that game designers can quickly try out different combinations of relationships in the game.

- **It increases emergence:** More flexible game simulations lead to a greater degree of emergence in the game as a whole. In Smith's example, having a more detailed behavioral simulation for the birds creates more varied roles for the birds to play in the game. If the birds react to sound and not just to proximity, a game moment in which gunshots ring out and the flock of birds dramatically takes to the sky becomes possible. This is not necessarily something designed directly into the game, but it is an emergent effect of the simple rules governing bird behavior. With more detailed simulations, the space of possibility is enlarged and complexified.

- **It increases play options:** More generalized simulation systems give players more choices and more ways to solve problems. In Rouse's example of the dungeon pressure plate, a player that had no object to drop might create a magical snowstorm that created enough weight on the floor to affect the plate. Idiosyncratic play styles are encouraged, rewarding players for exploring the increased space of possibility. This leads to more distinct styles of play and more avenues for meaning. If a group of Smith's birds were present in a deathmatch game, for example, a smart player might strategically position himself, fire a shot to scatter the flock, and then use the motion of the birds as visual cover for an assault on the enemy.

### A Balanced Approach

Hallford and Halford, Rouse, and Smith strongly advocate the use of generalized simulation design techniques as opposed to a case-based approach. But is it really possible to integrate every aspect of a game within a generalized system? One eloquent thinker on this topic is game designer Marc LeBlanc:

By "simulation" I mean game systems that are rich, many-faceted analog models, whereas by "emulation" I mean coarse, case-based game systems. An example of "simulation" might be the physics system in a shooter or the damage model in an RPG. The rules of the system create a space for exploration, allow for emergent complexity, and so forth. An example of "emulation" might be, in either game, a button that opens a door. There is no physical simulation of the electro-mechanical process of opening the door. It's just a single rule: "If button pushed, then door opens." The button is more of a semantic object than a physical object; in some sense it exists only on a functional level. Emulation is the smoke-and-mirror approach.13

LeBlanc brings to light a point we made at the start of this chapter: *simulations can be embedded.* Even though an entire game can be considered as a single representation, it is important to be able to identify the smaller procedural representations that make up the larger whole. Although LeBlanc shares a preference for generalized simulation design strategies, he takes a measured approach and ultimately suggests a balance of case-based and generalized techniques. No game can rely
entirely on rich, open-ended, emergent simulation. Nor should it. To demonstrate this balancing act, LeBlanc takes an example from a title he helped create, the computer game Thief, in which players can use water to put out a torch:

This is an emulated game system; there's a single rule that says water puts out fire 100 percent of the time. There's no simulation; no chance of using too little water and just getting steam, and no chance of drowning the torch so that it can never be reli. However, the other systems that interact with that system are simulations; you typically douse a torch by tossing a "water balloon" at it, the motion of which is physically simulated. The way the light from the torch affects your vision, and more relevantly, the vision of the opponents you want to stay hidden from, is also simulated.14 (Remember that LeBlanc is using his own terminology—for example, he is using the term "simulation" to mean a generalized design approach, which is only one part of what we identify as the larger issue of simulation in games.)

LeBlanc's example contains intersecting systems of representation, some case-based and some generalized. In Thief, water puts out fire 100 percent of the time. Rather than a complex physics simulation with a variety of possible results, there is a predictable and consistent outcome: water extinguishes fire. But this case-based approach interacts with more complex structures. Tossing a water balloon at the torch involves the generalized simulation of motion; the diminished visibility from the extinguished torch results from a generalized simulation of light. The whole of the game representation emerges from the complex relationships of these parts.

In fact, most games use an approach that combines case-based and generalized representations. The Deux Ex characters of thug, civilian, and military agent were designed with generalized heuristics that produce emergent behavior. But other characters in the game were designed with simple, case-based behavior, such as a character that activates a cinematic cut-scene when encountered. Even simple board games such as Candyland combine a generalized movement system with a number of case-based special spaces on which players can land.

If Rouse, Smith, and Hallford and Hallford are correct, if a generalized approach saves time, increases emergence, and provides players with richer play, why would a digital game designer (or any game designer) use a case-based approach for structuring a procedural representation? One answer is that a truly generalized system could easily become overly complex to implement and might not save work time in the end. Another answer is that although case-based approaches can sometimes become simple and flatfooted, the opposite danger is true of generalized strategies. Taken too far, generalized simulation design can become fuzzy and ambiguous. If you are playing a game and you see a group of birds that suddenly takes to the air for no reason, their behavior will seem random and meaningless. It might be that their behavior is simulated in such detail that their internal clocks told them it was time to leave the scene and migrate south for the winter. Unfortunately, the accuracy of the simulation would be lost on the player, who has no way of knowing how the meaning of the birds' actions fit into the larger game experience.

Neither case-based nor generalized design strategies guarantee a successful game experience. The goal of game design is the creation of meaningful play, which should guide the selection of representational strategies in a game. In the example of the Zelda witch, a compelling character was created out of a central case-based procedure: if the player brings a mushroom to the witch, she will turn it into magic powder. This procedural representation was well-integrated into the other more generalized representational strategies of the game (the witch's hut is isolated, difficult to find, and dangerous to reach; mushrooms are hidden in the forest; magic powder has special effects in the game). Thus the meaning of the witch gained its power from the total game context in which it was experienced.
The fact that rich meanings can emerge from a representational context not based on software complexities offers an important insight into game design and simulation. As we’ve mentioned, games currently suffer from a narrow range of simulated subject matter. Although there are important historical reasons for the prevalence of military and economic conflict in games, other forms of conflict, such as social, cultural, or emotional conflict, can and should be represented as well. Some presume that an increase in technological complexity will make such representations possible. For example, a widely published quote from an executive at a major console manufacturer not too long ago looked forward to the day when the faces of game characters could represent emotion. According to the executive, on that day, games would become a mature and sophisticated form of cultural production. Clearly, media such as literature, theater, and comics have been capable of sophisticated representation for centuries without relying on high-resolution animation. Furthermore, even within the history of animation, many animated tears have already rolled down numerous animated cheeks. That fact alone is no guarantee that the story was meaningful to its viewers. In the case of computer games, although animated elements do play a part, the systemic and interactive qualities of the form have to be taken into account when envisioning future directions for the medium.

The procedural representation of new kinds of game content is within our grasp, but new content can only be discovered by paying attention to the fundamental principles of game design and meaningful play. Game designers need to cultivate a deeper understanding of the form in which they work. This is especially true in considering games as simulations. More than just choosing a representational design strategy, there is a complex interplay between a simulation and its simulated subject. It is to this relationship that our attention now turns.

The Value of Reality

We have come to a turning point in this chapter. Up to now, our primary focus in considering the play of simulation has been on how simulations operate: the mechanisms that generate procedural representation, the abstraction necessary when game designers simulate a particular subject, the selection of case-based and generalized strategies for structuring simulations. But there is a very different aspect of simulations for us to consider. Recall our definition: a simulation is a procedural representation of aspects of "reality." So far, our emphasis has been on the procedural component of simulations. But now we turn to the equally important and immensely complicated set of questions regarding simulations and "reality."

What is the relationship between the simulated content of a game and its real-world or imagined referent? At the 1998 Game Developers Conference, game designer Steve Jackson shared a fascinating anecdote about creating a driving combat computer game based on his classic paper game Car Wars. Using real-world physics and car data, the development team created an unusually detailed driving simulation that incorporated minute details of driving physics and a detailed simulation of the car engine. They also created a track based precisely on the geometry of an existing speedway. But when they test-drove their simulated car, using a steering wheel and pedal interface, they weren’t able to reach the top speeds of the car on which the simulation was based. One day a professional race car driver visited the company. He sat down at the game prototype and immediately drove the simulated car around the track at breakneck speed, completing it close to the real-world speed record.

The simulation was so "accurate" that it required expert manipulation in order to resemble the real-world phenomena it had been designed to replicate. What’s the lesson? Don’t forget the player. The designers of the game had assumed that simulation design meant only formally recreating a mathematical model of the car and the track. In fact, a game simulation not only includes the formal mechanisms of the system, but also the
ways that those mechanisms engender and permit player action. The rules never solely determine the play of a game. The rules are always set in motion within an experiential context that includes particular players with their own levels of desire, skill, and expectation. The Car Wars designers had created a certain space of possibility with their design, but it took the right kind of player to navigate that space in the way it was meant to be explored.

The Car Wars anecdote reminds us that questions regarding the "reality" of a representation are never as simple as they seem at first glance. Was the car simulation "accurate"? Or was it only accurate in the hands of a professional race car driver? Is sitting in front of a computer monitor anything like driving a car? Would the race car driver have been able to reach top speeds playing with a standard console controller? Does the fact that the experience was "only a game" impact the answers to these questions?

When players interact with a simulation, they are never playing with the real thing. If they were, it couldn't be called a "simulation." At the same time, a simulation does reference its depicted subject through images, sound, and procedures. But how do these representations relate to their referents? In language, we know that the letters C-O-W don't resemble a cow in any way. But a photograph of a cow does bear some similarity to our own perception of a cow in the real world. How do games relate to their depicted subject matter? To answer these important questions, we return to the concept of metacommunication.

**Framing the Simulation**

*Children know that they are manipulating their thoughts about reality, not reality itself; and they know that their play self is not the same as their everyday self.*—Brian Sutton-Smith, *The Ambiguity of Play*

In *Games as the Play of Meaning*, we introduced Gregory Bateson's concept of metacommunication, the unique form of communication that takes place in the context of play. To use Bateson's own example, when a dog nips another dog, the nip signals two things. On the one hand, the nip signifies a bite; it is a stand-in for the action of a real bite. On the other hand, the nip signifies just the opposite of a bite: it signals the fact that the two dogs are playing and not actually fighting. This kind of metacommunication—communication about communication—is present not just in informal play but in games as well. It is a significant part of the complex mechanisms games use to construct meanings for their players.

Why is the concept of metacommunication so important, especially in the context of simulation? Metacommunication makes it clear that to play a game is to take part in a kind of double-consciousness. Game actions refer to actions in the real world, but because they are taking place in a game, they are simultaneously quite separate and distinct from the real world actions they reference. A kiss in Spin the Bottle or a frag in a Quake deathmatch refer to kissing and killing, but at the same time are actions that communicate *I'm not kissing or killing you, I'm just playing.* The magic circle is the space within which such paradoxical signals become meaningful.

In "A Theory of Play and Fantasy," Bateson uses the following diagram to illustrate the paradoxical state of mind embodied in play:15

```
All statements within this frame are untrue.
I love you.
I hate you.
```

This schematic is a riff on Epimenides' Paradox, also known as the Liar Paradox. The Liar Paradox is the philosophical problem of someone asserting "I am lying." If the speaker is a liar, then he is telling the truth, and vice versa: the liar's statement is a logical
paradox. In the diagram, the first sentence, *All statements within this frame are untrue*, echoes this classical logic problem. But significantly, it locates the statement within a *frame*, a limited context within which the paradoxical sentence asserts its meaning.

For Bateson the frame is a psychological and philosophical construct that delimits the peculiar space of play. For game designers, Bateson's frame offers another way of understanding the magic circle of a game. It is a boundary that makes the paradoxical meanings of play possible. At the same time, the frame is only sustained by virtue of the continued metacommunicative assertions of play. In Bateson's illustration, the frame enables the statement's meaning, even as the frame's own meaning comes directly from the statement itself.

The magic circle is both a prerequisite and an effect of play. It is a robust context for the exhilarating experiences of game play. But it is similarly fragile, and vanishes quickly when a game ends. Bateson's diagram is a schematic of the cognitive frame of play, a visual retelling of the state of mind of a player in the midst of a play context. As a way of understanding what happens when a player enters into the magic circle and plays with a game simulation, it is a subtle and powerful illustration.

What about the other two statements, *I love you* and *I hate you*? These statements are also part of the paradoxical meanings captured within the frame. The two sentences address a larger point Bateson makes about set theory, and whether some or all of the statements within the frame could be considered true or untrue. For our present purposes, we will sidestep his larger argument to make a point of our own. Bateson could have included any two contradictory sentences in the frame. But he chose emotional statements about love and hate, statements seemingly addressed to someone else outside the frame.

These two little sentences, signals of pure emotion, remind us that the questions of play and meaning, of metacommunication and paradox, are not just abstract philosophical chatter. In understanding how games construct meaning, we are address-

ing the deeply felt ways that players engage with games and the emotional and social realities games reflect and construct. The metacommunicative state of mind is deeply intertwined with the unique pleasures and experiences of play.

**The Immersive Fallacy**

All forms of entertainment strive to create suspension of disbelief, a state in which the player's mind forgets that it is being subjected to entertainment and instead accepts what it perceives as reality. — François Dominic Laramée, "Immersion"

We will return to Bateson's ideas about metacommunication and meaning in just a moment. But for now, let's bring the discussion back to the play of simulation, specifically the relationship between a game and the "reality" upon which it is based. The preceding quote is from a book on game design, appearing in an essay on "Immersion." Game designer and programmer François Dominic Laramée argues for a particular relationship between a game player and a game, between the player's state of mind and the perceived reality of the experience. He asserts that a game should strive to create an experience in which the player forgets that he or she is experiencing designed entertainment and instead believes that playing the game is experiencing reality firsthand. In fact, Laramée states that "all forms of entertainment" function in this way. This is a point of view very much at odds with our own.

We don't mean to unfairly single out Laramée. His ideas about how a player experiences the "reality" of a game are extremely common in the digital game industry, the game press, and even in the public at large. Game designer Frank Lantz has called these kinds of ideas about immersion "widely held but seldom examined beliefs. We wholeheartedly agree, and in the pages that follow we refute these beliefs, referring to them as the immersive fallacy. The immersive fallacy is the idea that the pleasure of a media experience lies in its ability to sensually transport the participant into an illusory, simulated reality.
According to the immersive fallacy, this reality is so complete that ideally the frame falls away so that the player truly believes that he or she is part of an imaginary world.

Although the immersive fallacy has taken hold in many fields, it is particularly prevalent in the digital game industry. Common within the discourse of the immersive fallacy is the idea that entertainment technology is inevitably leading to the development of more and more powerful systems of simulation. These technologies will be able to create fully illusionistic experiences that are indistinguishable from the real world. In an online discussion about the future of gaming, game designer Warren Spector speculated on this topic:

Is the Star Trek Holodeck an inevitable end result of games as simulacra? The history of media (mass and otherwise) seems pretty clearly a march toward even more faithful approximations of reality—from the development of the illusion of perspective in paintings to photography to moving pictures to color moving pictures with sound to color moving pictures with sound beamed directly into your home via television to today’s immersive reality games like Quake and System Shock. Is this progression inevitable and will it continue or have we reached the end of the line, realism-wise?\(^{17}\)

To be fair, Spector self-consciously exaggerated his views in order to spark discussion. But in the debate that followed, it was clear that many participants take for granted the propositions that Spector articulated.

Spector’s selective history of entertainment technologies offers one reading of the development of media. But there are others. History rarely provides such a linear progression, and in regard to immersion, cultural developments tend to be cyclical. As theorist Marie-Laure Ryan puts it, “The history of Western art has seen the rise and fall of immersive ideals.”\(^ {18}\) According to Ryan, immersion as a representational goal has gone through a number of stylistic cycles over the centuries. In the last several decades, she asserts, immersion has in fact become less prominent and respected in fields like art and literature. Ryan may be correct in regard to larger cultural movements, but within the digital game industry, belief in the immersive fallacy remains alive and well.

**Metacommunicative Media**

The immersive fallacy is symptomatic of contradictory ideas about technology. On one hand, there is a technological fetishism that sees the evolutionary development of new technology as the saving grace of experience design. On the other hand, there is a desire to erase the technology, to make it invisible so that all frames around the experience fall away and disappear. Nowhere are these contradictory ideals more clearly expressed than in the concept of the holodeck, a fictional technology that first appeared in the television show *Star Trek: The Next Generation*. The holodeck is the dream of the immersive fallacy, a room in which matter and energy are manipulated to create a simulated environment of sight, sound, touch, smell, and taste that is a representation completely indistinguishable from lived reality.

What is wrong with this picture, and how does it relate to games? On one level, the immersive fallacy actually does make intuitive sense. When we play a game, we feel engaged and engrossed, and play seems to take on its own “reality.” This is all certainly true. But the way that a game achieves these effects does not happen in the manner the immersive fallacy implies. A game player does become engrossed in the game, yes. But it is an engagement that occurs *through play itself*. As we know, play is a process of metacommunication, a double-consciousness in which the player is well aware of the artificiality of the play situation.

During the same online conversation in which Spector posted his intentionally provocative question, film studies scholar Elena Gorfinkel responded:
Remediating Games

In some sense, the layered, metacommunicative state of play is similar to our experience of all media. In their book *Remediation*, theorists Jay David Bolter and Richard Grusin analyze the mechanisms by which media function, arguing that media operate according to a double logic. On one hand, media participate in what Bolter and Grusin call *immediacy*, the ability to authentically reproduce the world and create an alternative reality. At the same time, media also remind their audiences that they are constructed and artificial, a characteristic that Bolter and Grusin call *hypermediacy*.

Like other media since the Renaissance—in particular, perspective painting, photography, film, and television—new digital media oscillate between immediacy and hypermediacy, between transparency and opacity. Although each medium promises to reform its predecessors by offering a more immediate or authentic experience, the promise of reform inevitably leads us to become aware of the medium as a medium. Thus, immediacy leads to hypermediacy.  

For example, as Bolter and Grusin point out, a web cam promises immediacy though authentic, real-time access to another part of the world. But the fact that users have to view the web cam on a computer, in an operating system, in a browser, on a web page, inside an interface, reminds them that they are not transparently experiencing the locale where the web cam exists, but are instead interacting with a highly artificial media construct. The main argument made by Bolter and Grusin is that all media combine these two processes into what they term *remediation*, an experience of media in which immediacy and hypermediacy co-exist.

We can also analyze games within this model. The double consciousness of play finds a strong parallel in the process of remediation, which mixes transparent immediacy with a hypermediated awareness of the constructed nature of play. In Cops and Robbers, players willingly take on the theatrical roles of criminals and police, even as they infuse those playful representations with meaning through their actions. In a first-person
shooter such as Halo, part of the experience is the sensual vertigo of navigating a coherent, imaginary 3D space. But playing the game also involves an awareness of the game interface, the strategic use of the frame-breaking options, the use of text-based chat, fluctuating server speeds, and the sharing of tips with friends in the larger social context of play. These frame-related aspects of the Halo experience remind the player that the game is a constructed, hypermediated experience.

The value of Bolter and Grusin’s model is that it doesn’t do away with illusionistic immersion, but includes it as one element within a more complex process. There is no doubt that the immediacy of sensory engagement is part of the pleasure of playing a game, particularly digital games with detailed representations that respond in real-time to player action. The immersive fallacy grossly overemphasizes these forms of pleasure, and in so doing, misrepresents the diverse palette of experiences games offer.

The Character of Character
The danger of the immersive fallacy is that it misrepresents how play functions—and game design can suffer as a result. If game designers fail to recognize the way games create meaning for players—as something separate from, but connected to the real world—they will have difficulty creating truly meaningful play. To highlight these complexities, we now take a detailed look at just one aspect of a game’s representation, character, to see how an understanding of metacommunication can impact the game design process.

We have already looked at character once in this schema, examining the way that procedural representations construct fictional personas in Zelda: Link’s Awakening, Virtua Fighter 4, Deus Ex, and The Blob. Now we’ll take aim at the other part of the simulation equation, pointing out the way that character representation relates to the “reality” outside the game. Two key questions arise: How does the player relate to a character in a game? And how can this relationship be understood in terms of the “reality” of the represented world? Just to keep things focused, we will limit our analysis to protagonist characters that a player directly controls, such as Mario in Super Mario World or Pai Chan in Virtual Fighter 4.

The immersive fallacy would assert that a player has an “immersive” relationship with the character, that to play the character is to become the character. In the immersive fallacy’s ideal game, the player would identify completely with the character, the game’s frame would drop away, and the player would lose him or herself totally within the game character.

These ideas have some validity, but they represent only one element of a much larger and more complicated process. A player’s relationship to a game character he or she directly controls is not a simple matter of direct identification. Instead, a player relates to a game character through the double-consciousness of play. A protagonist character is a persona through which a player exerts him or herself into an imaginary world; this relationship can be intense and emotionally “immersive.” However, at the very same time, the character is a tool, a puppet, an object for the player to manipulate according to the rules of the game. In this sense, the player is fully aware of the character as an artificial construct.

This double-consciousness is what makes character-based game play such a rich and multi-layered experience. In playing the role of Cloud in Final Fantasy VII, the player has a portal into the complex narrative world of the game. Through Cloud, the player encounters the settings, characters, and events of the game world; many players report a strong emotional attachment to their digital counterpart. At the same time, Final Fantasy VII is a complex role-playing game. The play experience occurs by watching cutscenes, navigating Cloud and his comrades though virtual spaces, managing a detailed inventory of weapons, items, and magic, taking part in constant strategic battles, and engaging with the game’s intricate spreadsheet-like interface. Through these diverse activities, the performance of play acknowledges and celebrates its own hypermediated construction.
The psychologist Gary Alan Fine, in his excellent book *Shared Fantasies*, offers a model for understanding the complex relationship between player and character. *Shared Fantasies* is an ethnographic study of tabletop role-playing game communities. Borrowing from psychologist Erving Goffman’s theories of Frame Analysis, Fine identifies three “levels of meaning” within which the player/character game experience takes place:

First, gaming, like all activity, is grounded in the “primary framework,” the commonsense understandings that people have of the real world. This is action without laminations. It is a framework that does not depend on other frameworks but on the ultimate reality of events.

Second, players must deal with the game context; they are players whose actions are governed by a complicated set of rules and constraints. They manipulate their characters, having knowledge of the structure of the game, and having approximately the same knowledge that other players have.

Finally, this gaming world is keyed in that the players not only manipulate characters; they are characters. The character identity is separate from the player identity.  

This three-fold framing of player consciousness—as a character in a simulated world, as a player in a game, and as a person in a larger social setting—elegantly sketches out the experience of play. The player and character frames both take place within the magic circle, whereas the person frame gains its primary meaning from the cultural context outside the immediate space of play. Fine makes the important point that movement among these frames is fluid and constant, and that it is possible to switch between them several times in the course of a single verbal statement or game action.

In digital games, the same multi-layered phenomena occurs. Imagine a player, holding a joystick-like controller, looking at a glass screen. The player is deeply engaged in a game activity, sweating and anxious, focused completely on the space in front of him, leaning his body in sync with the visceral rhythms of the game, smiling and grimacing as he battles opponents and his actions play out in the world on the other side of the glass. What game is he playing? Try on both of these answers for size:

He is playing *Tomb Raider*. Our hypothetical player is looking at a television screen and manipulating a console controller. In one sense, our player immerses himself in the game’s narrative world, taking on the identity of Lara Croft with her requisite strengths and weaknesses (*I feel lost… I can’t believe I survived that trap*). Simultaneously, he views her exaggerated anatomy from behind, pushing buttons and manipulating her like a puppet on his quest to find power-ups, overcome obstacles, and unlock doors to reach the next level (*What was that cheat code again? This cutscene sucks*). He is both character and player. In addition, the larger social and cultural context in which he plays constitutes Fine’s category of the player as person. Maybe he is trying to impress a friend with his skillful play. Or perhaps he is taking mental notes for a lecture he is going to give at an academic conference. In any case, the player is always present as a person connected to and situated in the real world.

He is competing in Comedy Central’s *BattleBots*. In this case, the player’s character is a battling, remote-controlled robot moving about the real world, the pane of glass not a television screen but a large sheet of plexi that protects the players and audience from flying scraps of metal. The BattleBots player is immersed in his activity too, and like the Tomb Raider player he is always aware that his actions are governed by the rules of the game. During game play, he might switch between the character/player/person frames many times, moving between emotional identification with his robot character (*Ouch! I just got slammed!*) and his role as player in the game contest (*Let’s see if I can get my bot out of the corner*). He might even break the frame of player to wave to a friend in the crowd or to offer a sound bite to the television host.
Fine's three-layer model is an extension of the double-consciousness of play. Players always know that they are playing, and in that knowledge are free to move among the roles of person, player, and character. Players of a game freely embrace the flexibility of this movement, coming in and out of moments of immersion, breaking the player and character frames, yet all the while maintaining the magic circle.

This model applies even when players are not directly controlling a game protagonist. In any game, players move constantly between cognitive frames, shifting from a deep immersion with the game's representation to a deep engagement with the game's strategic mechanisms to an acknowledgement of the space outside the magic circle. Devotees of the immersive fallacy tend to see this hybrid consciousness as a regrettable state of affairs that will only evolve to its true state of pure immersion when the technology arrives. Play tells us otherwise. The many-layered state of mind that occurs during play is something to be celebrated, not repressed—it is responsible for some of the unique pleasures that emerge from a game.

Hacking the Holodeck
The questions surrounding games as simulations are always more complex than they first appear. There is no simple relationship between player and character, or player and game, or game and outside world. This is one reason why the immersive fallacy continues to colonize most design thinking about the future of games and the role of technology in creating compelling experiences: it is simply an easier position to take.

But the immersive fallacy is more than an idea. It is also a stumbling block to advances in game design, as it represents an overly romantic and antiquated model for how media operate. As long as game designers are caught up in a desire for the technology of the holodeck, they lack the vision to appreciate the potential for game innovation today. What if game designers focused their efforts on actively playing with the double-consciousness of play, rather than pining for immersion?

Imagine the kinds of games that could result: games that encourage players to constantly shift the frame of the game, questioning what is inside or outside the game; games that play with the lamination between player and character, pushing and pulling against the connection through inventive forms of narrative play; games that emphasize metagaming, or that connect the magic circle so closely with external contexts that the game appears synchronous with everyday life. Innovation is only bound by a failure to see the fundamental principles of play.

A common complaint among game developers is that games are not recognized as a significant form of culture, and that they lack a diverse mass audience. Instead, games seem to be relegated to the backwaters of culture. A sea change in cultural status will only occur when game designers acquire a more sophisticated understanding of how their media operates. Robust forms of contemporary pop culture are not premised on naïve ideas of immersion. Just take a look at the explicit self-consciousness of hip-hop, fashion, and Animé. These forms of popular culture have a deep understanding of the way media cultivates immersion while making explicit the mechanisms through which the representation is experienced.

This, of course, brings us back to simulation. Even though simulations are premised on the notion of fidelity to their referent, the very fact that they are dynamic systems means they allow for the exploration of alternate permutations. Simulations allow players to explore a space of representational possibility through the very act of play. Certainly there are a great many game designers driven by a desire to tell stories and provide narrative worlds for players. Framing games as simulations, as dynamic systems of procedural representation, unlocks the potential of games as a powerful representational and narrative medium. But games have only just scratched at the surface. Questions remain: What can games represent? How can games engage players through meaningful play? And how can games challenge, critique, and contribute to the world outside the magic circle?
Further Reading

*Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media,* by Marie-Laure Ryan

An eloquent articulation of the relationships between literary theory, hypertext, and VR, *Narrative as Virtual Reality* focuses on what Ryan sees as the competing interests of immersion and interactivity. Although Ryan's sophisticated approach elevates her above the usual pitfalls of the immersive fallacy, she is an apologist for immersion, and her discussions of interactive design suffer as a result. That said, for the topics that this thick volume covers, it is essential reading.

*Recommended:*
- Part II: The Poetics of Immersion
- Part III: The Poetics of Interactivity

*Remediation: Understanding New Media,* by Jay David Bolter and Richard Grusin

Bolter and Grusin introduce the concept of remediation, the process in which new media forms define themselves by borrowing from and refashioning old media. This process also works in reverse: older media forms borrow from new media forms, such as television remediating the windowed world of computing. The book's most useful concepts for game designers are *immediacy,* and *hypermediacy,* which refer to the way media forms can both authentically reproduce the world while simultaneously reminding the audience that the reproduction is both constructed and artificial.

*Recommended:*
- Introduction: The Double Logic of Remediation
- Chapter 1: Immediacy, Hypermediacy, and Remediation
- Chapter 2: Mediation and Remediation
- Chapter 4: Computer Games

*Shared Fantasy,* by Gary Alan Fine (see page 417)

*Recommended:*
- Chapter 1: FRP
- Chapter 2: Players
- Chapter 3: Collective Fantasy
- Chapter 6: Frames and Games

"A Theory of Play and Fantasy" by Gregory Bateson (see page 373)

Notes

6. Robinett, *Inventing the Adventure Game.*
8. Ibid.
12. Harvey Smith, "The Future of Game Design"
14. Ibid.
22. Frank Lantz, *Hacking the Holodeck,* unpublished manuscript.
- A simulation is a procedural representation of aspects of "reality." Simulations represent procedurally and they have a special relationship to the "reality" that they represent.

- There are many kinds of simulations that are not games. However, all games can be understood as simulations, even very abstract games or games that simulate phenomena not found in the real world.

- Game simulations usually operate metaphorically: they do not literally recreate a representation of their subject matter. The difference between a game simulation and its referent can be a source of pleasure for players.

- A procedural representation is a process-based, dynamic form of depiction. Procedural representation is how simulations simulate their subject matter. These forms of representation emerge from the combination of the formal system of a game and the interaction of a player with the game.

- An entire game can be considered a procedural representation of a particular subject. In addition, games include smaller procedural representations that make up the larger depiction.

- The subject matter of game representations is linked to the kinds of conflict that a game can represent. Games typically represent territorial conflict, economic conflict, or conflict over knowledge. Most games combine two or all three of these categories. It is possible to represent other forms of conflict as well.

- Simulations are a powerful way of thinking about narrative because procedural representation is an approach to storytelling that directly emphasizes the player's experience.

- Simulations are abstract, numerical, limited, and systemic. A simulation cannot be both broad and deep. Because designing a simulation means radically reducing the simulation’s subject matter, a game designer must carefully select which aspects of a phenomenon to depict and how to embody them within the system of the game.

- Simulations, especially in digital games, can be structured according to a case-based logic, in which relationships between every element of a system are specified in advance, or a more generalized logic in which system elements share a set of general attributes. Generalized structures can save work time and lead to more emergent games where players have greater options for action. However, a balance between the two kinds of structures is usually necessary in any given game.
- The phenomenon of metacommunication implies that game players are aware of the frame of a game and that a player's state of mind embodies a kind of double-consciousness that both accepts and refutes that frame.

- The immersive fallacy is the belief that the pleasure of a media experience is the ability of that experience to sensually transport a player into an illusory reality. Although the immersive fallacy is prevalent in the digital game industry, it does not take into account the metacommunicative nature of play.

- Media theorists Bolter and Grusin argue that all media operate through the process of remediation. The two opposing elements of remediation are immediacy, which promises true and authentic representation, and hypermediacy, which emphasizes the constructed nature of media representation.

- Psychologist Gary Allen Fine identifies three layers of game player consciousness: direct identification with the game character, engagement with the game procedures as a player, and existence in larger social contexts as a person.
Conflict arises naturally from the interaction in a game. The player is actively pursuing some goal. Obstacles prevent him from easily achieving this goal. Conflict is an intrinsic element of all games. It can be direct or indirect, violent or nonviolent, but it is always present in every game. —Chris Crawford, The Art of Computer Game Design

competition and cooperation
systemic cooperation
player cooperation
game goals
victory and loss conditions
level playing field
Introducing Conflict

What does it mean to consider games as Systems of Conflict? First of all, we agree with Chris Crawford. Conflict is an intrinsic element of every game. Conflict, a game as a contest of powers, is a core component of our very definition of the term “game.” While conflict outside of games can sometimes be destructive, in games we find the wonderful paradox of a staged conflict, resulting in meaningful play.

Game conflict emerges from the unique circumstances of a game. The magic circle imbues games with special meanings. One of the most important meanings to emerge is the game’s victory conditions. Winning the game might only have value within the magic circle, yet players pursue it. By virtue of their participation in the game, they have taken on as meaningful the game’s presumptions and proscriptions, including everything associated with winning. The struggle among the players to achieve the goal of the game and become winners is the competitive activity that drives a game’s system of conflict.

The fact that this activity is a struggle derives from the intrinsic challenge presented by the conflict of a game. As we know from our study of the lusory attitude, games are constructed so that their goals are difficult to achieve. The conflict of a game arises as the game players struggle toward achieving the goal, often in opposition to each other, sometimes struggling together or in parallel.

What are the shapes of conflict that occur in games? Struggle in a game can take many forms:

- Single player vs. single player: a Chess game or Boxing match
- Group vs group: Basketball, Soccer, and other team sports
- One against many: Tag or Mother May I?
- Every player for themselves: a footrace or the strategic board game Risk
- Single player competing against a game system: Solitaire or Tetris
- Individual players competing side by side against the game: casino Blackjack
- A group of players cooperating against a game: Lord of the Rings Board Game

Many games mix and match these forms, such as a wrestling meet, in which individuals compete against each other in pairs, but their scores are added up and applied to the team as a whole. Some games can accommodate more than one of these game modes, such as the arcade game Double Dragon, in which one player can compete against the program, or two players can cooperate against it. Still others have competitive structures that change over time, such as the television show Survivor, in which players are initially divided into two competitive teams, eventually becoming a single group from which a single winner emerges.

Conflict in a game can be direct or indirect. In an arm wrestling match, players are pitted directly against each other, trying to pin the other player’s arm while avoiding being pinned themselves. The back and forth movements of the players’ locked hands is a direct meter of the struggle, indicating how near or how far either one of them is from achieving the winning conditions of the conflict. In a figure skating contest, the conflict is indirect. Competitors each have their own turn to perform and be judged. They cannot directly interfere with each other’s success, and winning the competition means receiving the highest score from the judges.

Still other games mix direct and indirect conflict. In a real-time strategy game such as multiplayer Starcraft, players compete against each other, though they are not always directly interacting. Players have to think offensively and defensively, building their resources and defenses, anticipating the actions of other players. As a game proceeds, the solo activities of each player evolve into direct conflict, as the units controlled by the players come into contact. Further, there is more than one way
to configure the conflict in Starcraft: the game lets players set up team vs. team games, one player vs. many players, human players vs. computer opponents, and so on.

**Conflict Case Studies**

As these examples illustrate, more than one form of conflict can exist within the scope of a single game design. Next we take a detailed look at three different games, focusing on the ways each one configures competition and cooperation between players. All three games are arcade games from the 1980s: Centipede, Joust, and Gauntlet. Each game weaves its own surprisingly complex fabric of player conflict.

**Centipede**

Our first example is the arcade game Centipede, in which the player uses a trackball controller and fire button to move a character at the bottom of the screen and shoot at objects coming down from the top. Centipede might seem at first glance to have a simple and straightforward structure of conflict. But in fact, the formal system provides many ways for players to struggle and pursue goals.

- As a single-player experience, you compete against the program. The game compiles an ongoing “score” based on your performance, and the presumed goal of the game is to achieve the highest score.

- There are many ways that you might pursue goals related to the high score goal of the game. You might have a general idea of what constitutes a “good score,” which you try to achieve. Or you might try to surpass your previous game’s score, or attain a new personal best score.

- You might set other goals besides those involving your score. For example, you might try to play for a certain amount of time, get to a certain level in the game, or destroy every enemy of a particular type that appears. Several of these goals might co-exist with each other and with the score-oriented goals.

- Centipede can be played as a two-player game. Both players alternate play, switching places when the current player loses a life. If you compete against another player in this way, the quantifiable outcome of the game (your score) has new meaning. It is no longer only an indicator of your personal success but becomes a way to compare your performance to that of the other player. Two-player Centipede is a zero-sum competition, where one player wins and the other player loses. In this sense, the actual game scores are important only insofar as they are used to determine the winner. The numeric scores of the players are translated into binary win/lose values.

- Aspects of the single-player competition can be combined with aspects of the two-player competition. You might have lost to your opponent, but you might also have gotten your best score ever, in which case you won in your own self-competition, even while losing to the other player in the zero-sum conflict of the two-player game.

- The fact that players can enter their initials into a high score list creates a different kind of competition: you compete against previous players, whom you probably have never met. This competition is more indirect: you compare your score with their scores, and if you are one of the top eight players, you get to enter your initials into the game for other players to see, bumping off the player at the bottom of the list. However, you might later be bumped off as well. Here, your numeric score is translated into a scaled rank: either your score wasn’t high enough to put you on the list, or you entered the list at a specific rank.

- There are other competition scenarios as well. For example, you might play as a single player and set the goal of making it onto the high score list. In this case, you turn the game into a system of competition with a binary win/loss condition: either you make it onto the high score list or you don’t.
- You might have an ongoing rivalry with a friend about who can achieve the higher score on Centipede. The two of you are not good enough to get on the high score list, but you can still keep track of your relative scores. Your score in this scenario is translated into a rank between you and your friend, a rank that changes as one of you bests the other’s higher score.

**Joust**

Who knew so many different forms of conflict were lurking under the surface of a simple arcade game? Our next example adds even more. In Joust, two players maneuver bird-mounted knights, attacking enemies controlled by the program. Both players can play the game simultaneously, instead of alternating turns. This structure opens up whole new forms of competition.

- Joust can be a single-player game. Individual players receive a score and there is a list of player high scores, including separate rankings for daily high scores and “all time” high scores. Most of the forms of competition in Centipede also occur in Joust.

- Two players could compete to see who gains the higher score over the course of a game. Because players do not alternate turns but compete simultaneously, the scores of both players are visible at all times, heightening the drama of this form of competition.

- The simultaneous two-player structure opens other possibilities for conflict. Two Joust players can attack each other if they wish. One way to play the game is as a fighting game, where players directly attack each other, killing their opponent with a successful attack. The goal of the competition in this case is to kill your opponent more times than you are killed. Playing the game in this way turns Joust into a zero-sum game. Numerical scores do not matter, only who is left alive at the end.

- It is also possible for two players to refrain from attacking each other and instead work together to defeat the computer-generated enemies, strategically coordinating their actions. In this case, the two players compete together against the computer. They might set a goal of reaching the highest level or for playing as long as possible.

- Even if players cooperate, they might still compete in other ways. For example, two players coordinating their actions against the computer might compete to get the higher score.

- Often, these different kinds of competition overlap. The game design of Joust makes it easy for a player to kill another player: if they collide, the one that is in a higher position destroys the other one. Even cooperating players sometimes accidentally kill each other, an event that usually affects the competitive flow of the game. After an accidental killing, one player might become resentful and aggressive and the game might transition into the “fighting game” version of Joust. Or the accidental killer might let his opponent kill him one time, just to balance things out. The game might also just continue as usual.
- Competitive tensions persist throughout the game. Because both players are operating on the screen at the same time, there may be competition about where and how they should play, even if they are not actively trying to kill each other. For example, two players might both wish to occupy a certain section of the screen or attack a specific group of enemy characters. An accidental player-killing (or the threat of one) can enter the game as a result, opening up additional competitive complexities.

**Gauntlet**

In Joust, the two-player simultaneous structure adds new layers to the possibilities of game conflict. In Gauntlet, our third arcade game example, up to four players can play at once. The players take fixed roles (Warrior, Valkerie, Thief, or Wizard) as members of a team. Together the team explores the game spaces, fights computer-generated enemies, and gathers resources that boost their abilities to let them explore further.

- Like Joust and Centipede, Gauntlet can be played by a single player. Gauntlet players also receive a score; if the score is high enough, players record high scores and player initials. All of the single-player and high score list forms of competition apply to Gauntlet as well.

- Unlike Joust, Gauntlet players can only attack computer-generated opponents—their attacks do not affect the other players. As a result, Gauntlet lacks the "fighting game" as a possible form of conflict. Instead, the players consistently work together, usually with the goal of seeing how many levels of the game they can explore.

- Because Gauntlet players receive a score, players might also compete to see who has the highest score at the end of the game. As with Joust, the scores are displayed throughout the game, allowing players to constantly check their relative scores.

- Whenever players clear a level of the game, the game pauses to display the relative points of each player and their overall performance in the game, showing, for example, which player received the most treasure in the last level. These moments highlight score-based and statistic-based competition between players, encouraging them to compare their performances against each other and invent competition around the many kinds of statistics in the game.

- During the actual play of a game, another form of competition takes place over in-game resources. As players progress in the game, a number representing their health is slowly reduced. When a character touches an enemy, health is lowered even more. However, there are many "food" items scattered throughout the game that raise a character's health. Players sometimes compete directly for food, trying to be first to reach the item. Players might also discuss who among them needs the food most and let that player acquire the item. The same is true for other special objects in the game, such as keys and magic potions. These forms of resource-based competition are heightened by the statistic comparisons between levels: at these moments, players take stock of how resources have been distributed among the group and can accuse each other of being "unfair" or "greedy."
A final form of competition unique to Gauntlet involves players spending money on a game. In many arcade games, prestige comes from being able to play for a long time on a single quarter. But unlike Joust and Centipede, Gauntlet lets players extend their current game via cash additions. Players can put quarters into the game during play to add to their characters' health or to resurrect their characters after they have died. This means that as long as players want to continue spending money, they can keep on playing, exploring more game levels. The escalating difficulty of the game ensures that players will need to spend more and more money as they play. This can turn Gauntlet into a completely different kind of conflict, one in which players compete to demonstrate their tolerance for putting money into the game, a form of conspicuous consumption much like high-stakes gambling. Conversely, players might compete to see who can play the longest before having to spend more money to continue, because skillful players will avoid being killed. In this case, spending less money for the same amount of time would be the goal.

There are obviously many, many more models of competition in games. However, even within these three similar examples there is a wealth of ways that conflict can manifest. The point of these examples is to demonstrate how the design of a game leads to forms of conflict. In each case, formal decisions about the game's structure directly shape the nature of conflict emerging from the game. For each game, the following kinds of questions determine the essential formal structures:

- How many players can play?
- Do they play simultaneously or do they alternate playing the game?
- Is there a high score list?
- Are players given constant feedback about their relative scores?
- Does the game pause to allow players to directly compare their scores and other game statistics?
- Are there computer-generated opponents and obstacles that players face together or do the players serve as opponents for each other?
- Does the structure of the game allow players to have direct conflict with each other?
- Are there resources for which players can compete?
- Can players spend money to continue the game or enhance their play?

The forms of conflict we observed follow directly from the way that each game answers these design questions. Take Gauntlet: if players were allowed to damage each other through attacks, the game would lose its enforced cooperative spirit, and interplayer fighting might become common. If players could not continue their game by paying another quarter, competition for in-game resources would be much fiercer, as players would vie against each other to stay alive until the game ended. What is surprising in all three examples is just how rich and multi-layered conflict can be in a game. This richness comes from the fact that players can derive and construct their own forms of conflict in a game. Some of the goals we outlined are explicitly defined by the game rules. Others are emergent forms of competition that arise from the player's active engagement with and manipulation of the game structure.

Back in Defining Games, in discussing whether or not Sim City was a game, we concluded that it was a borderline case. Although Sim City does not formally define goals with quantitative outcomes, it does provide a space within which players form their own goals and arrive at their own outcomes. As the investigation of Centipede, Joust, and Gauntlet demonstrates, in many ways all games can function like Sim City, with players inventing their own goals and layering these goals on top of those defined directly by the rules of a game.
A game’s space of possibility is a space of possible conflict. Part of playing a game involves selecting game goals as a means of navigating and exploring forms and degrees of conflict. What is the best form of conflict to provide your players? As with other aspects of games, there is no single formula that will work best for all players in all contexts. However, providing a rich space of possibility that supports a range of conflict increases the potential variety of players and the ways that they might find your game meaningful.

**Competition and Cooperation**

So far, we have spoken somewhat loosely about competition and cooperation as they relate to the conflict in a game. But what do these terms really mean? Competition occurs when players struggle against each other within the artificial conflict of a game. Perhaps our clearest model of competition comes from game theory: the zero-sum game. In a zero-sum game, one player’s winnings equal another player’s losses. If one player is the victor in a two-player zero-sum game, the other player will necessarily lose. Winning is always equally balanced by losing, making the end sum zero.

A common criticism leveled against games is that they are all competitive, and that competition is somehow undesirable. Framed in this way, competition is something to avoid in order to ensure a positive play experience. Bernard DeKoven, game designer and author of *The Well-Played Game*, states this position eloquently:

> It is clear to me now, that the result of such a union [playing to win] is separation, always separation. It divides us into winners and losers, those who have achieved and those who have failed. The division then leads us into further division. It becomes difficult, now that some of us have won and some of us have lost, to find a game that we are all willing to play well together. It was never our focus at all. Though what we have always cherished most is the game in which we are playing well together, winning takes precedence.

DeKoven’s point is that when the winning and losing of competition enters into the conflict of a game, it becomes the paramount concern of the game’s participants, eclipsing everything else the game has to offer. With all due respect, we disagree. It seems quite clear to us that competitive games can offer genuinely meaningful experiences. Sometimes that meaning can stem from the joy of play itself (DeKoven’s “playing well together”), but certainly much meaning derives from the competitive struggle of a game, from trying to become a winner while avoiding a loss.

The competitive striving toward a goal is fundamental in giving shape to the structure of a game and the way that the game creates meaning. The idea, for example, that in meaningful play a player’s actions are integrated into the larger context of a game is dependent on the competitive nature of games. Without a goal toward which players strive, it is very difficult for a player to measure his or her progress through the system of a game. Without a measure of progress to give a player feedback on the meaning of his or her decisions, meaningful play is not possible. Remember the “horrible” game The Grid in *Games as Emergent Systems?* That game had no goal, and no way for players to compete with each other. There was nothing to motivate players to move their pieces this way instead of that way. Meaningful play was impossible.

Our opinion is that all games are competitive. All games involve a conflict, whether that conflict occurs directly between players or whether players work together against the challenging activity presented by the game system. Without a clearly defined goal, games generally become less formalized play activities. However, just because all games are competitive does not mean that they are not cooperative as well. Although we can assert with confidence that all games are competitive, it is equally true that all games are cooperative. Are these two statements contradictory? Can all games be both competitive and cooperative? The idea that games are both competitive and coop-
ervative is only contradictory if the two terms are mutually exclusive, which they are not. The root of the word “compete” is the Latin con petere, which means “to seek together.”

In what ways are all games cooperative? Recall the magic circle and the lusory attitude, and the way that these aspects of a game create meaning. To play a game is to submit your behavior to the rules of the game, to enter into the time and space that the game demarcates, to traffic in the special meanings that the game offers up. To play a game is to participate in the discourse of the game with the other players. Players can play Basketball together because they both speak the “language” of Basketball. When two players hit the courts for a game of one-on-one, that is exactly what they are doing.

Therefore, to play a game is to cooperatively take on the artificial meanings of the game, to communicate to the other players through the artificial discourse that the game makes possible.

**Terminological Aside: Two Forms of Cooperation**

We use the term “cooperation” here in a slightly different way than at the beginning of this chapter. Saying that all games are cooperative refers to the mechanisms that underlie all games, and the way these structures ensure a shared discourse and cooperative spirit among players. We call this form of cooperation systemic cooperation because it occurs in all games at a fundamental level.

However, when we said that the Lord of the Rings Board Game, in which players work together to defeat the game system, was cooperative, we used the word in its more common sense. Unlike a directly competitive, zero-sum game such as Chess, players in Lord of the Rings win or lose as a group. We call this form of cooperation player cooperation because it describes specific player relationships that do not occur in all games.

The two uses of the word are not ultimately dissimilar. Player cooperation is really just a literalized manifestation of systemic cooperation. Systemic cooperation, as a phenomenon intrinsic to all games, occurs “under the hood” of the experienced game structure, whereas player cooperation happens at a higher level, incorporated more consciously into a player’s understanding of a game.

In this sense, the very act of playing a game is an act of cooperation. It is only through the shared efforts of the players that a game’s fragile magic circle takes shape and is sustained over the course of play. There is a wonderful paradox here. Within the magic circle set aside for the game, within the arena spelled out by the rules, a conflict takes place. The players cooperatively form the space of the game, in order to create a competition for their own amusement. Game conflict is like a duel between actors in a play; it is an elaborately staged competitive artifice, enjoyed in part because of its artificiality. There is genuine conflict in a game, but only within a larger cooperative frame sustained by the participation of the players.

**New Games**

In the earlier critique of Bernard DeKoven’s ideas about the negative aspects of competition, we were not quite playing fair. It is true that DeKoven questions traditional forms of competitive play. It is also true that we do not agree with all of his ideas on the subject. But DeKoven’s concepts have to be understood within the larger context of his important work on games. In his book *The Well-Played Game*, DeKoven argues for a new understanding of play, governed by a shift in emphasis away from competition. Instead, DeKoven is an advocate for more improvisational games in which players take on the role of game designers.

DeKoven was not alone in his ideas. He was one of the early members of the New Games Movement, a group of game designers and play advocates that had a tremendous impact on the culture of games. Founded by Stewart Brand (the same man who started *The Whole Earth Catalog*) in the late 1960s, the New Games Movement was an organization dedicated to the promotion of play and its positive impact on society. During the late 1960s and 1970s, the New Games Movement organized a number of large-scale public game “tournaments” in the San Francisco Bay Area and other parts of the world. Part art happening, part community action, and part playground carnival,
New Games Movement Tournaments embodied a uniquely game-centric, community-based politics of a scale that has not been seen since.

The New Games Movement had a large impact on physical education and the integration of games and play into schools. If you grew up playing with a parachute or huge rubber “Earth Ball” in your elementary school gym class, it is probably due to the direct or indirect influence of the New Games Movement. The New Games Movement published two books (The New Games Book and More New Games) that cataloged their playful game designs. How does the New Games Movement fit into an understanding of games as systems of conflict? The New Games Movement confronted the idea of competition and cooperation head on, creating games and ways of thinking about game design that challenged conventional notions of games as conflict.

Many people think of New Games as non-competitive. Of course this isn’t the case. Most of the games in this book involve competition—it’s what gives New Games its vitality. The effort each player makes to overcome the resistance and achieve the goal is the heart of the game and what makes it enjoyable and gratifying. In most games, the resistance is supplied by your opponent trying to achieve her goal. Your opponent is therefore your partner in the game. The best games are those in which you can play your hardest and still count on your opponent to meet your effort—to compete with you.

Although DeKoven may rail against competition in some of his writings, he also helped instill in New Games the more balanced notions of competition embodied in the quote above, taken from an essay he wrote for the New Games Book. DeKoven’s main point is that in the context of a game, the struggle of players against each other is also a struggle with each other, as players meet the challenges that they provide for one another. In this way, New Games affirms the interdependent relationship between competition and cooperation, the systemic cooperation that is part of all games.

But the central focus of New Games wasn’t game philosophy: it was the design and play of games themselves. The movement produced some extraordinary game designs. Take, for example, a game called Catch the Dragon’s Tail:

You’ll need a good-sized area for this event, clear of sudden pits and immovable oaks. About eight to ten people line up, one behind the other. Now, everyone puts their arms around the waist of the person in front of them. (You can’t be ticklish around dragons.) The last person in line tucks a handkerchief in the back of his belt. To work up steam, the dragon might let out a few roars—fearsome enough, we wager, to put Hydra to shame.

At the signal, the dragon begins chasing its own tail, the object being for the person at the head of the line to snatch the handkerchief. The tricky part of this epic struggle is that the people at the front and the people at the back are clearly competing—but the folks in the middle aren’t sure which way to go. When the other finally captures the tail, who’s the defeated and who’s the victor? Everyone! The head dons the handkerchief and becomes the new tail, while second from the front becomes the new head.

Catch the Dragon’s Tail purposefully blurs the lines between competition and cooperation. On the one hand, all of the players are cooperating to hold on to each other to become a single dragon. But at the same time, the front part of the dragon is chasing the rear part, with the people in the middle not given a clear role to play in the conflict. Catch the Dragon’s Tail makes playfully explicit the ways that players must work together even as they compete within the limited space of a game. Catch the Dragon’s Tail also embodies an important lesson for game design: all of our preconceptions about games can be questioned. Normally we might think that all players of a game must have a clearly defined goal, or that lines of competition must be sharply defined, or that a game with player coopera-
tion cannot also have vigorous competition—but Catch the Dragon's Tail debunks all of these assumptions. If nothing else, game design is about playing with ideas, and even seemingly fundamental ideas about competition in games are subject to playful intervention.

The Goal of a Game
In addition to competition and cooperation, another essential component of a game as a system of conflict is a goal. Goals are fundamental to games. In the explication of Centipede, Joust, and Gauntlet, goals figured into each form of conflict. At the outcome of a game, the goals are either reached or not reached, and this quantifiable outcome is part of our definition of games. Very often, it is a clear and quantifiable goal and outcome that distinguishes games from other play activities. Add a goal to informal play and usually you will have a game. Casual skiing for fun is a leisure play activity. But race your friend to the bottom of the mountain and suddenly you're taking part in a game.

A game's goal is defined by its rules and is tightly interwoven into the formal structure of the game as a whole. A game's goal is a central feature of its formal structure. When players come together to play a game, the goal is at the center of the magic circle, the pole that holds aloft the circular tent of the game while the players are inside the structure, at play with one another. The goal sustains their interest, their engagement, and their desire. Without a clear goal, meaningful game play is not possible; if players cannot judge how their actions are bringing them closer to or farther away from winning the game, they cannot properly understand the significance of their actions, and the game collapses into a jumbled heap of ambiguity.

A game's goal defines its endpoint; once it is reached, the game is over. In this sense, a game's goal is the death of play, the mark of the end, foretelling the moment the magic circle will disappear. There is a curious poetic quality to the struggle of game players as they make their way through the system of a game, playing to no end but the one provided by the game itself, even as their joyful pursuit of that end means the death of their pleasure. Until, of course, the next game begins.

Most games have an end in which one or more players achieve victory. However, in games such as Space Invaders, in which the game structure repeats itself with increasing challenge to the player, there is no single victorious endpoint. In this form of game, the goal is to play as long as possible or achieve the highest score. This formal structure heightens the sense of inevitable death. The player is living on borrowed time, staving off the inevitable end of a game that occurs when conditions of failure are met.

The space of possibility of a game is a plane stretched between two anchorage points: the beginning and the end of the game. The players journey from one end to another, making their way from the start to the finish. In a well-designed game that supports meaningful play, this journey between points should be taut and efficient, with every element contributing directly or indirectly to the larger experience.

In case this all sounds too goal-oriented, we must acknowledge that goals are not the only reason people play games. Play can be an end in itself, or a way to achieve social interaction, or affect cultural change. We address each of these motivations for play in later chapters. But seen as a formal system, the goal of a game needs to be recognized as a primary structure that shapes the game as a whole.
Case Study

Beating Loop

It is easy enough to state that the goal of a game is an important part of the overall design. But it is often very difficult to figure out the exact victory and loss conditions for a particular game. One play pattern can lend itself to different winning conditions, each one shaping the game experience differently.

LOOP is an online single-player game about catching butterflies. In the development of LOOP, the core game play was invented before a decision had been made concerning victory and loss conditions. The game development began with the idea of drawing lines to loop butterflies, scoring points by looping special groups of them (such as butterflies of the same color). The first prototype demonstrated the game interactivity but the start and finish conditions of the game had not been defined. Many options for the victory conditions were discussed and were narrowed down to three scenarios:

1. The player has to catch a certain number of butterflies to finish a level.
2. The player has to clear the screen of all the butterflies to finish a level.
3. There are no levels; the player has to keep on catching butterflies forever.

Each of these endpoints entailed different kinds of player experiences. For example, scenarios 1 and 2 presumed that the game proceeded as a series of discrete levels, whereas scenario 3 provided a single, continuous game, as in Tetris. There were other unsolved questions too: for example, in scenario 2, what would happen if a player were left with only one butterfly? And what was the loss condition? If the game did have levels, there needed to be victory and loss conditions for finishing a level; if it did not have discrete levels, there only needed to be a loss condition for the game as a whole. In considering the loss conditions, two primary schemes were proposed:

1. A player has a certain amount of time to attain the victory conditions.
2. Butterflies appear as the player proceeds. If too many appear, the game is over.

Many of these variations were playtested in different combinations. For the loss conditions, a time limit was selected, rather than the option of having the screen overrun with butterflies, because it gave the game a clearer sense of progression. In Tetris, it is clear that when the bricks pile up near the top of the screen, you are about to lose. In the prototype of LOOP, it was never quite clear when there were too many butterflies on the screen. A
The Level Playing Field of Conflict

Competition and cooperation, goals and struggle, victory and loss: how does it all add up? What are the general conditions of a game conflict? One core principle of conflict in games is that it is fair. Game conflict is impartial conflict: it is premised on the idea that all players have an equal chance at winning, that the game system is intrinsically equitable, that the game’s contest takes place on a level playing field, which does not favor one side over the other. Anthropologist Roger Caillois points this out in speaking about competitive forms of play: “A whole group of games would seem to be competitive, that is to say, like a combat in which equality of chances is artificially created, in order that the adversaries should confront each other under ideal conditions, susceptible of giving precise and incontestable value to the winner’s triumph.”

Why would a game strive so forcefully to create equality in this way? As our definition states, a game is an artificial conflict. The game structure creates an artificial arena, in which everything is removed except for the factors involved in the conflict. Chess is a context for intellectual strategic competition. In a gymnastics competition, only gymnastics skills matter.

In real life, the conflicts and struggles faced are never so clearly articulated and understood as in a game. The idea that players are entering into a fair conflict, where they won’t be fooled or tricked by the game itself, is a key component of the lusory attitude. Even though games may have elements of uncertainty, the structure within which that uncertainty plays out is known in advance. The qualities of rules themselves make this so. As we know from Defining Rules, rules of a game limit player action, are explicit, unambiguous, binding, and shared by all players. Within the magic circle, players experience a kind of equality and fairness that is not present outside games.

Is it really true that games strive to create spaces of equality, where only the play of the game can determine the winner of a game? Are games so pure and separate from the real world that no other factors possibly enter into the play? Generally speaking, no. But in this chapter we are analyzing games on a
"meter" that kept track of the number of butterflies and displayed how close the game was to being overcrowded could have been added, but that seemed to unnecessarily complicate the game interface.

Having made a decision about the loss conditions (a timer), it was determined that the game would be given discrete levels, with the goal of each level to catch a certain number of butterflies. It was decided that the number of butterflies did not decline over the course of a level; every time you caught a group, the same number of butterflies immediately flew in from the side of the screen. There were several reasons for these decisions. First, there was a desire to create a game that seemed full of possibilities. If the goal of the game was to clear the screen, then the game would have gradually emptied out as the player proceeded through a level. The game would have felt less dynamic and alive: given any set of butterflies, there would have been a single best solution for clearing the screen—and as a level proceeded, the possibilities for different actions decreased. The solution to keep the number of butterflies constant kept the game exciting and full of alternate strategies.

Even though a level ended when a quota of butterflies had been captured, it was possible to exceed the level's victory conditions by going over the quota. You could, for example, strategically capture enough butterflies so that you only had to capture one more to reach your quota—and then snatch a large group of them for your final loop, putting you well above the quota. Butterflies captured beyond your quota earned you large bonus points. The best players managed their butterflies carefully, creeping up on the quota and then scoring a large group at the end that took them well over the quota amount. This kind of careful play was only made possible by the particular structures that had been designed into the game.

Digital games tend to proceed as a series of levels and have loss conditions that end the game, LOOP being one such example. Deciding what constitutes success in a level and what ends the game are absolutely crucial game design decisions. Too often, game designers take these decisions for granted, following design conventions instead of inventing new ones. What about a game in which there are multiple loss or win conditions? What if the goal of the game is to lose as quickly as possible, or run out of points? Just as Catch the Dragon's Tail played with traditional ideas of competition and cooperation, your games can play with traditional ideas of winning and losing as well.
formal level, removed from consideration of their relation to
the outside world. And in a formal sense, yes, games are spaces
of pure conflict, separate from the outside world.

Given the artificial nature of games, is fairness possible? Does
the magic circle offer a truly level playing field? Cailliois thinks
not:

> As carefully as one tries to bring it about, absolute equality does not
> seem to be realizable. Sometimes, as in checkers or chess, the fact
> of moving first is an advantage, for this priority permits the favored
> player to occupy key positions or to impose a special strategy.
> Conversely, in bidding games, such as bridge, the last bidder profits
> from the clues afforded by the bids of his opponents. Again, at cro-
> quet, to be last multiplies the player's resources. In sports contests,
> the exposure, the fact of having the sun in front or in back; the wind
> which aids or hinders one or the other; the fact, in disputing for
> positions on a circular track, of finding oneself in the inside or out-
> side lane constitutes a crucial test, a trump or disadvantage whose
> influence may be negated or modified by drawing lots at the
> beginning, then by strict alternation of favored positions. 6

There is indeed a contradiction at work in the idea of equality
within games. As Cailliois points out, equality is something that
is sought after in games, but somehow never quite achieved—
at least, in the non-digital game examples he cites. What about
digital games? In some ways, they have an advantage when it
comes to creating a level playing field. The constrained context
of a computer system allows for greater control over the exact
conditions of play. On the other hand, the complex automated
nature of digital games can place players at some distance
from the rules. Players can easily grow suspicious of an unfair
network lag, a "cheating" AI, or a low processing speed "stutter."
This kind of player distrust, whether or not it is based in reality,
can ruin a game.

The magic circle is fragile, easily dispelled when players fail to
invest faith in the game. If your players feel that your game is
unfair, that it lacks a level playing field, it is unlikely that they
will want to play. Within the magic circle, a game is suspended
between the ideal notion of a level playing field and the reality
of inevitable unfairness, a reality that creeps into every game,
even while the magic circle's border holds it at bay. Perhaps
games do not take place on an absolutely level playing field.
But they are premised on the very real idea of fairness and
equality. This struggle is part of what gives games their vitality.

All games participate in this conflict between fairness and
unfairness, a struggle that reaches its climax in rule-breaking, a
phenomena explored fully in the next schema. But even within
a more limited game design context, establishing a sense of
fairness is crucial to successful game play. The following case
study looks at one example of this problem in detail.

**Pig Redux**

We first examined the game Pig in *Games as Systems of
Uncertainty*. In Pig, the goal of the game is to score points by
rolling a die and adding to your score until you reach 100
points. If you roll a 1, then you lose the points you earned that
turn and pass the die. Otherwise, you keep rolling to try to
increase your score. You can always decide to stop rolling, at
which point you add your current total to your overall score
and pass the die to the next player.

The game is very simple. But is it truly fair? Does the first player
have an advantage? Because Pig is about accumulating a
score, turn after turn, it does favor the first player: that person
has an added chance of reaching 100 first. If there are five play-
ers and on the tenth turn the player that went first scores 100
points or more, that player wins. But some of the other players,
who only got to play nine rounds, might have reached 100 if
they had been allowed a tenth turn as well.
Pig embodies the classic game design problem of creating a level playing field. Ideally, every player should have an equal chance of winning. So what is the solution? There are a few possible game design adjustments. One solution is that the winner of the previous game gets to go first, as an added reward. But this does not solve the problem of deciding who goes first in the very first game. Should the winner be rewarded in this way? Doing so creates a positive feedback loop, which might unbalance the game. Should the player with the lowest score go first? Neither of these solutions create fairness for all players.

Another solution is using a random die roll to determine player order. The player that rolled the highest number goes first. Even though a great many games use this method, it is not necessarily the best solution to the problem. Will it feel fair to all of the players? For example, because play proceeds around the circle of players, the player that rolled the lowest number may end up as the second player, if that player is sitting next to the person that rolled the highest number. In any case, even if the order of players is randomly determined, the player moving first still has an advantage over the other players, and the inequality remains.

Yet another solution would be to play the game a number of times equal to the number of players at the table, rotating which player goes first. If there were ten players, they would play the game ten times. Each time, a different player would go first. Players would either add up their scores for a grand total or the player who won the most times would be named the overall winner. This solution works mathematically to equalize the game, but it suddenly transforms the casual experience of Pig into a structured tournament. What if you only want to play a game or two and not an entire series of games?

We borrowed the game of Pig from Reiner Knizia’s book *Dice Games Properly Explained.* Knizia notes this very inequality and suggests the following as a variation: when a player reaches 100, all of the rest of the players get to roll once more and finish the round. If more than one player ends up exceeding 100, the player with the highest score wins. This is better than any of the previously proposed adjustments, but even this well-designed solution is flawed. In Knizia’s solution, it is best not to be the first player to reach 100, because all of the other players know exactly the score that they need to win, and they will push their luck in order to beat the player that is about to win. It is actually best if you roll last during the final round. Because the player that went first at the beginning of the game is mathematically more likely to be the first to reach 100, it ends up being a slight disadvantage to be the first player.

Even in a very simple game such as Pig, there is no perfect solution that offers absolute equality for all players. But luckily, players are not perfectly rational beings. They are human, and the best solution is not necessarily the best mathematical answer to the question of equality, but the one that *feels* right within the context of a game. Absolute equality, like pure randomness in a computer algorithm, may be a myth. But as long as the *feeling* of equality persists within the game, players will keep the faith and enter into the magic circles you design for them.

As a final thought, is fairness itself something that can be put at play in a game? We have suggested that other components of game conflict, such as competition and cooperation, or achieving game goals, could be challenged through innovative game designs such as *Catch the Dragon’s Tail.* Does this extend to the level playing field of a game as well? Perhaps. But it is a very complicated question: In the next schema, *Breaking the Rules,* we do our best to answer it.
Further Reading

*Homo Ludens*, by Johann Huizinga (see page 99)

**Recommended:**
Chapter 1: Nature and Significance of Play as a Cultural Phenomenon

*The New Games Book*, by Andrew Fluegelman and Shoshana Tembeck (see page 21)

**Recommended:**
"Creating the Play Community," Bernard DeKoven
"Theory of Game Change," Stewart Brand

Notes

4. Ibid. p. 47.
- Conflict is an intrinsic element of every game. The conflict in a game emerges from within the magic circle as players struggle to achieve the goals of a game.

- Game conflict comes in many forms. Conflict can be individual or team-based, cooperative or non-cooperative, direct or indirect. Many games mix and match forms of conflict within a single game structure.

- The forms of conflict occurring within a game are a direct outgrowth of its rules. One way of framing a game's space of possibility is that it is a space of possible forms of conflict. Players take part not just in the forms of conflict that the game design prescribes, but will also find their own forms as well.

- All games are competitive in that players struggle against each other or against a game system as they play. Without this sense of competition, meaningful play would be difficult because players would not be able to judge their progress through the space of possibility of a game.

- All games are cooperative, in that playing a game means engaging with the shared meanings of the game, "speaking the language" of the game with other players in order to play.

- Systemic cooperation refers to the fundamental, discursive cooperation that is intrinsic to all games. Player cooperation refers to games in which players all work together to achieve the goal. Not all games exhibit player cooperation.

- The goal of a game is a fundamental element that shapes the game's formal structure. The goal is at once that toward which players strive, while also that which represents the end or symbolic death of a game.

- Shaping victory and loss conditions is an important component of game design. Victory and loss conditions directly shape the possible outcomes of a game.

- Game conflict is premised on a level playing field where all players have an equal chance of winning. A truly equitable game is virtually never possible in the real world, creating an intrinsic tension in regards to the fairness of any game. Players will generally refuse to play a game they perceive to be unfair.
[The rules that players verbalize] are an idealized set of rules—they are the rules by which people should play rather than the ones by which they do play....we may have to know two sets of rules: the ideal ones and those by which the ideal rules are applied, misapplied, or subverted.—Kenneth Goldstein, "Strategies in Counting Out"

When you have to win, you’re willing to break whatever rules you can if that would help you get closer to the goal. When you have to win, you’re not concerned with fairness, feeling, the community, or even play. When you have to win you can’t leave the game until you have finally, ultimately won.

What’s amazing to me about all this is that the game itself doesn’t change. The rules and the conventions are the same. But the manner of playing the game is completely different.—Bernard DeKoven, The Well-Played Game
Introducing Rule-Breaking

This schema opens with a pair of quotes from two thinkers we have heard from before. Folklorist Kenneth Goldstein first appeared in the schema on Uncertainty, where he looked at the ways that children subvert the ritual of counting-out through a number of subtle and devious strategies, such as adding an extra “eenie-meenie-minie-moe” in order to avoid becoming “it.” We introduced Bernard DeKoven in the previous schema on Conflict as a leading figure in the New Games Movement.

Goldstein points out that although games have rules, they should be considered to have two sets of rules: the ideal rules of play and the actual rules of play, which sometimes misapply and subvert the ideal rules. DeKoven comes at the same set of issues from a different point of view. He points out that some players are so motivated to win that they disregard usual notions of fairness. What seems to intrigue DeKoven the most is that such opposing styles of play can occur alongside normal play within the same game structure.

Whether we are talking about ideal rules versus actual rules or honest players versus cheating players, both writers point to an important game phenomenon. So far in this book, we have described game players in an almost naïve way: we have assumed that every player is an earnest player, carefully and honestly playing by the rules. Although this does describe many game players, it is certainly not true of every single one. Take the children that Goldstein studied in his analysis of counting-out games. In manipulating rhymes in order to achieve certain desired results (he is going to be “It,” not me!), what were these players actually doing? Were they stretching and altering the rules of counting-out in order to win? Were they cheating at the game? Or were they simply playing the game very well? This final formal schema, Breaking the Rules, takes a direct look at how players bend, cheat, and break those carefully crafted systems of rules that we have so thoroughly investigated in the last several chapters.

In so many different ways, breaking the rules seems to be part of playing games. Whether it is trying to sneak in a foul while the referee isn’t looking, altering a board game to play with a special set of “home rules,” or making use of an ace of spades hidden up your sleeve, reconfiguring, breaking, and ignoring the rules seems to be an intrinsic part of games themselves. But what guides a player to break the rules? What is the effect of rule-breaking on game play? How does a game’s design either encourage or discourage players from breaking the rules? Lastly, can rule-breaking be used as a creative strategy for game design? We investigate these questions in the following pages.

Kinds of Rule-Breaking

Rule-bending and rule-breaking manipulate the structure of a game. To cheat or transgress in a game means to break the rules, to have a relationship to the formal system that is different than the relationship that the formal system itself presupposes and endorses. In considering the ways that game rules are broken, we can divide players into different player “types.” Each type of player is defined by his or her relation to the formal systems of a game, along three related axes of behavior and attitude:

- The rule-breaking player’s adherence to the rules
- The rule-breaking player’s interest in winning
- The rule-breaking player’s degree of lusory attitude

Player Types

The Standard Player: This player type is a “standard” and honest game player that plays the game as it was designed to be played, following the rules and respecting their authority.

The Dedicated Player: This close cousin of the standard player studies the formal systems of a game in order to master and perfect his or her play of the game, often finding and exploiting unusual strategies in order to win. Examples: professional athletes, hardcore gamers.
The Unsportsmanlike Player: This third type of player follows the rules of a game, but does so in a way that violates the spirit of the lusory attitude. Examples: The older sibling that never lets the younger sibling win, or the baseball catcher that tries to distract the batter's concentration at the plate.

The Cheat: The cheater, unlike the other kinds of game-players, actually violates the formal rules of the game, but does so in order to win the game. Example: The hide-and-seek player that peeks while the other players are hiding.

The Spoil-Sport: This kind of game player is hardly a player at all. Unlike the cheat, the spoil-sport refuses to acknowledge the magic circle of the game and does not care about winning or about following the rules. Example: The frustrated Twister player that ruins a game by pushing over the other players.

In the sections that follow, we describe each kind of player in more detail. But before moving on, it is important to recognize that these categories are neither fixed nor mutually exclusive. The boundaries between them are quite fuzzy, and often contextual. A player that is a dedicated hardcore gamer among gamer friends might be seen as an unsportsmanlike, overly competitive "power gamer" when playing a game with more casual players. Likewise, a player might shift between categories over time, or even within the course of a single game. Despite the fluid boundaries between them, however, these categories provide a useful typology for understanding the ways players stretch, bend, and break game rules.

Standard Players
The standard player is the test case against which all other types of players are contrasted. The standard game player attempts to follow the rules as best he or she can, respecting their authority and honoring the limits they set. In terms of rules, goals, and possession of the lusory attitude, the standard player is a most law-abiding citizen.

Do most players fit this description? Actually, they do. The magic circle is fluid, but when most players play a game, especially a game with other players that can be seen face-to-face, they respect the rules and play the game from beginning to end. Why is face-to-face interaction important? A game is a kind of social contract. The presence of other players is important to maintaining the authority of the magic circle, because if a group of players are all obeying the rules, they implicitly police and enforce proper play. Why? Because if they have decided to invest the game with meaning in order to play, they all have a vested interest in maintaining the level playing field of conflict created by the rules. This does not mean that most players are mindless slaves to the rules of a game, but generally speaking, looking across all phenomena of games, players do follow the rules. If this were not the case, then cheating at games would be the rule and not the exception.

You may well disagree with our contention that most players do not break the rules. One could also take the position, for example, that cheating exists in all players, that the force of game-playing desire that drives a player to win contains the seeds of cheating. Cheating, in this view, would be an intrinsic aspect of game-playing, even if it did not always rise to the surface in the form of genuine rule-breaking. But whether the "standard player" is really the majority case or a fiction that doesn't exist in the real world, the notion of the "standard player" is still important. The idea that there is a standard player, a game player that earnestly follows the rules without trying to bend and break them, provides the backdrop against which less rule-governed styles of play can be understood.

Dedicated Players
The next type of game player is the dedicated player. The dedicated player is really more of a special case of the standard player than a completely different player type. The dedicated player desires to become an expert at a game, and diligently studies the rules of play in an attempt to maximize the chances
of winning. Whereas standard game players exhibit a desire to win and an interest in the rules of a game, dedicated players apply themselves to this task with a certain kind of zeal, to a degree that more casual players might not find enjoyable. If the game permits, dedicated players tend to practice their play, testing out strategies and perfecting their knowledge of the game.

A typical Las Vegas tourist who wants to enjoy BlackJack might play a few games here and there, browsing different casinos and tables, relying on intuition to guide him as he plays. A dedicated BlackJack player, on the other hand, won’t merely play a few casual rounds of the game, but is likely to study a BlackJack “system” or two and implement it diligently in play, finding tables with advantageous rule variants, counting cards during play, and spending long hours at the BlackJack table in order to balance out his odds of winning. The difference between dedicated and standard players is a matter of degree, not kind.

Recall that the differences between types of players is drawn along three axes: their relationship to the lusory attitude, their respect for the authority of the rules, and their interest in attaining the goal of the game. Within each of these categories, dedicated players resemble standard players. But dedicated players have a deeper engagement with the game, a greater zeal for play. It is more important for dedicated players to win, and in order to do so, they will generally learn and master the rules of a game. At the same time, dedicated players tend to invest the magic circle with more authority, because of the value of their investment in the game as a whole. They possess extra amounts of the lusory attitude, relishing the inefficiencies of games as important challenges to overcome as proficiently as possible.

Who are dedicated players? Professional athletes and professional gamblers—those that make their living as game players. So are so-called “hardcore gamers,” from grognard historical wargamers to deathmatch clan leaders with tricked-out custom PCs. In general, dedicated players require more depth and complexity, a richer space of possibility in their games. This is why non-gamers often find the gaming fare of hardcore gamers bafflingly complex and unapproachable.

Dedicated players tend to play with a zeal that often puts off less dedicated players, who sometimes wonder if dedicated players are taking the game just a bit too seriously. The dedicated BlackJack player we described, who might spend most of a Las Vegas vacation at the BlackJack tables, might seem incomprehensible to the casual, standard player, who looks at games as a form of relaxation and leisure. A casual player does not wish to spend so many waking hours inside the magic circle of a game.

There is a very fuzzy line between dedicated game players and standard game players, and the difference is often contextual. Among your dedicated bowling buddies, you might fit in just fine as a standard player, scoffing at the league players that wear matching shirts and play the game “too seriously” to have fun. But when you end up in a game with a group of beginners who want to abandon a match in the middle to go see a movie, you might find yourself being accused of playing “too seriously” when you demand that they stay to the tenth frame and finish what they started.

As game designers, it is important to understand the range of player types that encounter your game, and the kinds of relationships they have to the rules, goals, and magic circle that your game delineates. Some games clearly appeal to both standard and dedicated players, such as Scrabble. Scrabble is often played as casual family fare, but it also supports an international tournament culture of hardcore players. Other kinds of games tend to attract one kind of player over another. The players that enjoy the low-pressure, exploratory pacing of Myst are generally not the same kind of dedicated player audience that would spend the many hours required to understand and
master Myth: The Fallen Lords. There is a similar divide off the computer between players of party games such as Pictionary and fans of complex wargames and role-playing games.

The first two categories of game-players—standard and dedicated players—are not ultimately rule-breakers. They are "classical" game players, the kinds of players for whom designers usually design games, loyal functionaries of the rules. Like standard players, dedicated players are indeed rule-abiding. But as we'll see soon enough, even though they seem more invested in the magic circle of a game, their dedication takes them one step closer to actual rule-breaking.

**Unsportsmanlike Players**

The third type is the unsportsmanlike player. Unsportsmanlike players do anything they can to win. They try to find shortcuts to victory, exploiting the rigidity of the rules to locate holes that they can slip through to end up ahead. An unsportsmanlike boxer, for example, might constantly grab at the ropes or go into a clinch whenever the opponent advances aggressively. Note that the boxer stops short of actually violating the rules of the game. In fact, some might consider this approach a valid strategy for Boxing. But somehow, the unsportsmanlike boxer violates the spirit of the contest of Boxing, marring the purity of the battle between the athletic skills of the two players.

Unlike standard and dedicated players who generally engage openly with the "fun" quality of play, there is something negative about unsportsmanlike behavior. The unsportsmanlike player turns the special zeal of dedicated players into something that seems to run counter to the joyful nature of play and games. An unsportsmanlike player is not a cheat. The unsportsmanlike player does follow the rules of a game, but in a way that violates the spirit of the game. By attempting to shortcut the challenges of a game, the unsportsmanlike player refuses to surrender completely to the lusory attitude, in which the inefficiencies of play are readily accepted.

Unsportsmanlike behavior is a violation of the "unwritten" rules of a game, the implicit rules that are not actually written out, but are observed by all players. This is how the unsportsmanlike player "technically" avoids designation as a cheater, while still failing to completely respect the lusory attitude. One of the implicit rules of Tic-Tac-Toe we discussed in *Rules on Three Levels* is the implied time limit between turns. Even though the operational rules do not mention a time limit, the idea that a player must take a turn in a "reasonable" amount of time is an implicit rule of the game. Imagine an unsportsmanlike player that is about to lose a game of Tic-Tac-Toe, but refuses to take a turn. The player might state that he is "thinking" about his next move, and claim that because the rules do not state a time limit, he can take as long as he wants, even years, before he has to move. This kind of behavior, although not violating the operational rules, clearly violates the spirit of the game.

**Degenerate Strategies**

Dedicated and unsportsmanlike players have particular ways of engaging with the system of a game. One common behavior these player types exhibit is to make use of degenerate strategies or exploits. We first encountered degenerate strategies in *Games as Game Theory Systems*. A degenerate strategy is a way of playing a game that takes advantage of a weakness in the game design, so that the strategy guarantees success.

Degenerate strategies often appear in complex games, where the numerous permutations of play sometimes afford shortcuts in the space of possibility. For example, you are playing a real-time strategy game against the computer and you realize that the program's AI does not handle pathfinding well. (Pathfinding refers to the aspects of the program that plot navigational paths for the computer-controlled characters through obstacle-filled terrain.) Whenever the computer-controlled troops move around obstacles, they begin the march in formation but end up disorganized, with individual units trapped in irregularly
shaped pockets of the terrain. It is not difficult for you, however, to make the small corrections necessary to keep your units together. If you decided to take advantage of this weakness by strategically leading the computer-controlled opponents into obstacle-filled parts of the map, you would be using a degenerate strategy.

Taking advantage of the game's weakness in this way would not exactly constitute cheating, but it does exploit the game's structure as a means of winning. Although games are not designed to be exploited by players, what makes a degenerate strategy degenerate is not just that it goes against the intentions of the designers. Using an exploit is a way of playing that violates the spirit of the game, similar to taking advantage of the implicit rule governing time between Tic-Tac-Toe turns.

Degenerate strategies appear in non-digital games as well. In early editions of Magic: The Gathering, certain card combinations were simply too powerful and could destroy a player on the first turn, before a match had a chance to develop. Wizards of the Coast, the publishers of the game, declared certain cards "officially" illegal, most notoriously the Black Lotus card, in order to keep this kind of play experience in check. In regulated tournament play, the outlawed cards were not used. But in more casual games, players continued to include them in their decks for years.

Why isn't using a degenerate strategy considered cheating? Degenerate strategies take advantage of weaknesses in the rules of a game, but do not actually violate the rules. What kind of player would play in this way? The answer is both a dedicated player, who is overzealously seeking the perfect strategy, and an unsportsmanlike player, who has found a hole in the rules to exploit, even though he understands that he is not playing the game the way it was intended. These two kinds of players can both make use of degenerate strategies, depending on the context.

The difference between a dedicated player and an unsportsmanlike player is the degree to which the player subscribes to the luxurious attitude. Dedicated players follow rules on all levels. Unsportsmanlike players follow the operational rules, but they do not follow all of the implicit ones. Dedicated players loyally uphold the magic circle of a game, but unsportsmanlike players fail to do so, occasionally stepping just outside its borders in order to bend the rules.

Often, whether or not a degenerate strategy is a "proper" way to play depends on how the game experience is framed. When it was discovered that Pac-Man could be played by memorizing patterns of movement instead of through improvisational moment-to-moment tactics, player reaction fell into two camps. Some frowned on using memorized play patterns as a violation of the spirit of the game. Other players, however, capitalized on patterns in order to get higher scores. These pattern players did not consider themselves to be unsportsmanlike at all: they saw themselves as dedicated players who had simply found a better (and more demanding) way to play the game.

One more example: remember the hypothetical fighting game from our earlier investigation of degenerate strategies? The game could be beaten by using one technique over and over, rather than exploring the carefully orchestrated system of fighting moves created by the game's designers. It could be said that the player making use of this degenerate strategy is behaving in an unsportsmanlike manner, improperly playing the game, sacrificing "fun" in exchange for a shortcut to victory. It could also be said, however, that the exploit was being used by a dedicated player who had "solved" the fighting game like a puzzle. As with the Pac-Man pattern players, instead of playing the game the way it was designed to be played, the dedicated player simply invented a new method of interaction. This is arguably an example of transformative play, an important game phenomena we will investigate in chapters to come.
Whether or not a particular degenerate strategy is considered proper is often contextual. For example, the use of the single-technique exploit to beat all of the computer opponents in our hypothetical fighting game might be admired by a group of players for its elegance. On the other hand, if the degenerate strategy were used against other human players, fighting bouts would devolve into uninteresting games, with both players relying on the one exploitable technique again and again. In this social context, the exploit would be frowned upon as unsportsmanlike behavior, a violation of the implicit rules and the enjoyable spirit of the game. The meaning of a game action, even if the action is the selection of a general strategy, is always influenced by the context in which it occurs. In a social context, the exploit unbalances the level playing field of conflict and shrinks the space of possibility to a very narrow range, threatening the meaningful play of the game.

Degenerate Strategy Ecosystems
As a rule of thumb, you want to be on the lookout for degenerate strategies and keep them out of your game. The ability to win a game by playing in a singular way demonstrates a poor game design, a space of possibility with an unintended, limiting short-circuit. There is, however, an extremely fuzzy line between degenerate strategies and imaginative ways to play a game. There is something exciting about having players explore the space of possibility of your game, rooting around for new strategies and new ways to play. If the game is complex enough and the community of players is large enough, degenerate strategies that do emerge can be countered by new strategies created specifically to oppose the exploits. An ecosystem emerges from the community, in which different styles of play compete for dominance.

In real-time strategy (RTS) game player communities, for example, players constantly look for ways to get ahead on the rankings boards. Command and Conquer, like most RTS games, was intended to emphasize steady planning and gradual development. But over time a degenerate strategy evolved called the “tank rush.” Instead of slowly building up forces, a player using the “tank rush” strategy could quickly create a group of tanks and wipe out his opponent’s base camp in the early game, before his opponent had a chance to prepare his defenses. Although the tank rush degenerate strategy ruined the games of many players that desired a more typical long-term conflict, it also spawned new kinds of defensive strategies. The introduction of a degenerate strategy enlarged the overall space of possibility of the game.

Although some player communities are resourceful enough to create their own antidotes to degenerate strategies, it is often necessary for the designers to step in and correct the breach themselves, as in the case of Magic’s Black Lotus card. With popular games, play strategies sometimes evolve in a way that necessitates a refinement of the formal structure, like a gardener pruning branches of a tree to improve the overall health of the plant. The process of degenerate strategy correction is ultimately part of the iterative process of game design. One game that has undergone constant refinement is professional Basketball in the U.S.

Over the last several decades, Basketball has undergone a number of rule changes. For example, in the 1960s and 1970s, most of the action took place right under the basket, where the chance of scoring was greatest. Play was dominated by tall players that could control this space with the greater offensive and defensive capabilities their height provided. Two rules were introduced that shook up the play of the game and defused degenerate strategies that were beginning to crop up. The three-point line incentivized players to play away from the basket, daring them to risk a longer shot in order to gain an extra point. At the same time, the three-second rule, which kept offensive players from spending more than three seconds
parked in the paint under the basket, helped unclog the scoring zone traffic jam. The end result of these two rules is that quick players who could weave into the zone and out from under the basket, perhaps darting back to the three-point line to take a shot, became more important than static, towering giants. The space of possibility of the game expanded to include not just more diverse strategies of play but more diverse physical types of players to implement them.

Basketball has plenty of other rules that have been modified over time as well, from the introduction of dribbling near the beginning of the century to the more recent innovation of the shot clock and the back-and-forth controversies over zone defense. In his essay “The Heresy of Zone Defense,” cultural critic Dave Hickey eloquently addresses this process of rule iteration:

The “illegal-defense rule” which banned zone defenses, however, did more than save the game. It moved professional basketball into fluid complexity…leaving the college game with its zoned parcels of real estate behind. Initially, it was feared that this legislated man-to-man defense would resolve competition in terms of “natural comparative advantage” (as an economist might call it), since if each player is matched up with a player on the other team, the player with the most height, bulk, speed, or quickness would seem to have a permanent advantage. But you don’t have to guard the same man all the time; you can switch, and this permission has created the beautiful “match-up game” in which both teams run patterns, picks, and switches in order to create advantageous situations for the offense or the defense—to generate shifting interplay.¹

Degenerate strategies can lead to iterative design. It is beautiful to think of a game design as a design in process, which can grow and evolve over time, remaining fresh in response to changing needs and invented strategies. As the athletic abilities of players and the strategic acumen of coaches tested the limits of the system, the rules of Basketball were refined. Changes in rules maintained the tautness of the space of possibility while allowing players to move freely within it. Even today, regular changes in the rules continue to keep the game fresh. The act of rule-modification itself—by game designers, players, or administrative bodies—is an important kind of game design which will be addressed further in the pages to come.

Cheats and Spoil-Sports

The player who trespasses against the rules or ignores them is a “spoil-sport.” The spoil-sport is not the same as the false player, the cheat; for the latter pretends to be playing the game and, on the face of it, still acknowledges the magic circle… the spoil-sport shatters the play-world itself. By withdrawing from the game he reveals the relativity and fragility of the play-world in which he had temporarily shut himself with others.—Johann Huizinga, Homo Ludens

The final two categories of players are the cheater and the spoil-sport. Up to this point, we have had to look very carefully at the players’ behavior to decide whether or not they are violating the formal system of the game and are actually breaking the rules. With these final two categories of players, things become more explicit.

What defines the cheating player? The cheater breaks rules. Unlike the unsportsmanlike player, who merely violates the implicit, unspoken rules of a game, the cheater transgresses the operational rules, the actual rules of play. The cheater is the player that secretly moves a piece when her opponent looks away from the board, the player that steals Monopoly money from the bank and hides it for future use, the player that uses a non-regulation golf ball in a tournament in order to gain a little more distance. The cheater surreptitiously takes actions that are not proscribed by the rules, in order to gain an advantage.

Does cheating destroy a game? The unexpected paradox of cheating is that, as Huizinga points out, the cheater is still in
some way playing the game. The cheater breaks rules, but only to further the act of winning. So while the cheater sheds enough of the lusory attitude to disrespect the authority of the rules, the cheater still has faith in the sanctioned conflict of the game: being the victor still has meaning to the cheater. This may seem like bizarre behavior. What is the point of hanging onto the authority of the quantifiable outcome when the prescribed steps for getting there are thrown out the window?

It turns out that the cheater is only one step removed from the dedicated player. It is possible to sympathize with a cheat, for he or she too has a passion for winning. A cheater craves winning, but too much, committing crimes in order to attain the object of desire. Of course, the motivations for cheating are many. Cheating might grow from a desire to beat the game system itself, to show up other players, or to reap rewards of glory external to the game. But no matter what the psychological motivation for cheating, all cheating behavior shares a particular set of formal relationships to rules, goals, and the magic circle.

The spoil-sport is the category of player furthest from the standard player. As game designer Mark Prensky explains, "What spoils a game is not so much the cheater who accepts the rules but doesn’t play by them (we can deal with him or her), but the nihilist who denies them altogether." The cheater breaks the rules but remains within the space of play. The spoil-sport is more destructive, refusing to acknowledge the game altogether. The spoil-sport is the frustrated player that knocks all of the pieces off the Chess board, the player that reveals the hidden information of Charades, the player that answers when it isn’t his turn, the player that hacks into the game database to erase all of the player records. The cheater is a conniving actor, a spy within the magic circle, carefully pretending to obey all of its regulations even as he breaks them. But the spoil-sport has no such compunction. His destruction of the game does not require concealment, because the rule structure that would condemn his action as illegal is exactly the authority the spoil-sport wishes to undermine.

When a set of Chess pieces are placed in their proper positions on the board and a game begins, the pieces gain meaning. But if, during a game, the action of a spoil-sport wipes the Chess pieces from the board, meaning is violently erased. Removed from their grid positions, the Chess pieces merely represent a collection of scattered figurines. The spoil-sport returns the game to its pre-game state as a collection of parts, no longer the embodiment of the space of possibility set out by the rules of the game.

The spoil-sport, more than any other kind of player, demonstrates the fragility of the magic circle. Not bound by a faith in the game, an interest in the lusory attitude, a respect for the rules, or even a concern for the outcome, the spoil-sport is the representative of the world outside the game. Armed with a powerful lack of belief, the spoil-sport has no qualms about ruining the play of others. The cheat may hack into a multiplayer deathmatch to up his ping time and secretly improve his play performance. But the spoil-sport will unleash a virus that brings the game servers to a halt, making play impossible for all players.

Five Player Types Compared

On the following page is a table that summarizes the five kinds of players discussed in this schema. Several fascinating patterns arise when we compare player types in this way. The slippery slope between the dedicated player and the cheat becomes particularly clear. An enthusiasm for playing a game can quickly become a zealot winning-for-its-own-sake, which can lead to unsportsmanlike behavior and outright cheating. In their shared investment in the outcome of the game, players and cheaters have a great deal in common.
<table>
<thead>
<tr>
<th></th>
<th>Degree of lusory attitude</th>
<th>Relationship to rules</th>
<th>Interest in winning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Player</td>
<td>Possesses lusory attitude</td>
<td>Acknowledges authority of rules</td>
<td>Typical interest in winning</td>
</tr>
<tr>
<td>Dedicated Player</td>
<td>Extra-zealous lusory attitude</td>
<td>Special interest in mastering rules</td>
<td>Intense interest in winning</td>
</tr>
<tr>
<td>Unsportsmanlike Player</td>
<td>Sometimes resembles the Dedicated player, sometimes resembles the Cheat</td>
<td>Adherence to operational rules, but violates implicit rules</td>
<td>Intense interest in winning</td>
</tr>
<tr>
<td>Cheat</td>
<td>Pretends to possess lusory attitude</td>
<td>Violates operational rules in secret</td>
<td>Intense interest in winning</td>
</tr>
<tr>
<td>Spoil-Sport</td>
<td>No pretense about lack of lusory attitude</td>
<td>No interest in adhering to rules</td>
<td>No interest in winning</td>
</tr>
</tbody>
</table>

It is sometimes difficult to identify exactly when an instance of cheating is a true transgression of the magic circle or merely part of the play of a game. Is hacking into an online server to inflate a high score on a public ranking board cheating? The transgression is not taking place within the magic circle of a particular game, but it certainly demonstrates an overly serious interest in the act of winning. How about fouls in sports? And what about games that encourage rule-breaking as part of their play? Where do they fit into our understanding of formal transgressions? We end this chapter by looking at a series of games that incorporate rule-breaking into the game design itself.

**Sanctioned Violations: Professional Sports**

In most games, rule violations threaten to destroy the magic circle. However, there is one category of game in which rule-breaking by players and punishments for violations of the rules are an important part of the overall game structure: professional sports. Double-dribbling in Basketball, icing in Hockey, using hands in Soccer—these are all rule-violations, but they are violations that are punished within the game itself, in ways that let the play continue. It is expected, and even anticipated that these kinds of events will occur in a sports game. It would be extremely unusual for an entire Basketball game to occur without a single foul being committed.

What is interesting about the way that sports handle rule-breaking is that there is always a sliding scale of severity for different rule violations, and often extra punishment for repeated offenses, as when a basketball player "fouls out" and cannot play in a game after committing six personal fouls. A single foul might be the result of an "honest mistake" and is therefore treated somewhat lightly. Six fouls, on the other hand, creates a pattern of rule-breaking behavior, and the player is ejected from the magic circle entirely. Sports referees, as extensions of the formal system of a game, have authority to decide when violations occur and how to interpret the rules to mete out punishment. For example, referees generally have the authority to throw players and coaches out of games if their behavior becomes too extreme.
When rule-breaking becomes sanctioned, as it is in sports, a whole new layer of implicit rules enters into the space of play. Whereas it is considered aggressive play (and a foul) to elbow an opponent on a Basketball court, it is truly bad sportsmanship to punch that same opponent in the face. As rule-breaking is integrated into a game, it is incorporated into the space of possibility. Depending on the particular game, players may strategically transgress rules, accepting a short-term punishment for a long-term strategic or psychological advantage.

This intentional brokering of rule-breaking can be quite complex. In Basketball, the players can attempt to “draw fouls” from opponents. This risky practice can result in the player who is attempting to draw a foul committing a foul himself. Players who charge the basket on offense hoping to be fouled on their way to the hoop are often called for “charging,” an offensive foul that results in the loss of the ball for the offensive team.

In professional sports, the complex system of violations and punishments within a game is also reflected in the professional legislative bodies which can sanction penalties for larger violations. Outside the scope of an individual game, these organizations govern more serious offenses. If a professional athlete is found to be fixing games or is convicted of a criminal act, he can be banned from the sport for life by the game’s professional body.

Why is there so much attention to breaking the rules in sports, particularly professional sports? One answer is the nature of athletic game play. On a Chess grid, there is little or no ambiguity about which square a piece occupies; a Chess player will not gain an advantage by having a little corner of his Rook peak into an adjacent square. But in the infinitely granular space of the real world, milliseconds and millimeters can mean the difference between winning and losing. The runner does not want to start running before the starting gun fires, but springing forward as close to that moment as humanly possible will certainly offer an advantage. As a result, many false starts occur in races. Most sports fouls are motivated by an attempt to maximize an offensive or defensive advantage.

In looking for a motivation behind the prominence of rule-breaking in sports, we must also acknowledge the economic component of the games. A great deal of capital is connected to professional sports, from player salaries to ticket sales to network advertising. When the external stakes of a game are high, it is especially important to maintain and enforce the level playing field of conflict. The premise of a professional sport, even more than with most games, is that it is being played fairly. This emphasis on fairness extends naturally to its opposite: an emphasis on breaking the rules.

**Sanctioned Cheating: Illuminati**

For a different approach to the integration of rule-breaking into a game, we turn to Illuminati, a humorous strategic tabletop game based on the Illuminatus books by Robert Anton Wilson. In the game, players take on the role of all-powerful Illuminati, the shadowy power brokers pulling the strings behind world governments. The original edition of Illuminati contained an optional set of rules for cheating:

*Cheating:*

Some fiendish people think Illuminati is even more fun when nothing, not even the bank, is sacred. In this variant of the game, most forms of cheating are permitted.

*Exceptions:*

- You may *not* tip over the table or disarrange opposing power structures.
- You may not bring in counterfeit money or money from other sets.
- You may not cheat on the amount of money drawn from the bank during setup or the income phase (this would slow things down too much).
- Anything else goes. Anyone caught in the act must undo that cheat. There is no other penalty.

Suggested methods for cheating include:
- Accidentally misread the dice.
- Steal from the bank (other than during the income phase).
- Lie about the amount of power or resistance your groups have.
- Stack the deck or peek ahead.
- If anyone leaves the table, anything goes!
- We recommend you play the cheating game only with very good friends or with people you will never see again.³

These "rules" for cheating in Illuminati provide a fascinating example of the relationship between rule-following and rule-breaking. Normally, cheating is considered something that runs counter to the spirit of the game rules. But in Illuminati, the sanctioned formal system of the game actually contains rules for cheating.

Illuminati's rules for cheating are different than rule-breaking in professional sports. In sports rule violations, most fouls are committed by players performing as close as possible to the limits of what the rules allow. In the real-world context of athletic performance, sometimes players miscalculate and end up breaking a rule. But in Illuminati, the suggested modes of cheating focus explicitly on player deception. The rules above directly suggest out-and-out, down-and-dirty cheating. The rules are not descriptions of penalties for fouls: they are proscriptions for different ways to cheat! In fact, there is no explicit penalty for being caught cheating, other than undoing the effect of the cheat.

Sanctioned cheating can easily destroy a game. Are Illuminati's "cheating rules" a recipe for anarchy, or are they a well-designed extension of the rest of the rulebook? It seems like a contradiction that the rules themselves contain suggestions for transgressive play. But a close look at the rules reveals the care taken in crafting this section of Illuminati's formal structure.

Illuminati places numerous formal restrictions on the scope of possible cheating. Forbidding players from tipping over the table (a classic spoil-sport action) lets players know that they cannot completely disrupt the game for the other players. Keeping players from inflating their income ensures that the game will not get too bogged down in mathematical squabbling. Permitted cheating focuses on keeping the rule-breaking play constrained, so that things do not swing too wildly outside the magic circle. For example, the rule that keeps players from smuggling money in from other sets of the game performs a number of regulatory functions. It keeps the designed economy of the game intact, while not letting players with "outside" resources (such as their own copy of the game) from gaining an unfair advantage. The result is that even with cheating, the game is contained within the magic circle, so that all of the players have an equal chance of being skillful cheats. The magic circle is such a strong focus of the cheating rules that when a player actually leaves the physical space of the game by getting up from the table, the rules state that "anything goes." Players are clearly discouraged from exiting a game in progress.

In addition to formal restrictions, the cheating rules go so far as to shape the lusory attitude of the players that might want to use them. The statements that begin and end the cheating rules place it within a particular context. The opening statement, which implies that only "fiendish" players would play this game variation, and the suggestions at the end, which imply that only good friends or near-strangers play this version of the game, are revealing. By removing the artificial nature of the game conflict, cheating can destroy the implicit camaraderie of the magic circle, letting its conflict leak out to infect the real-world relationships of players. Only friendships strong enough to weather such an experience or more disposable relationships in which further contact is not desired are appropriate.
The very notion that the rules could sanction cheating is a bit outrageous, but it ultimately fits the spirit of the game and its narrative world quite well. Illuminati is a parodic game about hidden organizations that rule the world, where the players are secret power brokers manipulating governments, media, and culture to their own devious ends. Seen in this light, the idea that the rules themselves are also subject to manipulation fits within the overall narrative trajectory of the game. Rule-breaking is a way of expressing the humorous critique of power that Illuminati the game embodies.

In the right context, sanctioned cheating can be an innovative way to enrich a game design. But it must be done with great care. Beneath the light-hearted tone of Illuminati's rules is a careful design allowing only those forms of cheating that leave the game intact, playable, and meaningful. Cheating in Illuminati does not remove all rules and boundaries from the game: it serves to re-draw them. Although the new boundaries might be drawn in lines that are considerably more fuzzy, a clear formal system remains. Even cheating is something that can be intentionally designed to facilitate meaningful play.

**Hacks, Cheats, and Mods: Digital Rule-Breaking**

When it comes to forms of rule-breaking incorporated into the design and experience of games, computer and video games offer a cornucopia of examples. Following are some sample instances of digital game rule-breaking, ranging from the timidly transgressive to the truly unlawful.

**Easter Eggs**

Easter eggs are secrets hidden in a game that players can discover. The first Easter egg was created by game designer and programmer Warren Robinett for the Atari 2600 game Adventure. In defiance of Atari's refusal to give credit to the creators of their games, Robinett programmed a secret room that could only be found with great difficulty. When players reached it, his initials were displayed. Hidden messages, images, and spaces are now a standard feature of digital gaming. In a mild kind of way, Easter eggs break a game's rules because they violate the otherwise internally consistent world of a game. Part of the pleasure of finding an Easter egg is a sense of transgressive discovery: by bending the rules of the game in just the right way, the player gets to see or experience something that more lawful players would not.

**Cheat Codes**

Although Easter eggs usually do not impact the strategic play of a game, cheat codes do. Like Easter eggs, developers design cheat codes into a game. Some of the best-known instances of cheat codes come from the first-person shooter DOOM, where a player can type special key combinations to gain weapons, health, and invulnerability. Sometimes a cheat code is a leftover tool from the game's development process, but often they are added just for the benefit of players. Although the name "cheat code" implies that these shortcuts to power are rule infringements, cheat codes frequently appear in game magazines and on official game websites, making them a form of officially sanctioned "cheating." The result is a rich culture of insider game knowledge, with fans scouring magazines and websites for the latest, coolest cheats.

**Game Guides and Walkthroughs**

Related to cheat codes are the sources of information that players turn to for help with a difficult or lengthy game. These resources appear on the web and in print, and range from elaborate color maps and strategy guides to fan-generated text files that cover every conceivable aspect of a game. Game walkthroughs are step-by-step instructions for finishing a game, particularly useful to players of adventure games and role-playing games that have a more linear structure. Some players view these resources as unfair techniques that breach the spirit of a game. At the same time, walkthroughs have raised the bar of difficulty and complexity in certain game genres. Many digital games are so challenging that they seem designed to require a guide.
Workarounds
The complexity of digital games often makes it impossible for designers to test or anticipate every possible permutation of play before releasing a title to the public. Furthermore, players are infinitely creative in finding ways of “legally” working around game structures. In “The Future of Game Design,” Harvey Smith writes about how players discovered new ways to play Deus Ex. For example, the proximity mine object is an explosive device that can be “stuck” onto walls in the game space. After the game’s release, players realized something that the game’s developers did not anticipate. Exploiting the game’s physics and interactivity, players learned to climb up on proximity mines, and using (or misusing) a series of these objects like a ladder, they could ascend the game’s vertical surfaces, ruining many of the carefully designed levels. Workarounds are on the borderline between dedicated play and unsportsmanlike play, and include degenerate strategies. Is it cheating to purchase game power by buying an EverQuest character on eBay, or is it simply a workaround that converts labor to capital?

True Cheating
In addition to fuzzier types of “cheating” behavior, there is plenty of bona fide cheating in digital games. More than clever workarounds or sanctioned cheat codes, true cheating breaks the rules of the game. In a multiplayer environment, guidelines for what constitutes cheating are generally made known to all players; cheaters are usually removed immediately and permanently from a game. In SISSYFIGHT 2000, the most common form of cheating is multi-sessioning, in which a single player opens up two game windows on two different computers, playing two characters at once and gaining very strong play advantages. Although it is difficult to spot, multi-sessioning is outlawed in the game, and there are vigilante fan websites devoted to maintaining lists of known game cheats.

Hacks
Hacking into a digital game goes beyond simply breaking the rules—it does so through intervention at the level of code. A player might hack a high score list, for example, to place her name at the top. Or she might modify the code of a first-person shooter to gain an unfair advantage in a deathmatch. If too many players hack a game, all sense of fairness can be destroyed. Therefore, the administrators of commercial multiplayer games put great effort into eliminating cheating and hacks from their games. According to massively multiplayer online game designer Ralph Koster, tracking down cheaters and hackers can occupy approximately half of all of the resources spent on maintaining and improving an online game.

Spoil-Sport Hacking
Most hacking is done in the spirit of the cheat: players want to do well in a game and do not mind breaking the rules in order to get ahead. Occasionally, game hackers can take the role of a spoil-sport as well, bringing down an entire game or game network. In this case, the aim is to dispel the magic circle for all players involved, not to better one’s own performance.

Why are digital games so fertile a ground for these varieties of rule-breaking? First and foremost, code is a plastic and pliable medium. The complex processes that give digital games their uniquely automated quality leave gaps for hacking into the system, whether it is through officially distributed cheat codes, clever workarounds, or genuine code-breaking. The anonymous nature of digital game play, where computers and networks mediate players, encourages rule-breaking as well. The reduced physical presence of other players permits a greater sense of social autonomy, which can facilitate the surreptitious activities of rule-breaking. Lastly, digital games are pop culture with a rich fan base: game fans deconstruct and reconstruct the codes and structures of the works that interest them. Cheating and hacking in this sense is similar to the ways that Star Trek fans re-mix the narrative universe of the television show to invent new stories and characters.
The blessing and curse of digital gaming media is that they provide a pliable space in which to play. With so many ways to gently bend and forcefully break the rules of a game, in playing a computer or video game players must decide what constitutes proper game behavior, navigating the space of possible rule violations. Is it acceptable to download a walkthrough guide? Do you use cheat codes to short-circuit your way through tough game levels? If you were offered a cracked version of the game that let you cheat, would you use it? As a digital game designer, you need to decide what kinds of rule-breaking you want to engender and what kinds you want to outlaw. Can you foster fan communities by offering sanctioned ways to violate the game without letting things get out of hand altogether? Ethics and game design collide in this rich space of rule-breaking possibility.

Rule-Breaking as a Game Design Practice
Our discussion of rule-breaking is not just an explication of the ways in which players break the rules of a game. It is a game design schema, a way of looking at all games that offers a framework for solving particular game design problems. However, it is a different kind of chapter than the other formal schema we encountered in our investigation of RULES. Framing games as systems of rule-breaking questions many of the unspoken assumptions of earlier schemas. We did not, in considering games as emergent systems, information, or cybernetic feedback loops, ever consider that players might disrespect or transgress the authority of the rules and the magic circle.

Player behavior is not universally law-abiding. Given any particular game, there are many ways to play it and many ways to bend and break its rules. For game designers, this means that you should never take players' behavior for granted. You need to assume that your game will be played not just by earnest rule-followers, but by zealously dedicated players, inappropriately unsportsmanlike players, brilliantly secretive cheaters, and uncaringly nihilistic spoil-sports. Some of these player types can help expand your game's space of possibility, whereas others can wreck the game for everyone involved. How do you take these possibilities into account in your game design? As always, there is no single solution. But framing your game as a system of rule-breaking lets you formulate your own answers.

There is yet another way to frame rule-breaking: as an attitude toward playing and designing games. We have seen a number of examples of how rule-breaking can enhance meaningful play. In professional sports, digital games, and in the cheating variant of Illuminati, breaking rules is part of the game itself. In all of these cases, through rule-breaking the space of possibility fills with alternative modes of play. What is the lesson here? Perhaps it suggests a shift in the way that we think about game design. In The Well-Played Game, Bernard DeKoven advocates a fundamental adjustment in players' attitudes towards the rules of a game:

You're not changing the game for the sake of changing it. You're changing it for the sake of finding a game that works.

Once this freedom is established, once we have established why we want to change a game and how we go about it, a remarkable thing happens to us: We become the authorities.

No matter what game we create, no matter how well we are able to play it, it is our game, and we can change it when we need to. We don't need permission or approval from anyone outside our community. We play our games as we see fit. Which means that now we have at our disposal the means whereby we can always fit the game to the way we want to play.

This is an incredible freedom, a freedom that does more than any game can, a freedom with which we nurture the play community. The search for the well-played game is what holds the community together. But the freedom to change the game is what gives the community its power.
Rather than obeying game rules as an ultimate authority, DeKoven would like players to assume authority over the rules. Once they feel confident and in control of the rules, players can break them and modify them in the course of playing a game. They do so not out of a mischievous desire to disrupt the authority of the rules, but out of a directed attempt to create a deeper experience of play. This beautiful vision for games does not describe the way that most people normally play. However, there is one type of game player that already has this attitude: game designers. Game designers, particularly those that design through an iterative process, already possess a methodology in which playing a game means breaking, tweaking, and modifying rules. In a sense, DeKoven is advocating that game players become more like game designers.

How are game designers rule-breakers? Being a game designer means that you are constantly testing the limits of a game you are creating. Which aspects of the rules are working and which are not? Do you need to add a feedback loop, or modify the amount of randomness in the game? Are players being faced with meaningful decisions at every moment? The best way to answer these game design questions is by changing the rules of your game, trying out new variations, and seeing what happens.

Of course, DeKoven’s vision for dethroning the authority of a game extends beyond just professional game designers. He would like to see all game players adopt this attitude toward play. What would it mean if all players felt free to break the rules of a game, to play not just inside the space of a game, but to modify and change the shape of that space itself? One answer to this important question is that it would require a fundamental alteration in the attitudes of game players and game designers. If players regularly break the rules, are they really rules at all? If players no longer stay inside the magic circle, are they really playing a game? Making this shift might be liberating, but it would certainly change the way we conceive games, game play, and game design.

Yet another answer to DeKoven’s challenge is that perhaps the phenomenon he describes already exists. Perhaps all players already play, not just inside the frame of a game, but with the frame of a game itself. If this is indeed the case, then all the varieties of rule-breaking players, from dedicated and unsportsmanlike players to cheaters and spoilsports, are natural extensions of the flexibility of game structures. Rule-breaking is simply one of the ways that we play.

Lastly, rule-breaking can be considered not just a way to play or design games, but a more general attitude about game design itself. If the conventions and genres of game design are the rules by which most designers “play,” then the innovators are those designers that manage to break the rules. Games hold great promise, but only if we are bold enough to truly break the rules of our field. This is harder than it seems. We know that to skillfully break rules requires an intimate knowledge of the rules themselves. And our hope is that this book provides some of those “rules of play”—rules that you will mercilessly and playfully violate in order to expand the space of game design’s possibilities.

With this chapter, we finish our first Primary Schema. In RULES, we consciously limited our gaze to the strictly formal boundaries of the magic circle, generally ignoring the player experience and the larger contexts in which a game takes place. But as we move forward, we will slowly widen the scope of our investigation, as we include those aspects of games that have been left out. How stable is the authority of a game’s rules? How permeable is the boundary of the magic circle? How is it possible to not just play a game but play with the very structures of gaming? We directly address these questions and many more in the PLAY and CULTURE schemas to come.
Further Reading

Grasshopper: Games, Life, by Bernard Suits (see page 98)

Recommended:
Chapter 4: Triflers, Cheats, and Spoilsports

The Well-Played Game: A Player’s Philosophy, by Bernard DeKoven
(see page 21)

Recommended:
Chapter 2: Guidelines
Chapter 3: The Play Community
Chapter 5: Changing the Game

Notes

• Breaking the rules is a phenomenon that occurs in almost every kind of game.

• Relative to rule-breaking, there are five player types. Each type of player has a particular relationship to the following aspects of a game:
  • adherence to the rules
  • interest in winning
  • degree of lusory attitude

• The standard player is the typical rule-following player that obeys the restrictions of the game and possesses the lusory attitude. Even if the standard player is a theoretical fiction, it is important to acknowledge this player position, which stands in contrast to the other four types.

• A dedicated player is similar to the standard player but has an extra zealousness toward succeeding at a game. The dedicated player follows the rules, is interested in winning, and possesses the lusory attitude.

• Unsportsmanlike players violate the implicit rules of a game without actually breaking operational rules. Their strong interest in winning gives them license to violate rules of etiquette and proper game behavior.

• Cheaters break operational rules of a game in order to win. Cheating players thus possess a strong interest in winning, but will forgo the normal means of achieving victory. Acknowledging that other players can invoke the authority of the rules, cheaters break rules secretly.

• A spoil-sport is a player that refuses to acknowledge the authority of a game in any way. These nihilistic players do not hesitate to destroy the magic circle of a game.

• The five player types are not always distinct. During a single game, a player can move from one category to another. The same behavior in different contexts can fall into different player categories.

• A degenerate strategy or exploit is a way of playing a game that ensures victory every time. Dedicated players and unsportsmanlike players make use of degenerate strategies. In general, degenerate strategies are detrimental to a game. However, within a community of players, degenerate strategies can sometimes act to expand the space of possibility.
- There are many examples of the integration of rule-breaking into game design and player experience, including professional sports, digital games, and games that sanction cheating such as Illuminati.

- Game designers need to recognize that rule-breaking is a common phenomenon in gaming and incorporate it into their game design thinking. One solution, which comes from the New Games Movement, is to empower players to be more like game designers by creating games with rules that are meant to be broken and modified.
It's not that you have to "care" in order to get good, but rather that you have to be kept caring. You've got to be kept in the right state so you'll get to some places a little bit better all the time, so that a goal remains alive by always moving just ahead out of reach and you keep wanting to attain it without having to spend a fortune.—

David Sudnow, Pilgrim in the Microworld
Introducing the Play of Pleasure

Video game arcades are sites of lucratively programmed caregiving, worlds of fun nourished by a seemingly endless stream of quarters, tokens, and plastic swipe cards. Players enter, they play, and if the game designers have done their job well, they stay to play some more. The carefully crafted arc of rewards and punishments that draws players into games and keep them playing connects pleasure to profitability.

Such intricate games of pleasure and play are not unique to the arcade. Pleasure is, perhaps, the experience most intrinsic to games. From the visceral excitement of an online deathmatch to the satisfying clink of a Go stone on wood, games provide an abundant variety of pleasures. We often take it for granted that games are fun to play, that they provide pleasure, that they embody enjoyable experience. Players derive many kinds of satisfaction from play, from the imaginative adventures of a narrative role-playing game to the social camaraderie of a team sports match. But what is the pleasure that underlies the appeal of games, the pleasure at the core of game play, the pleasure that provides the enticement to begin play and to continue playing? What connects pleasure to the design of meaningful play?

The word “pleasure” evokes associations with activities of leisure or self-indulgence. Sex, drugs, and rich foods come to mind, as do stolen naps, deep friendships, or dancing to a favorite song. Pleasure is commonly understood as a fundamental feeling that is hard to define but that people desire to experience. Words such as delight, amusement, gratification, satisfaction, or happiness describe the kinds of feelings pleasure evokes. When we speak of pleasure in games, we are referring to the fundamental feelings derived from the intense concentration of a game of Memory, the exhilaration of a winning touchdown, the charged socio-sexual maneuvers of Twister, the hypnotically satisfying patterns of Tetris. Pleasure can include any physical, emotional, psychological, or ideological sensation. Of course, pleasure’s opposites (pain, frustration, despair) are equally important in understanding the play of pleasure in a game.

Within *Games as the Play of Pleasure*, a game’s space of possibility is defined as more than a mathematical entity. It is a space in which a player’s emotions and sense of desire undergoes manipulation and coercion, teasing and seduction, frustration and reward. As the sculptor of the space of possible pleasure, the game designer faces a truly challenging set of problems. Managing the pleasure of a game’s players means translating the formal intricacies of the rules into an engaging experience of play. Although the emergent math of formal rulesets may be complex, the tangled puzzles of pleasure and desire are surely enigmatic dilemmas of an even higher order.

Rule-Bound

Picture a child poised excitedly at the starting line of a footrace, ready to run down the track, breathlessly awaiting the starting signal. Rather than giving in to her intense desire to leap from the starting line, she waits for the signal that the race has begun. What’s going on here? Why does our player anxiously hold back when she really desires to run?

Developmental psychologist L. S. Vygotsky notes that “Play continually creates demands on the child to act against immediate impulse, i.e., to act on the line of greatest resistance.”2 Certainly the child in our example wants to begin running, but the rules of the game order her to wait. At the same time, the runner knows that the rules are artificial, describing systems that are in some way outside ordinary life. So why follow the rules? Vygotsky argues that players accept the rules of the game not in order to restrict pleasure, but instead to maximize

Footnote to a Footnote

In *Defining Play*, we noted that we would not be investigating the purpose or function of play in this book. Rather, we focus on the way that play creates meaningful experiences for players, when considered from a game design perspective. Likewise, in the study of the play of pleasure, we will not suggest a root cause or mechanism, nor argue a unified theory of pleasure. There is a tremendous amount of existing research on the philosophical, psychoanalytic, cognitive, and cultural qualities of pleasure, some of which we reference in this chapter.
it. “To observe the rules of the play structure promises much
greater pleasure from the game than the gratification of an
immediate impulse.” Through mechanisms of restraint and the
withholding of immediate impulses, games transform the player’s experience of constraint into one of abundant pleasure.

The notion that pleasure is an effect of submitting to the rules
of a game, that pleasure delayed and constrained is pleasure
enhanced, offers a powerful model for understanding all kinds
of pleasure. Think of examples from your own experience: waiting
to eat a particularly enticing dessert until completing the
main course, or not skipping ahead to the end of a suspenseful
murder mystery. The delayed gratification of orgasm is height-
ened when it is initially resisted, as is the urge of opening a fine
wine before it has properly aged.

Submission to constraint is certainly not the only way to
understand pleasure, but it is an appropriate starting point for
a discussion of the play of pleasure in games. Consider, for
example, how the notion of constraint intersects with several
core game design concepts:

- **Rules and Play.** The idea that players subordinate their
  behaviors to the restrictions of rules in order to experi-
  ence play—and its pleasures—is a fundamental aspect of
  games. The restrictions of rules facilitate play, and in doing
  so, generate pleasure for players.

- **Free Play.** A player’s sense of pleasure is explicitly derived
  from being a part of the system of a game, from being “at
  play” within the more rigid structures of a game. In Man,
  Play, and Games, Caillois makes an explicit link between a
  player’s free action within the limits set by the rules and
  player gratification: “This latitude of the player, this mar-
  gin accorded to his action is essential to the game and
  partly explains the pleasure which it excites.” Free play is
  dependant on, yet also resists, the rigid structures that give
  rise to it.

- **The Lusory Attitude.** Playing a game means abiding by arti-
  ficial restrictions, which make game actions seemingly
  inefficient. Runners not only wait for the starting gun, but,
as Bernard Suits points out in Grasshopper, they also run
around a circular track, instead of cutting through the
middle of the field to reach the finish line first. Games are
constituted by these kinds of constraints, which simulta-
neously restrain and enable pleasure. The willingness of
players to step into these artificial systems in order to
experience the resulting pleasure is at the heart of the
lusory attitude.

- **Stylized Behavior.** Although play is a free and improvisa-
tional activity, the rules of a game stylize the actions and
behaviors of players in very particular ways. Think about
the patterned movement of players engaged in a game of
Ping Pong, or the tightly constrained movements of Simon
Says. There is something very pleasurable in the way that
games stylize play through a ritualistic, collective orches-
tration of movement and action. Children derive pleasure
not just from the dramatic tension at the start of a race, but
also from the collective experience of running together in
formation, pumping their arms and kicking their heels
about toward the finish line.

Rules give rise to the dramatic structure of pleasure, the link
between constraint and pleasure binding tightly the formal
and experiential qualities of a game. But players don’t simply
stumble into a game. Unlike other forms of ludic activities (such
as playing with a toy), a game demands that players know the
rules before play begins. What provides the enticement to
begin play? What makes players stay in a game once it starts?

**Autotellic Play**
The magic circle of a game is, by definition, removed in some
way from what Huizinga calls “ordinary life.” The victories and
losses, the triumphs and failures that a player experiences in a
game are in a very real sense contained within the magic circle.
As DeKoven puts it, a game provides “a common goal, the
achievement of which has no bearing on anything that is outside the game.⁵ We know, of course, that there are many ways winning or losing games can impact players: affecting their lifestyles, their sense of self, their relationships to friends, even the amount of money they have in their pockets when the game is over. There are certainly extrinsic ways that winning a game matters. At the same time, every game implicitly asserts the premise that the value of the game is intrinsic, that the game is self-contained, that the fiction of the magic circle will be upheld, that winning or losing the game is separate from everyday lived experience.

If one considers the self-contained nature of the magic circle, the way that games create their own meanings and provide their own goals, it is clear that games are strongly autotelic. We borrow the term from psychologist Mihaly Csikszentmihalyi, who in his book Flow explains that “The term ‘autotelic’ derives from two Greek words, *auto* meaning self and *telos* meaning goal. It refers to a self-contained activity, one that is done not with the expectation of some future benefit, but simply because the doing itself is the reward.”⁶ When an experience is autotelic, it is participation in the activity alone that counts. Games are, to a greater or lesser extent, pursued for their own sake, for their own intrinsic stimulation. Although there are always some extrinsic reasons for play, there are always intrinsic motivations as well. In playing a game, part of the incentive is simply to play—and often, it is the prime motivator.

Because they have such a strong autotelic component, games are largely non-utilitarian. Most forms of design serve an external function, or utility. Architecture houses and shelters our families, government, and industries. Typography enables visual communication. Automotive design supports mobility through the design of cars. Game design, on the other hand, simply enables its own play. Please note, in saying that game design exists in contrast to other forms of design, we are not proposing that games do not serve external functions, or that other forms of design don’t also serve non-utilitarian ends. Our point is that games posit their own intrinsic needs or goals, such as abstract winning conditions, which gives them a distinctly artificial and non-utilitarian status.

Contrast an online medical database program with an online multiplayer game. A hospital worker looking for a particular patient’s record comes to the software experience with a clear extrinsic goal in mind, such as finding out what meds the patient needs to take that day. The database does not contain its own set of goals; it supports the goals of the user. The database program is used as a tool, as a means to an end, rather than as an end in itself. When the worker finds the record he is looking for and extracts the prescription information, the database has successfully fulfilled a goal that was brought to the system from an external context. In an online multiplayer game, on the other hand, there is no clear utilitarian purpose that the game serves. Why is the player exploring the game world, customizing her character, killing monsters, and accumulating treasure? Because she is playing the game. The game is not a tool being used to fill an external, utilitarian need. The player is not playing the game in order to feed her cats, or tune her car’s engine. The explicit interaction of the game is not a means to an end, as in the case of the medical database program; rather, the play of the game represents an end in itself. We play, in some measure, for play’s own sake.

Consider the way that the experience of play as an end, rather than a means, has affected the development of digital game technology. One of the reasons why games have been so innovative—pushing the envelope of computer processing power, creating experimental hardware interfaces, pioneering graphics rendering and spatial audio—is because games must provide their own motivations and pleasures. The medical worker will suffer through an awkward interface and ugly visual design in order to find the record he needs. A game player, on the other hand, is a much more fickle user: why play a game that isn’t fun? The computer and video game industry is continually spurred on by an audience hungry for ever-more spectacular games
and ever-more meaningful interaction. People play games because they want to; game designers must create experiences that both feed and satisfy this sense of desire.

Enter. Play. Stay.

Why go to such lengths about the non-utilitarian nature of games? In order to make a larger point about the challenge of bringing players into a game and keeping them at play. Because games are premised on needs intrinsic to the game, it is necessary for game designers to both entice the player into crossing the boundary of the magic circle and also keep them there until the goals of the game have been met.

Beginning a game means entering into the magic circle. Players cross over this boundary to adopt the artificial behaviors and rituals of a game. During the game, the magic circle persists until the game concludes. Then the magic circle dissolves and players return to the ordinary world. These two actions, crossing into the magic circle as well as maintaining its existence, represent two of the chief challenges of designing meaningful play. The two actions require a carefully orchestrated double seduction. First, players are seduced into entering the magic circle of a game. Second, players are seduced into continuing to play.

Both events are challenging to design. The first seduction, bringing players into the magic circle, requires players to cross a threshold that will take them out of their ordinary lives and into the world of the game. The difficulty in making this happen comes from the formal quality of game play. It is much easier to slip into and out of ludic activities that aren't games. Are you eating peanuts and feeling playful? Just toss one up and see if you can catch it in your mouth. How about those building blocks on your desk? Stack them up, knock them down, or just let them be. In The Magic Circle, we looked at the way a child might play with a doll, at how smoothly a player can slip in and out of play, at how permeable the borders are between playing and not playing. In games, however, the transition between not playing the game and starting to play the game is more clearly defined. Games usually require formal preparation: finding players, reading the rules, opening a saved game file, shuffling cards, setting up the board, and so on. Players must learn the system and "officially" enter into the game and begin play. This is a genuine hurdle for players of your game: they must attend to the initial set of chores that lie on the border of the magic circle; they must properly perform the rituals of entry.

What does this mean for game designers? Designers create not just the game itself, but also the ways that players enter into the game system. This event involves consideration of not just the formal elements of the game, but also the way that the game interfaces with external contexts. How and when does a player enter into a game? Where does the initial seduction begin? Does it begin the first time a player sees a commercial or reads a review of the game that encourages or discourages the player to make a purchase? Does the seduction emerge from peer pressure and social values (Barbie Fashion Designer is for girls! Quake is cool! Everybody is playing P.O.X.!). Does it begin with the installation of a downloaded game, the first reading of a game's rules, or the menu screen of a console title? Does it start the moment a newbie shoots his first monster?

Clearly, there is no single factor to which the act of seduction can be attributed and no single, isolated moment when the player decides to begin play. Designing the seduction of a game means understanding all of the formal, social, and cultural factors that contribute to the player's experience. It is important, for example, to understand how marketing, promotion, and distribution work in the game industry. It is important to scout out what other game developers are creating and how it may impact the game you are designing. It is important to understand how the culture at large perceives and regards games and how new audiences might be brought to your games. There are no simple answers to the question of whether or not a player will decide to begin playing your game. This is one more challenge game designers face.
On the other hand, once players start playing your game, they
have stepped out of the world at large and entered into the
magic circle of the game’s design. As we will see, keeping play-
ers in a game, understanding and sculpting their experience of
pleasure, offers at least as great a challenge as getting them to
play in the first place.

**Typologies of Pleasure**

It is difficult to generally speak about pleasure. It is especially
hard to find the words that describe the pleasures we experi-
ence in games.

Games evoke emotions of struggle, of competition. The kinds of
things you feel aren’t often given common names in our usual
everyday parlance but they are important emotions that we feel
and go through and enjoy and find in some mysterious ways
enlarge our spirit. How about the anxiety that you feel when your
chest suddenly swells as you realize you are going to be a master?
How about the sense of self that develops as you concentrate all
your being and the various parts of your body upon the task of
overcoming obstacles? How about the dejection you feel, the
despair when you fail utterly? And how about the exultation and
the sense of triumph you feel when you actually succeed? And
sometimes a little bit of awe as you maybe find that path out
there. And there’s another name for these emotions and game
developers call them fun.

Game designer Hal Barwood organizes all the varied emotions
a game can produce under the heading of “fun.” This term does
make some sense. Good games are fun. Fun games are what
players want. A fun game makes for a pleasurable experience,
which is why people play. But not everyone sees the value of
this word. Game designer Marc LeBlanc simply hates the term
“fun.” In several of his talks at the Game Developers Confere-
cene, he has called for a moratorium on the word. “Fun,” accord-
ing to LeBlanc, is merely a stand-in term for a more complex phenom-

Perhaps LeBlanc is right. Perhaps we are falling into a similar
trap by our use of the word “pleasure.” Is it possible to unpack
the more general notion of fun and create a structure for cate-
gorizing pleasure? LeBlanc has done some thinking along these
lines and has created his own typology, proposed as an anti-
dote to the singular concept of “fun.” In his typology LeBlanc
lists eight categories that describe the kinds of experiential
pleasure players derive from playing games.

1. Sensation: *Game as sense-pleasure*
2. Fantasy: *Game as make-believe*
3. Narrative: *Game as drama*
4. Challenge: *Game as obstacle course*
5. Fellowship: *Game as social framework*
6. Discovery: *Game as uncharted territory*
7. Expression: *Game as self-discovery*
8. Submission: *Game as masochism*

Most of these categories are self-explanatory. Note, however,
that “masochism” doesn’t refer to sexual pleasure, but instead
to the more general pleasure of submission to a system. Part of
the hypnotic pleasure of Bejeweled or Solitaire, for example,
comes from the ritualized act of behaving in a rule-based, styl-
ized manner. That is what LeBlanc means by his category of
Submission.

LeBlanc’s model is intended not only to assist game designers
in understanding the range of forms that “fun” can take, but also
to provide a common language for marketing digital games. He
has proposed, for example, that by rating each of these catego-
ries on a zero-to-ten scale and putting that information on
the back of a product package, a consumer could quickly get a
sense of the kinds of pleasures the game provides. A first-per-
son shooter, for example, might have a high rating in Sensation,
Fantasy, and Challenge, but a low rating in Expression, Narrative,
and Fellowship. The challenge, of course, is that many of the
categories seem to overlap. There is a very fuzzy line, for example, between Fantasy and Narrative. Other ambiguities persist as well. Categories such as Discovery and Expression might easily be applied to other categories: can’t a social framework be uncharted territory? Doesn’t self-discovery occur in a challenge? Moreover, even if these theoretical problems could be resolved, “officially” rating a game’s pleasure in this way would be a highly subjective endeavor. Despite all of these criticisms, however, LeBlanc’s eight categories do identify many of the components of game-induced pleasure and are useful as a way of understanding the range of pleasures games provide.

A different approach comes from psychologist Michael J. Apter, in his essay “A Structural-Phenomenology of Play.” In focusing on the cognitive arousal play provides, Apter compiles the following list, amended with our brief paraphrasing in italics:

1. Exposure to Arousing Stimulation: intense and overwhelming sensation  
2. Fiction and Narrative: emotional arousal from character identification  
3. Challenge: difficulties and frustrations arising from competition  
4. Exploration: moving off the beaten track into new territory  
5. Negativism: deliberate and provocative rule-breaking  
6. Cognitive Synergy: imaginative play  
7. Facing Danger: risk within the “protective frame” of play?

Apter admits in his essay that these categories offer only a partial list of cognitive arousals, and that there is considerable overlap between categories. Despite these delimitations, Apter’s model gives us another framework within which to consider pleasure in games, one that emphasizes cognition. Some of his categories, such as Challenge and Exploration, appear similar to LeBlanc’s. Others, such as Negativism and Facing Danger, clearly identify alternate approaches.

A third typology of pleasure comes from the classification of games by anthropologist Roger Caillois. In *Defining Play*, we introduced his four “fundamental categories,” which purport to describe the phenomena of play:

1. *Agôn*: competition and competitive struggle  
2. *Alea*: submission to the fortunes of chance  
3. *Mimicry*: role-playing and make-believe play  
4. *Illinx*: vertigo and physical sensation

In some ways, Caillois’ compact categories offer a succinct distillation of the models LeBlanc and Apter propose. In āgon, alea, mimicry, and illinx, there is a fusion of experiential and cognitive components that creates a useful critical framework.

There are many other typologies we could consider as well. Last chapter, we looked at Brian Sutton-Smith’s five categories describing the psychological processes of video game players: concentration, visual scanning, auditory discriminations, motor responses, and perceptual patterns of learning. These too might be considered a list of the means by which games generate and support pleasure. There is no need to choose a single typology to represent pleasure in games. You should feel free to mix and match different models of experience and pleasure, depending on the needs of your design. These typologies are less useful for theorizing about pleasure or for classifying games, but they can be very handy as a way of organizing observations about the kinds of pleasures that a particular game provides. One model is not necessarily better than the others; each offers a different way of thinking about pleasure and its many motivations.

For example, let us employ one of these typologies—Caillois’ four categories—in looking at an Unreal deathmatch. Do they apply to the pleasures of playing Unreal? Certainly the game contains a great deal of competitive, agonistic struggle. Mimicry plays a strong role as well, in the fact that each player is represented to the others through a customizable avatar in a fictional, virtual space. Unreal and games of its ilk are well
known for representing physical movement through three-dimensional space in real-time, often creating vertigo in the form of motion sickness. There are arguably even elements of chance in Unreal as well, such as the particular players that happen to join an online deathmatch, or the layout and distribution of items on a level.

We can similarly apply the categories of LeBlanc and Apter. A game of Unreal provides all of the pleasures they list too, from the Fellowship that emerges out of hard-fought competition, to the creative Negativism of cheats, hacks, and mods. Pleasure is always already exceedingly complex: where we find one form of pleasure in a game, we will almost always find others. In general, most games provide many or all of the pleasures listed in any typology of game play experience. But at the same time, there is always a balance of factors, a particular ratio of ingredients that adds up to the unique flavor of an individual game experience. What meaningful pleasures is your game providing, or failing to provide? This is the utility of a typology of "fun:" offering a vocabulary for charting out the complex play of pleasure.

Game Flow
Listing categories is one approach to describing pleasure in a game. Are there other approaches? Is it possible to look at game pleasure in a more abstract way to synthesize the diverse pleasures of gaming into a single concept? Think again about the experience of playing a game. One aspect of game pleasure lies in the intensity with which it is experienced, the almost overwhelming sensation of play. Whether the pleasure rests in a cognitive response, an emotional effect, or a physical reaction, the experience of play, and especially play in games, can be strikingly deep. As writer J.C. Herz writes of classic arcade gaming, "Just the emotion, the survival nature of the videogame—you're tapping into the most powerful human instinct. Survival. Fight or flight. That is so hugely intense that in some ways it becomes too intense. People really lived the games. They dreamed the games."

All game players have experienced this feeling at one time or another, even if for only a short time. This level of engagement with a game suggests that the player has transcended an ordinary psychological state to arrive at a more profound relationship with the game. The psychologist and theorist Mihaly Csikszentmihalyi is best known for his research on what he calls the flow state—a particular state of mind in which a participant achieves a high degree of focus and enjoyment. His book Flow: The Psychology of Optimal Experience offers a great general introduction to his ideas. Flow is filled with anecdotal accounts of individuals achieving a flow state, documented over the years of Csikszentmihalyi's research.

The phenomenon of flow comes in many forms. Some people reported to Csikszentmihalyi that they reached flow through the rigors of perfecting an assembly line work task, or through the immersive problem-solving of law library research. Others say they achieved flow during the solitary exertion of rock climbing or through the exacting vocation of surgery. What exactly is flow? Csikszentmihalyi suggests that flow is something we have all experienced. It is a feeling of being in control of our actions, masters of our own fate. Although rare, when we achieve a state of flow we are deeply exhilarated. Csikszentmihalyi refers to this phenomenon as an optimal experience.

It is what the sailor holding a tight course feels when the wind whips through her hair, when the boat lunges through the waves like a colt—sails, hull, wind, and sea humming a harmony that vibrates in the sailor's veins. It is what a painter feels when the colors on the canvas begin to set up a magnetic tension with each other, and a new thing, a living form, takes shape in front of the astonished creator.11

Flow is, more than anything else, an emotional and psychological state of focused and engaged happiness, when a person feels a sense of achievement and accomplishment, and a greater sense of self. What might be the relevance of flow for game design? In many ways, the heightened enjoyment and
engagement of the flow state is exactly what game designers seek to establish for their players. In fact, many of Csikszentmihalyi’s examples come from games, such as professional Chess players, which were an early focus of his research. The connection between game design and the flow experience clearly appears in Csikszentmihalyi’s description of the components of flow, the conditions that make flow possible. He lists eight components:

First, the experience usually occurs when we confront tasks we have a chance of completing. Second, we must be able to concentrate on what we are doing. Third and fourth, the concentration is usually possible because the task undertaken has clear goals and provides immediate feedback. Fifth, one acts with a deep but effortless involvement that removes from awareness the worries and frustrations of everyday life. Sixth, enjoyable experiences allow people to exercise a sense of control over their actions. Seventh, concern for the self disappears, yet paradoxically the sense of self emerges stronger after the flow experience is over. Finally, the sense of the duration of time is altered: hours pass by like minutes, and minutes can stretch out to seem like hours.12

It should be immediately striking how every one of these eight components corresponds to an aspect of games. We can look at each in more detail, making use of Csikszentmihalyi’s own language.13

A Challenging Activity that Requires Skills: Csikszentmihalyi emphasizes the fact that the flow activity is not passively experienced; it requires active and directed engagement. “The overwhelming proportion of optimal experiences are reported to occur within sequences of activities that are goal-directed and bounded by rules.” This sounds remarkably like a description of a game.

The Merging of Action and Awareness: One distinctive feature of the flow state is that a person is so absorbed in the activity that it becomes “spontaneous, almost automatic; they stop being aware of themselves as separate from the actions they are performing.” This component of the flow experience is something that can occur in games as well. David Sudnow’s account of his engagement with Breakout clearly describes this state of mind.

Clear Goals and Feedback: These two components evoke the goal-oriented nature of games and the discernable action-outcome sequence necessary for making meaningful choices. Meaningful play seems to be intimately related to flow.

Concentration on the Task at Hand: A common effect of flow is “a complete focusing of attention on the task at hand, thus leaving no room in the mind for irrelevant information.” Like a game that removes itself from “ordinary life,” flow activities carve out their own experiential spaces for participants.

The Paradox of Control: In an optimal experience, the participant is able to exercise control without completely being in control of the situation. If there is no chance of failure, the activity is not difficult enough. “Only when a doubtful outcome is at stake, and one is able to influence that outcome, can a person really know whether she is in control.” As game players struggle against the system of artificial conflict, they attempt to assert control by taking actions. Yet the outcome of a game is always uncertain.

The Loss of Self-Consciousness: In flow, the participant’s sense of self becomes subservient to the greater whole of the experience. “When a person invests all her psychic energy into an interaction... she in effect becomes part of a system of action greater than what the individual self had been before. This system takes its form from the rules of the activity; its energy comes from the person’s attention.” The fact that Csikszentmihalyi emphasizes the systemic quality of a participant’s connection with the flow activity is reminiscent of the system-based nature of games.
When we consider a game as an experiential system, the player is a component of that system—a formulation echoed by Csikszentmihalyi.

The Transformation of Time: The participant’s sense of time can stretch or shrink. Sometimes this feeling comes directly from the activity itself: “Most flow activities do not depend on clock time; like baseball, they have their own pace, their own sequences of events marking transitions from one state to another without regard to equal intervals of duration.” Games not only change our perception of time but also offer freedom from its tyranny; losing track of time adds to the exhilaration we feel during a state of complete involvement.

In each of the eight components of the flow activity Csikszentmihalyi mentions, there are clear parallels with games. This doesn’t mean that flow applies only to games, or that every game produces a flow state for its players. What it does mean is that games are one of the best kinds of activities to produce flow. The rules, goals, feedback, uncertain outcome, and other qualities of games make them fertile terrain for the flowering of a flow experience. We believe there is an intrinsic connection between game play and flow. Although the maximum flow “optimal experience” that Csikszentmihalyi describes is rarely achieved, all forms of play in some way partake of the flow experience. The conditions for flow are established as players find the interstices of a rigid structure, engaging with rules in order to play with them and transform them. Flow is one way of understanding that pleasure which draws players to a game and keeps them there.

Although he does not organize them this way, Csikszentmihalyi’s eight categories can be divided into two groups. Four of the eight components of flow describe the effects of the flow state:
- the merging of action and awareness
- concentration
- the loss of self-consciousness
- the transformation of time

All of these effects occur in the player’s experience once flow commences. These four facets of flow can diagnose whether a player has reached the flow state. If you are not sure if your game is truly producing flow, go down the list. If some or all of the four experiences listed are missing, you may need to adjust your design. But what kinds of adjustments are necessary? That’s where the other four components come into play. Rather than being effects of flow, they represent flow’s prerequisites:
- a challenging activity
- clear goals
- clear feedback
- the paradox of having control in an uncertain situation

These four prerequisite elements of flow are characteristics of the flow activity itself. Within them is the key to designing flow in games. Does your game contain the prerequisites of flow? Is there enough challenge to create real uncertainty? Do the players clearly understand the goals? As they move through the system, do their actions provide clear feedback and a sense of control? If your game supplies all of these mechanisms, you are well on your way to creating the necessary conditions for flow.

If your aim is to create a flow state for your players, we can summarize our advice quite simply: design meaningful play. The four prerequisites of flow bear a striking resemblance to the key components of meaningful play. “Clear feedback” is another way of stating the need for discernable choices and outcomes in an interactive experience. The goals, challenge, and uncertainty of a game provide the larger context within which choices are integrated and become meaningful. This is not to say that meaningful play is the same thing as flow. Flow is a state of mind and meaningful play is an approach to game design. But when it comes to games, the two are closely intertwined. If you
want to create flow in a game, meaningful play must be present. If you want to design meaningful play, flow can be a useful diagnostic tool in the process of making your game.

Why would game designers want to create a flow state for their players? Being in flow represents a rich and meaningful engagement with the activity at hand. Generally, as a game designer, you are creating game systems meant for deep exploration. We should all be so lucky that players of our games invest enough effort and attention to achieve a state of flow. Remember that flow doesn’t refer to just one kind of experience. The flow that surgeons feel is by all accounts radically different in sensation and emotion than the flow of a LARP combat. What unites all forms of flow, however, is the optimal happiness that participants experience. As an experiential goal for creating games, spreading happiness, focus, and a sense of well-being is certainly a worthy pursuit.

Sculpting Desire

If one problem with the concept of flow is that it is not as gamespecific as we would like, what would it mean to take a more game-centric look at pleasure? Thinking of games as systems of pleasure implies that the game designer is an artisan of desire, shaping the pleasure of the players of a game. The designed system of the game, set in motion by the participation of players, becomes not just an experience of play, but also an experience of sensual, emotional, and psychological pleasure.

Achieving such an experience requires that a game designer not only pay attention to the immediate feelings of pleasure a game may produce, but also the way that a player’s pleasure evolves and changes over the course of a single game, or across many games.

Anyone can sit down at Quake and start shooting things. As he gains more experience, he realizes that if he stands in one place, he’ll get killed, so he learns to start moving while shooting. Then he learns to circle-strafe. Then, to shoot while running backwards. Then, to figure out which weapons are better up close or far away.

Disclaimer: The Limits of Flow

As useful as it is, the concept of flow is not a skeleton key to unlock every mystery surrounding play and pleasure in games. Consider a few of the challenges in applying flow to game design:

Flow is not unique to games. As Csikszentmihalyi’s many examples from art, work, and non-game leisure demonstrate, flow can occur in many kinds of activities. Why is this a problem? If one of our goals as game designers is to understand and isolate the unique kinds of pleasures that only games can provide, then the flow state is not of much help.

Flow is more about the player than the game. According to Csikszentmihalyi, flow depends at least as much on the attitude of the individual participant as the activity itself. Chess masters may achieve flow, but most Chess players do not. There is no guarantee that the game you design will be put to use by players that are ready or able to experience flow. A player’s individual psyche is out of a game designer’s control.

Flow is not a universal phenomenon. It is easy to get carried away and assume that the flow state is the ultimate experience that every game design should try and induce. That is simply not the case. As Sutton-Smith points out in a critique of Csikszentmihalyi’s work, “To say flow is universal might be like saying that all peak sex is everywhere the same, and that ‘flow’ is to play what orgasm is to sex. But who would be innocent enough of all the different contexts and acts that make sex meaningful to say something like that?”

Although flow is a useful conceptual tool for creating pleasure in games, it is but one of many possible tools. Flow offers a rigorous investigation into one kind of meaningful engagement, even if it doesn’t represent a universal state of mind and even if it isn’t completely unique to games.
Then he learns to rocket jump. As he progresses, he learns the characteristics of each weapon. He learns to "lead" his opponent. Anyone can pick up Quake and start having a good time within minutes, yet the longer he spends mastering the game, the more enjoyable it becomes.15

Quake is easy to learn but difficult to master. As game designer Bob Bates points out, it offers players a gaming experience that is pleasurable in both the short and long term. Because the core mechanic is relatively simple—move and shoot—players gain immediate access to the pleasures the game affords. Because playing the game well requires subtle skills that can only develop through repetitive play, long-term engagement with the game brings its own kind of pleasurable reward.

How does pleasure emerge and evolve over time in a game? All of the possible states and experiences of a game are contained within the theoretical construct called the space of possibility. A game player begins his or her journey through the space of possibility at the same place every time: the start of the game. But the experiential path that a player takes through the space will vary each time the game is played. Every play of the game will be unique, even though the rules of the game, its formal structure, remain fixed. This quality of games, that a game provides the same consistent structure each time but a different experience and outcome every time it is played, is a powerful engine that sustains and encourages play. We refer to this concept by the shorthand term same-but-different.

The same-but-different experience of play occurs both within an individual game, as well as across more than one game. Inside a game of Breakout, the player engages with the core mechanic over and over, exploring its permutations many times within the changing context of the game. Hit the ball again and again. Can you get that bank shot a second time from the side? Or slow things down by hitting the square in the middle of the paddle? Can you control the flow of the volley? Within a game, given the same repetitive action of play, part of the pleasure that sustains the game is the player’s ability to engage repeatedly with the same kind of interactivity—but with different results.

The core mechanic of a game provides its own inherent pleasure, whether it takes the form of the sensual click of a Tiddley-Winks flip or the simple, randomized drama of each round of the card game War. If a game’s core mechanic is well-designed, players may not even care about winning. Players can enjoy Tennis just for the sake of a good volley, or Charades for the challenge of skillful pantomime and clever guessing. But to sustain pleasure over time, the repeated action of the core mechanic needs to embody the concept of same-but-different. It needs to continue to offer up new variations and experiences, even if they are as subtle as the gradual build-up of Tiddley-Winks skill or the deterministic playing-out of fate in War.

On a larger level, the same phenomenon occurs as a player plays a game more than one time. In this case, it is not the core mechanic that is repeated, but the entire formal structure of the game. The rules remain the same, but the play is different. It doesn’t always happen, but if the play is meaningful enough, if the pleasure is rich and flowing, then a player will want to play a game again. With repeated play, the structure is increasingly familiar, and the player continues to play out the possible experiential permutations of the game. Within and between games, players discover the comforting familiarity of a fixed structure and the challenge and danger of an uncertain outcome. This same-but-different mechanism makes for an extremely powerful engine of desire. It is the itch of the same-but-different that brings you back, time and time again, for just one more round of play.

Furthermore, transformative play assists this process. In Defining Play, we established transformative play as the special case of play, when the free movement of play alters the more rigid structure in which it takes shape. When the structure of
the game is altered, the possibilities for replayability increase. Even in a simple betting game, transformative play can ensure that new experiences continue to arise, as a player finds new patterns of betting, new places in which to bet, new patterns of life into which the betting activity fits, new circles of friends to support the activity. Still, the game remains familiar, even as it changes. Philosopher James S. Hans expresses this notion of the same-but-different experience of play quite well:

In this regard, all play shares one thing with games: a familiar structure that allows one to play with the unfamiliar. This familiar structure is not universal; it is contingent upon the particular context of play. Nor is this familiar structure always the same. Indeed, it changes every time it is played with, for the occasion for new play introduces different elements into the activity that become part of the structure of any future play. . . . The structure of the familiar then permits the introduction of the different; play in one sense is no more than the infection of the familiar by difference.16

Although Hans is talking about all kinds of play, every game by our definition shares this quality. He makes explicit the idea that play is transformative, that through repetition, play itself changes. Hans calls this play of desire "the infection of the familiar by difference"—perhaps the heart of what makes games pleasurable. Within the magic circle, rules endow actions with meanings. But the free movement that is play transforms these meanings, even as they are experienced, putting pleasure "at play" at each moment of a game.

Patterns of Pleasure
The patterns of pleasure that emerge within and between games offer special kinds of enjoyments. Game designer Brian Moriarty uses the word entrainment to refer to this kind of rhythmic pleasure. Entrainment comes from the French word "entrainer" and has two meanings: to carry along, and to trap. The word has commonly been applied to a range of physical and natural phenomena, from circadian sleep rhythms to the sonic play of a thunderstorm. According to holistic thinker Dr. Stephan Rechtschaffen, "Rhythmic entrainment is one of the great organizing principles of the world, as inescapable as gravity. It explains how one rhythm works with another, and how separate entities, from molecules to stars, will fall into rhythm as automatically as a pulse beats or a butterfly flaps its wings."17

As used by Moriarty, entrainment is the process of falling into a patterned activity, such as when baseball fans spontaneously create a stadium-wide "wave" in a co-authored, massively multiplayer spectacle. One can also apply the concept to the play rhythms evoked when playing a game. In 1998, Moriarty gave a talk at the Game Developers conference about entrainment and game design:

Rhythms and pattern exist in all games, if you watch. Watch someone playing a game sometime. Not the game itself, lest you be sucked in, but the player and the space around him or her. Watch the rhythms emerge, and how the player and the game interact. It will become clear that a game is really an entrainment engine. The job of the gamewright, therefore, is to reinforce patterns, and dampen dissonance.18

The notion of entrainment combines pattern, interactivity, and the same-but-different quality of games into a rich and powerful design concept. If entrainment is a form of pleasure, it is a pleasure at once structural and experiential, both mathematically regular and playfully flexible. Entrainment is not a phenomenon completely unique to games, but it does come very close to identifying the curious structural pleasure that all game experiences seem to contain: the meditative patterns of Tetris; the turn-taking, clacking cadence of Billiards; the rhythmic shooting pattern of Space Invaders; the pulsing flow of cards, hits, and chips of Blackjack. Each of these game experiences—every game experience—can be framed as an instance of entrainment.
Entrainment is the experience of the same-but-different. As players explore the space of possible game pleasures, progress through the space occurs through patterned repetition, the drumbeat driving the heart of a game experience. Entrainment sometimes literally takes on form: the recurrent bleep of a laser blast, or the relentless throb of a marathon runner’s steps. But ultimately, entrainment manifests in a more pervasive fashion, occupying not just perceptual sensations, but modes of thinking and feeling as well. The double-sided definition of the word entrainment, to carry along and to trap, is entirely accurate. The patterns of a game initially draw us in, moving us forward, encouraging us to play. But somehow, at some point, something changes. We find ourselves not just playing a game, but being played by the game as well. Pleasure is a mighty force, and it can carry along those trapped in its wake.

The Role of the Goal

How does this transition come about? How is it that a game can draw us in and take us hostage? Some of the most powerful mechanisms of pleasure that a game contains are derived from their constituent parts. The difference between games and other forms of play is most often the fact that a game has a goal and a quantifiable outcome. When it comes to understanding the pleasure of a game, the goal plays an absolutely crucial role.

A game’s goal is often the largest single element that drives the pleasure of a player. The goal is the ostensible reason for playing, but the goal is never easily attained; rather, it is the obscure object of desire, the carrot held just out of reach, pulling players forward through the varied pleasures of game play. The goal helps move players through the space of possibility, a space stretched between the starting state of the game and its outcome like a billowing cloth staked to the ground. The goal acts to guide the players along the axis defined by the beginning and the end, letting them know if they are advancing or falling behind. In Chutes and Ladders, the player’s position on the board clearly communicates proximity to achieving victory. In Chinese Checkers, the accumulation of pieces within a player’s goal serves as a competitive marker toward winning. In an online deathmatch, ranking players by frag count shows which players are closest to victory.

A core component of pleasure in games lies in the creation and maintenance of a player’s relationship to the goal. The game designer, by creating the game rules, indirectly engineers this relationship. Do you want to create a close game with a dramatic finish? Then engineer a negative feedback loop that punishes the lead player and closes the gap with players at the rear. Or perhaps you want to make use of hidden information, so that the outcome is in doubt until the very end. In the puzzle game Mastermind, each turn brings the player closer to the end of the track, where failure awaits after a set number of turns. Each turn also offers the player feedback and a glimpse of hope, signified by the black and white pegs, as he tries to puzzle out the correct solution. Between the inevitable, uncontrollable turn-by-turn movement, and the sometimes-backpedaling advancement toward the solution, the game unfolds as a journey of the player’s desire. The result is the genuinely compelling play of Mastermind, in which the physical structure charts the player’s progress through the game.

The classic arcade game Missile Command offers yet another example of the way games engineer dramatic experiences for players. As J. C. Herz writes in Joystick Nation: How Videogames Won Our Hearts and Ate Our Quarters:

The most intense thing about Missile Command, though, was this weird crazy moment near the end, when the ICBMs were raining down and you knew you were just about to lose it, that was totally euphoric. Because you knew that you were going to die, that you were within seconds of everything going black. You’re gonna die in three seconds. You’re gonna die at this instant. You’re dying. You’re dead. And then you get to watch all the pretty explosions. And after
the fireworks display, you get to press the restart button, and you’re alive again, until the next collision with your own mortality.\textsuperscript{19}

Within the game the goal takes on enormous importance, but the goal itself as a formal construct is not the point: the goal is important only insofar as it serves to shape a player’s experience. The goal is an artificial, invented condition that the players accept as their ultimate objective. In establishing the nature of that goal and the way that players overcome adversity to work toward it, the game designer has tremendous influence on crafting the character of the play experience.

**Goals Within Goals**

The goal is not the only source of pleasure in a game. In addition to the thrill that the pursuit of victory (or the agony of defeat) can provide, games offer many pleasures that are parallel, or even tangential, to winning. Just as important as the final win condition, the macro-level goal, are the tiny moments of directed play, the micro-interactions that move a player through a game. These smaller moments of play emerge as the player engages repeatedly with the core mechanic, the same-but-different experience sustaining the interest and desire of the player.

If the macro-level of a game’s pleasure is the player’s pursuit of the goal, and the micro-level is the player’s engagement with the core mechanic, then what is it that links these two levels of play? The answer is *short-term goals*. A game never simply provides a single long-term goal. Along the way, a player struggles toward short-term goals, each one providing a kind of pleasure that is less immediate than the instant gratification of the core mechanic, but more rapidly obtained than the long-delayed ultimate outcome of the game. Even in a simple game like Tic-Tac-Toe, there are short-term goals that help players gauge their progress through the system. Placing an “O” in the same row as another “O” to form two-in-a-row is a short-term goal that must be reached before one can achieve three-in-a-row (and victory). This short-term goal may sound uninterestingly simple, and for adult players it usually is. But for young children struggling to comprehend the strategic complexities of the game, understanding short-term goals and the way these goals link the core mechanic of mark-making with the long-term goal of three-in-a-row is crucial to their enjoyment of the game.

The kinds of short-term goals that a player can achieve depend on the nature of the game and the way the goals are suspended between the core mechanic and winning. In a wargame such as Tanktics, the short-term goal might be outflanking the enemy’s ranks in order to weaken their defensive position on the battlefield. In SISSYFiGHT 2000, it might be making a social alliance with another girl to shift the wrath of the player mob onto a particular player. In the digital trading game Dope Wars, a short-term goal might be saving up enough cash to move up from selling pot to selling heroin.

A game can explicitly provide short-term goals, such as the medals a Pokémon player periodically earns by beating the best trainers in particular gyms. However, it is also very common for players to generate short-term goals themselves, in response to their current situation. A Pokémon player might be concerned with earning every medal in the game, but perhaps he invents a different short-term goal, such as capturing every species of Pokémon, or moving his Pidgeotto up to level 50.

Encouraging players to conceive and achieve goals gives them a sense of control in the game, as Doug Church points out in his essay “Formal Abstract Design Tools:”

There are many ways in which players are encouraged to form their own goals and act on them. The key is that players know what to expect from the world and thus are made to feel in control of the situation. Goals and control can be provided and created at multiple scales, from quick, low-level goals such as “get over the bridge in front of you” to long-term, higher-level goals such as “get all the red coins in the world.” Often players work on several goals, at different levels, and on different time scales. This process of
accumulating goals, understanding the world, making a plan and then acting in it, is a powerful means to get the player invested and involved.

The way that players engage with goals as they play is a complex process. As Church mentions, at any moment during a game a player might be working on several nested, interrelated goals. As players construct and work toward short-term and long-term goals, they are actively charting a course through the space of possibility of a game.

In the landscape of a game defined by the space of possibility, short-term goals are navigational beacons that help orient players through two related experiential functions. First, players use short-term goals to make plans. Short-term goals allow players to plan ahead, scouting out future actions, generating hypotheses about how they should play the game. (I'm playing Risk. What happens if I focus on conquering South America next turn?) Second, short-term goals are sources of satisfaction for players. It is one thing to take on a short-term goal, but it is another thing to actually attain it. (I did it! Now I'll get bonus reinforcements for controlling a whole continent.) Short-term goals generate pleasure through both of these functions: making plans as well as achieving them.

Short-term goals are necessary because without them, a player can get lost in the landscape of a game. Are your players confused about what to do next? Perhaps you need to adjust the design to encourage the creation of short-term goals. An open-ended, massively multiplayer online role-playing game such as Ultima Online has an intimidatingly vast space of possibility, but it also provides innumerable opportunities for short-term goals. Are players trying to work their way into a guild? Trade up for an impressive suit of armor? Or just explore a particular section of the world? The game structure encourages each of these short-term goals, and the fact that players can author their own experiences in this way is part of the reason why UO provides such intense pleasure to its dedicated players.

The pleasure a player experiences in a game arises from many simultaneous factors: from the moment-to-moment core mechanic to the short-term accomplishments of play to the final outcome of the game. Each of these interrelated factors generates pleasure in its own way. The core mechanic might provide sensual entrainment, the short-term goals the satisfaction of gradual skill mastery, and the pleasure of winning the spoils of bragging rights. But what links these levels of pleasure is meaningful play. Without meaningful play, a player will never be able to take actions that have predictable outcomes, to choose this over that with a sense of how the choice plays out. Without the ability of players to progress, to have a sense of achievement and accomplishment, to know when they are moving toward or away from victory, your game's play experience will be dead in the water. The play of pleasure may seem free and spontaneous, the farthest thing from a careful, conscious design process, but creating a game that can nourish deep pleasure—that can truly entrain and enrapture players, that can lead to new forms of pleasure and meaning—is always a matter of sensitive and detailed game design.

**Conditioned Pleasure**

Meaningful play is key to designing pleasure in games, but it is only by making choices that meaningful play emerges. If you recall from *Interactivity*, a choice is made up of two primary components: the action that the player takes and the outcome of that action. Our exploration of the core mechanic focused on the action half of the equation: the actual activity that the player performs. So what about the other half—the outcome? One way of framing this facet of the moment of choice is that whenever a player takes an action, she ends up being rewarded or punished by the game as a result.

Psychologists have studied the connection between choice, action, reward, and punishment in a variety of contexts. One useful approach, known as behavior theory, emphasizes observable behavior, specifically the way that interaction with an envi-
ronment shapes behavior. Ivan Pavlov and John B. Watson were early proponents of behavior theory and developed a series of experiments designed to study learned behavior. In one famous experiment, a bell would ring as dogs were fed a meal. Eventually, the dogs came to associate the bell-ringing with food, and would salivate at the sound of the bell. The dogs had been “conditioned” to provide their natural response to food (which was to salivate) even when the food was not present. Pavlov and Watson believed that the same principles could be applied to human behavior. This kind of conditioning, in which innate reflex responses are tied to a new stimulus, became known as classical conditioning.21

Psychologist B. F. Skinner refined the ideas of Watson and Pavlov by rejecting their exclusive emphasis on reflexes and natural conditioning. Instead, Skinner attributed a more active role to the learning subject. According to Skinner’s theory of operant behavior, people learn to behave the way that they do because a certain kind of behavior has been rewarded in the past. If a lab rat learns that pressing a lever results in a food pellet appearing, it is going to develop a strong tendency to press that lever over time.

Behavior theory distinguishes between positive reinforcements (a positive reward, such as a rat getting a food pellet), negative reinforcements (the removal of something unpleasant, like silencing a loud, high-pitched noise), and punishments (the addition of something unpleasant, such as a sudden electric shock). Each kind of reinforcement can be effective in a particular context, usually when the reinforcement or punishment event immediately follows the behavior it is meant to condition. Reinforcements often function because their effects of pain and pleasure are linked to innate biological responses. However, punishments and reinforcements that operate on social and cultural levels can also have strong effects for people. For example, a nod and smile from a teacher can serve as powerful positive reinforcement. In games, these kind of non-biological reinforcements as the outcome of a game choice are common. For example, positive reinforcement in a game might involve giving a player bonus points or an extra life; a negative reinforcement might be eliminating a debilitating disease from a game character; a punishment might be a damaging attack on a player’s character.

Games are systems of meaning. It is within their artificial boundaries that rewards and punishments are interpreted as positive or negative and gain force to shape player behavior. Operant conditioning reminds game designers to pay attention to the way a game encourages or discourages certain behaviors. In creating rewards and punishments, game designers shape the actions players are likely to take in the future. This is an important game design concept, especially in digital games, where the program automates so much of the play activity.

Rewards and Schedules

Operant conditioning not only affects the kinds of choices players make during the course of a game, but also their general motivation to continue playing. More than just shaping good and bad behaviors, rewards and punishments shape a player’s sense of pleasure and overall play experience. Game designers Neal and Jana Halford point out this design challenge:

It’s surprising how many developers forget that it’s the victories and the treasures—not the obstacles—that make people interested in playing in the first place. If you stop giving out the carrots that will keep players excited, or even worse, if you start punishing them for their curiosity, you’re only going to drive away the very people who want to enjoy your game.22

Keeping players engaged in your game as they play is the second of the two seductions of game design. Halford and Halford are absolutely correct that players need to be rewarded, that they need to accomplish tasks and feel satisfaction as they play. Although punishments are important, as balance a play experience needs to be pleasurable. Otherwise, nobody is having any fun.
What kinds of rewards can games offer players? There are as many kinds of rewards as there are forms of play. Halford and Halfor list four general types. Although these categories were written about computer role-playing games, they suggest the kinds of rewards other kinds of games might contain.

- **Rewards of Glory.** Glory rewards are all the things you're going to give to the player that have absolutely no impact on the game play itself but will be things they end up taking away from the experience. This includes winning the game by getting all the way to the end, completing a particularly difficult side quest, or defeating the plots of evil monsters.

- **Rewards of Sustenance.** Rewards of this nature are given so the player can maintain their avatar's status quo and keep all the things they've gained in the game so far. This might include health packs that heal injuries, mana potions that increase a player's magical abilities, high-tech armor that shields a player from e-mag radiation, robots that remove curses or diseases, or even storage boxes or beasts of burden that allow a player's avatar to carry more resources along with them.

- **Rewards of Access.** Rewards of access have three critical features: they allow a player access to new locations or resources that were previously inaccessible, they are generally used only once, and they have no other value to the player once they've been used. Keys, picklocks, and passwords are typical examples of this kind of reward.

- **Rewards of Facility.** Rewards of facility enable a player's avatar to do things they couldn't do before or enhance abilities they already possess. When well handled, they should increase the number of strategies and options that player will have for playing the game. A good example of a facility reward might be a magic orb that lets an avatar walk through a stone wall or a cybernetic software upgrade that lets them shut down enemy gun turrets from a distance.23

Punishments, negative reinforcement, and positive reinforcement are important game design tools. They not only teach players what actions to take and not to take in a game, but also craft larger structures of pleasure. These structures assure that players are properly rewarded for spending the time to take part in the experience designed for them. But using reinforcement successfully in a game means more than just knowing what kinds of pleasures to provide. It is equally important to know how to integrate rewards and punishments into the experiential structures of a game. How often does reinforcement occur? How powerful is the reward or punishment? Do reinforcement factors change over time or remain the same?

Behavior theory has devoted much study to reinforcement schedules. A reinforcement schedule refers to the rate a subject is given reinforcement over time. These reinforcement patterns, along with a network of integrated rewards and punishments, help shape the fabric of any game experience. There are two basic kinds of reinforcement: fixed and variable.

**Fixed reinforcement** means that rewards or punishments are occurring at a steady, continuous rate. A fixed ratio means that the outcome occurs a set number of times that the behavior is performed, such as a player getting a chevron for every five waves of aliens defeated. A fixed interval refers to a regular amount of time between reinforcements, as when a power-up appears in a game every 30 seconds as long as a player can stay alive.

With **variable reinforcement**, the rewards and punishments are coming at irregular intervals. Variable ratio means that the outcome happens after an irregular number of intervals, like slot machine payoffs that occur at a random rate. With a variable interval, the reward or punishment occurs at random time intervals, as in mechanical children's games like Don't Wake Daddy, in which daddy will wake up (with negative consequences) after a random amount of time.
A Hypothetical Case Study

A Game Called Unlocker

Let's invent a fictional game called Unlocker to illustrate several points about classical conditioning. Unlocker is a straightforward 2D computer game where the player controls an avatar seen from a top-down point of view. Moving through a series of rooms, the player must avoid traps, collect weapons, fight pursuing enemies, and collect keys that unlock doors to additional rooms and levels. Although combat can occur, it is not the intended focus of the game. The goal of Unlocker is to unlock as many doors as possible and earn the most points before dying.

Even this relatively simple game contains many kinds of objects and events: open and locked doors, hidden keys, mobile enemies, movement, combat, manipulation of an inventory, and so on. Because many players will not read the instructions (and just as many will forget them soon after reading), how do you teach players what they are supposed to do in the game? Rewards and punishments are one means of shaping their behavior.

The overall trajectory of the game is to open doors and move on to new rooms and new game levels. So you want your players to open those doors. Imagine the first time a player finds a key and uses it on a locked door. A sprite animation plays and the player sees the door swing open in the game space. The problem is that there is nothing in the game to let a player know that unlocking the door is a valuable action that brings the player closer to a positive outcome. The solution? Reward the player! Give the player bonus points for unlocking a door and make sure to add a Ka-ching! sound to emphasize the event. Make a gold star appear in the interface when a player unlocks a door—maybe five gold stars earn an extra life. Or flash a message on the screen that congratulates the player and shows a map to the next locked door. Each of these possible solutions represents a different way to reward the player for the action of unlocking. All of them combine internal, system-based rewards (points, extra lives, information about the next door location) with external, audio-visual rewards (sound effects, gold stars, congratulatory text). If you can craft the proper reward for your player, you will create a desire to achieve that satisfying reward event again, and the player's actions will follow suit.

The same is true of negative reinforcements and punishments. By providing unpleasant feedback, you can teach your player what not to do in your game. Let's say your intention
Fixed schedules are best at shaping behavior if the subject is being punished: sending a child to his room every time he performs an undesired behavior is much more effective than sending him to his room only some of the time. On the other hand, for many kinds of reinforcements, especially positive ones, variable schedules are more effective. In gambling, players are usually rewarded at variable ratios. The repeat play of gamblers is strong evidence of the power of variable reinforcement.

One game designer known for integrating ideas of operant conditioning and reinforcement schedules into his work is Gabe Newell, lead designer on the computer game Half-Life. During a panel discussion at an MIT conference on gaming, Newell discussed the way that Half-Life's design integrates these concepts:

The rewards in Half-Life are getting to see new monsters, the plot is moving forward, getting to have a new fun weapon, getting to see something really cool. ... You want to make sure that throughout the course of the game that they're getting rewards. ... You want to look at it from the point of view of a reinforcement schedule and say OK, that makes sense. I mean there were points in the game before we shipped where there were long lulls where basically all you were doing was stuff that you'd already done before, which in our view didn't represent a reward. So we said we've got to put more fun stuff in here or eliminate that section from the game. 24

Newell makes a number of important points. He begins by listing some of the chief rewards that Half-Life provides for players, which include what Halford and Hallford call rewards of facility (new fun weapons), access (meet new monsters and experience plot twists), and glory ("getting to see something really cool"). He also emphasizes the fact that it is rewards that sustain players through the course of a game. When his design team felt like players were not consistently receiving substantive rewards, they either increased rewards in that section or removed it entirely from the game.

Half-Life utilizes principles of operant conditioning in other ways as well. Much of the success of Half-Life has been attributed to the way that it shapes player experience, creating a thriller-like tension while drawing the player slowly into its dangerous and mysterious spaces. For example, the intensely combative moments in the game are interspersed with uneventful stretches. As opposed to a typical "enemy lurking behind every door" structure, Half-Life creates uncertainty about when and where the terrifying mutant monsters of the game will appear. In some game levels, most doors do not open up to an opponent but instead to empty space. In these levels, Half-Life uses variable ratio punishments. Sometimes the player will be attacked when he opens a door or rounds a corner, but usually he is not. The experiential result of this design strategy is that in Half-Life, deadly threats seem to lurk in every dark shadow and beyond every closed doorway. Deploying enemies with restraint, creating a sparse pattern of unexpected, horrifying encounters, results in a more powerful experience through the use of fewer game elements.

The question arises: if rewards and pleasure are the keys to keeping players involved in a game, why don't tension-inducing punishments such as variable monster attacks drive players away? How can seemingly negative emotions create seductively positive play? We could answer by referring to similar pleasures in other media: the frightening ghost story or the gripping sci-fi cinema thriller. But there is a deeper principle at work. To play a game is to experience its pleasure. At the same time, we know from the lusory attitude that part of playing a game is to take on artificial challenges, inefficiencies adopted for no logical reason except that they make play possible. It is surely challenging to get a golf ball into a tiny little hole so many yards away on the green, but that hardly stops golfers from playing, just as the anxiety of Half-Life doesn't cause players to exit the program. On the contrary: challenge and frustration are essential to game pleasure. Without them, there would be no game conflict to struggle against and no pleasure would emerge from the process of overcoming adversity.
is that the game play of Unlocker is more about finding keys and unlocking doors than about combating enemies. In this case, you may need to find ways to punish players that go against the grain of the game. If combat is too rewarding, either in terms of structure (such as earning points) or experience (such as cool combat effects), your players will keep on fighting because the game intrinsically encourages them to do so.

What are some solutions to this dilemma? You could make the enemies tougher, but that might simply encourage players to rise to the challenge. There are better "punishments" that can steer your players away from fighting. For example, maybe players earn no points from killing enemies, only from finding objects and unlocking doors. Or perhaps when they die as a result of combat, they are sent back to an earlier level. Or when they die, they lose points and game resources. Obviously, you do not want to punish the player too harshly for fighting or dying, but you do want to nudge them in the direction that you designed for them. Of course, if you continue to observe your Unlocker playtesters engaging in too much combat and not enough door unlocking, it might be telling you something else. Maybe something about the game's core mechanic makes combat more compelling than the hunt-and-unlock activities you intended as the game's focus. Perhaps you should turn Unlocker into a combat game. On the other hand, you could always remove the combat component entirely. It's your design. You decide.
Reward and punishment are two sides of a coin, both of them necessary to craft the structure of meaningful experience for players. Finding that elusive balance between positive and negative experience—between anxiety and pleasure—is one of the deepest challenges of game design. In the following section, we engage directly with these important questions.

**Boredom and Anxiety: Flow Redux**

The concepts that come from behavioral psychology—operant conditioning, reinforcement schedules, positive and negative reinforcement—describe how individual rewards and punishments can affect a player’s present and future behavior. The challenge of putting these concepts to use is that the actual context of a game is always more complex than any individual instance of behavioral conditioning.

At any moment in a game, a player is pushed and pulled in many directions at once, experiencing a complex mix of pleasures. Think about the layered pleasures of Half-Life; the terror of the lurking unknown, the vertiginous pleasure of moving in a 3D space, the satisfaction of gradual skill mastery, the strategic exploration of combat possibilities, the glory of finishing the game victorious. Some of these pleasures frustrate the player; others provide banquets of sensual delight. Yet others find their main significance in contexts outside the game. In order to bring the diversity of pleasure into a single understanding of the play experience, we turn once more to the work of psychologist Mihaly Csikszentmihalyi. In his book *Flow*, Csikszentmihalyi provides a general heuristic for understanding how a participant is pulled into and out of the flow state. Although we know that flow is not identical to meaningful play (or even play in general), the model is extremely useful for conceptualizing the play of pleasure in games.

Csikszentmihalyi looks specifically at the degree of challenge that a potential flow activity provides. Does the activity provide tasks that meet the abilities of the participant? What happens when an activity is too difficult, or not difficult enough? What

---

**Terminological Aside: Behavior Theory and Cybernetics**

If all of this talk about positive and negative reinforcement sounds familiar, it should. Cybernetics uses similar language when talking about feedback loops, which we first encountered in *Games as Cybernetic Systems*. Although there is a connection between behavior theory and cybernetics, there is also the possibility for linguistic confusion.

Behavior theory and cybernetics both address the way that repeated actions and outcomes affect the state of a system. If a thermostat keeps telling a heater to blow hot air into a room, that cybernetic system of positive feedback will continue to progress toward higher and higher temperatures. If a dog is continually rewarded with food for performing tricks, it will keep trying to perform tricks in order to get more food.

However, there are also differences in the kind of phenomena studied by each. Cybernetics is concerned with deviation from a stable state; negative feedback returns a system to stability and positive feedback moves the system away from stability. Behavior theory is concerned with intelligent learning behavior, which makes associations between specific actions and outcomes. Although they share some structural affinity, these two ways of thinking are not at all identical.

There are also differences in the way that each uses the terms “positive” and “negative.” The term “positive” has happy connotations in behavior theory, as when a teacher rewards students with praise. Similarly, “negative” has unpleasant connotations, even implying punishment. But in cybernetics, the terms positive and negative are mathematical, and either one might result in an emotionally positive or negative experience for a player. If, in a car racing console game, the car in last place is given an automatic boost, that is an example of a negative feedback system (tendency toward a stable state where all cars are in equal place). The boost will most likely have a positive emotional impact on the losing player.

“Positive reinforcement” and “negative reinforcement” are not universal terms, and mean very different things depending on whether you’re referring to cybernetics or psychology. Be careful how you use them!
are the conditions under which someone can engage with the activity and enter a flow state? In answering these questions, Csikszentmihalyi charts a person’s experience along two axes. One axis represents the degree of challenge an activity offers. The other axis represents the skills a participant possesses.

Both factors can range from a low to a high value, and as a player moves to different positions on the chart, he or she is navigating through different experiences of the activity. The narrow diagonal strip represents a potential flow state, those moments when a player’s skills equally meet the challenges of the activity. On one side of this strip is the state of anxiety, where the activity’s tests exceed the participant’s skills. On the other side is boredom, the state in which the player’s abilities outstrip what challenge the activity can provide.

Csikszentmihalyi uses the example of someone learning to play Tennis. When a Tennis player begins her study, she is at position 1 on the chart, possessing low skills, but also facing challenges appropriate to her abilities, meaning that she may have some initial experiences of flow. As she proceeds, however, she is likely to fall out of flow. If her Tennis skills exceed the challenge of her lessons, the result is an experience that does not fully engage her, and she finds herself in position 2 (boredom). On the other hand, if the sense of challenge she feels from Tennis is overwhelming, the result is a negative and intimidating experience, position 3 (anxiety). Only by finding a new balance between skill and challenge can the Tennis player arrive at position 4 and regain the flow state.

Csikszentmihalyi’s model has a great deal of relevance to game design. How many times have you played a game and had a negative experience, either because it was too difficult to learn or play, or because it was not challenging enough for your skill level? Remember that for games, the concepts of skill and challenge should be interpreted broadly. Skill does not only mean hand-eye coordination or athletic ability. Increasing skill might take the form of greater knowledge of a game system’s rules, more detailed mapping of the game’s narrative world, or increasing confidence to bet larger and larger sums of money at a gambling table.

One of the useful insights gleaned from Csikszentmihalyi’s approach is that his model doesn’t simply address an isolated moment in a game, but tracks a player’s experience over time, over the course of many games. The best games manage to scale their challenge to the player. Ideally, games are simple to learn but difficult to master, providing an appropriate degree of challenge for beginners and advanced players alike. The two “traps” into which poorly designed games can fall—boredom and anxiety—are extremely useful ways of thinking about your players’ game experience. Can you provide proper challenge at every stage of the game, for all levels of players? During playtesting, keep a sharp eye out for players encountering boredom and anxiety and note when these moments occur. What was the context? What kinds of decisions and outcomes were happening during these moments? What kinds of player strategies led to boredom or anxiety?

We can also frame boredom and anxiety in terms of meaningful play. Both states represent poorly designed moments of choice. If a player is feeling boredom, for example, she is not meaningfully exploring the space of possibility of a game.
Perhaps the outcome of a decision is not tied closely enough to the action, and so events in the game feel arbitrary. Boredom is "dead space" in a game, moments when the taut space of possibility falls limp, when the player is not being confronted with a rich set of choices in an entraining pattern of experience. Playing a game is a dance between a player and the system of the game. When a game is boring, the player's dance partner feels like a lifeless mass that has to be dragged about the dance floor. That doesn't sound like much fun.

Boredom can come from many different sources in a game. In the passage from Gabe Newell, he recalled the disappointing discovery of "long lulls where basically all you were doing was doing stuff that you'd already done before." Even though the player was actually quite busy, the lack of new kinds of rewards wasn't creating sufficient motivation for the player. But too many rewards can create boredom as well. Players often crave power, what Hallford and Hallford would call rewards of facility and sustenance. But too much power leaves players feeling tediously omnipotent and enormously bored, despite their fantastic abilities.

Anxiety operates in a parallel manner. Dramatic tension, as noted in Half-Life, can be a wonderful experiential tool for design. But there is a very thin line between meaningful challenge and truly unpleasant anxiety. Imagine if the tension in Half-Life never let up, if the player never received rewards to balance variable punishments, if the player was never able to feel an actual sense of accomplishment. The game experience would feel like a series of gratuitous attacks, with no justification or end in sight.

When the play of a game becomes synonymous with anxiety, the experience is surprisingly similar to boredom. No matter what choice a player makes, it feels like negative outcomes will always result, and choices in the game therefore feel arbitrary. The space of possibility becomes stifling and inert. To use the dancing metaphor once again, when game experience becomes synonymous with anxiety, the system of the game takes over the dance completely, like an overpowering robot, and the player is trapped in a series of actions over which she has no control. Once again, meaningful play fails to occur.

Anxiety sometimes results because in seeking to design a challenging game, game designers create games that are too challenging, especially for novice players. Remember the level playing field of conflict from Games as Systems of Conflict? It is important that players feel a sense of fairness as they play, that they win or lose because of the application of their own abilities within an equitable game system. This is why many games have handicapping rules or player classes, so that players of equal skill can be matched up against each other. With single-player digital games, it can be more challenging to anticipate and balance challenge and anxiety. In Games as Cybernetic Systems, we looked at an important approach to challenge management called Dynamic Difficulty Adjustment, used in games like Crash Bandicoot and Jak and Daxter.

Csikszentmihalyi's model of boredom and anxiety applies to games in many ways. It can help us understand how a player navigates the terrain of skill and challenge within a single game, for example, or to understand the way game skills are slowly built up over time. Csikszentmihalyi's model can also serve as a tool within the game design process. An iterative design process allows game designers to locate moments of boredom and anxiety in their game and re-shape the game experience to minimize moments of less meaningful play.

**Anxiety and Boredom on the High Seas**

In a wonderful essay published on Gamasutra.com, Jesse Schell and Joe Shochet of Disney Imagineering write about the process of designing Pirates of the Caribbean—Battle for the Buccaneer Gold, a game "ride" where a group of players stands on a motion-platform pirate ship surrounded by video projections. During the game, one player steers the ship while the other players operate a number of cannons, firing at monsters, forts, and enemy vessels. Pirates of the Caribbean is designed as a condensed five-minute experience, and it was essential that players feel properly challenged at every moment of the game.
In their design analysis, Schell and Shocet detail a number of design problems that had to be overcome in order to maximize player enjoyment. For example, during playtesting they identified as a problem the fact that the player steering the ship could take the ship to what they call "dull places," leading to a less engaging experience for all of the players. In the selected quotes below, Schell and Shocet outline some solutions to this problem:

*Architectural Weenies:* "Weenie" is a phrase coined by Walt Disney himself. It refers to the technique used on movie sets of guiding stage dogs by holding up part of a sausage. In the case of Pirates, there are three main "weenies," one for each island: a volcano, an enormous fort, and a plume of smoke coming from a burning town. No matter which way the boat is facing, at least one of these "weenies" is in view. Since the coolest action takes place at the islands, [we wanted] to guide the captains to go there.

*Guide Ships:* Since the short-term goal of the game is to fire on other pirate ships, captains strive to get near these ships so that their gunners can get a clear shot. Many of the ships in the Pirates world are "on their way" to the islands mentioned above. Many captains, in just trying to stay near these ships find that just as they have destroyed the ship, they have arrived at one of the islands, without even trying to get there.

*Sneak attacks:* What if the captain ignores the guide ships? Even if he heads toward one of the "weenies" it might mean as long as a minute during which the gunners have little to shoot at. For this reason, [we] created special "sneak attack" ships that "magically" appear behind the players' ship, and quickly pull up alongside, when no other boats are in range.

*The Waterspout:* This was a nickname for a "last ditch" forcefield that surrounds the game play area. If a captain tries to sail out of the main game play area and out to open sea, they hit the forcefield, and the ship is "magically" pointed back to where the action is. The few guests who see this don't even realize that anything unusual has happened. They are just pleased to have their boat going somewhere cool.

Schell and Shocet are thinking in very experiential terms, using clever techniques to subtly guide player action in meaningful directions. At the time of its release, Pirates was a very high-tech production, featuring real-time 3D graphics, physically engaging cannon-firing interfaces, and a large motion platform to simulate a pirate ship rocking on the waves. Often in these instances, a desire to "properly" simulate a coherent 3D space or "correctly" output logical behavior for computer-controlled characters overshadows the design of the actual play experience. But Schell and Shocet had no hesitation in making pirate ships "magically" appear to guide the player, or abandoning "realistic" physics to have the player's ship turn on a dime to facilitate navigation. As they put it, "By choosing to be less concerned with reality and more concerned with what was fun, we created an experience that... is easier to adapt to, quicker to learn, and is a better show." In game design, player experience should always trump so-called "realism."

Boredom and anxiety, as game design watchwords, are wonderful because they speak directly to player experience. As you shape and sculpt your players' pleasure, you are guiding them between the Scylla and Charybdis of anxiety and boredom. This task is made all the more difficult because, as we know, the experience of play can only be indirectly designed. How do you create a set of rules that maximizes the play of pleasure for your audience?

**Meaningful Pleasure**

We can identify elements of the play of pleasure through concepts such as repetition and entrainment, short-term and long-term goals, rewards and punishments, and anxiety and boredom. Ultimately, however, a player experiences a more pervasive sense of pleasure and enjoyment; a total feeling of engagement that arises directly from play, the experiential
whole that is more than the sum of the parts. Pleasure is emergent. Constructing the rules of a game, the formal system that produces this pleasure, is the challenge of game design. As usual, the key to understanding is meaningful play.

The core of meaningful play lies in the relationship between action and outcome. As a player uses core mechanics to take action, outcomes accumulate. These outcomes take many forms: sensory feedback, strategic achievement, emotional gratification, social relationships, and so on. As a player advances through a game, it is crucial that the game provide meaningful play at every moment. For example, as a player achieves a short-term goal, the movement toward, through, and beyond that goal should be clear. The game must communicate where the goal is, how it might be achieved, whether the player is making progress toward it, exactly when it was reached and completed, and its impact on future play. There is room in this experience for uncertainty and ambiguity, but a certain kind of clarity must underlie every action in a game. Even in the inexact, messy realms of pleasure and desire, every game choice must be discernable and integrated.

When game actions are discernable, the events of a game and the outcomes of choices are always evident. Discernable outcomes drive the experience of meaningful play and facilitate pleasure. In DOOM, for example, the monster opponents that players battle hardly exist on their own. Until the player enters a room full of monsters, they will “idle,” walking in place, waiting for the player to enter so that they can spring to life and attack. Although some players regard this aspect of DOOM as comically impoverished, in fact it is key to the successful play of the game. Because monsters have little or no life “off camera,” all of their important activity happens “in the face” of the player as he encounters them in battle. The game events that result in rewards and punishments for the player are always clearly communicated because they almost always occur in the presence of a player.

The need for events to be integrated into the larger fabric of the game experience is perhaps even more important than discernability in sustaining player pleasure. As long as a player understands the implications of the game’s system of rewards and punishments, he or she can use that knowledge to set new short-term goals. This allows the player to maintain an overall sense of progress toward a long-term goal, such as winning. In the popular online game Neopets, a player is continually rewarded for taking game actions, exploring the game world, caring for her pets, playing simple games, and interacting with other players. Each of these simple activities rewards the player with a small amount of points. These points are then used to facilitate new purchases, which in turn make new activities possible. The steady stream of incremental rewards forms a tight loop of desire, a compelling system of pleasure where short-term and long-term goals are constantly forming on the horizon of player action.

When game events are not discernable or integrated, boredom and anxiety, the enemies of pleasurable flow, can result. Does the game program know that a player just took an action? Why did all of those important events happen off-screen? Does it matter which piece a player just moved? Why is the game so hard? Designing for meaningful play comes down to treating players with great care and concern at every moment of the game. Too often, for example, a digital game just doesn’t feel right. The interface is clunky, the player is not sure what to do when the game begins, or the first level is too hard. Retail digital games are usually designed for 30 to 40 hours of play. That kind of commitment demands a tremendous amount of trust. If the first five minutes are unpleasant, why would a player want to continue?

There is a reason why Myst was superior to all of the CD-ROM multimedia game clones that followed it, or why Super Mario 64 is still better than the scores of 3D over-the-shoulder, character-based console games that are released every year. Myst
NeoPets

and Super Mario 64, although very different in the experiences they provide, have one thing in common: they both treat the player with a tremendous amount of care. From the moment the game begins, the player has clear direction and purpose. As players explore their expansive worlds, both games provide a satisfying increase in challenge, while never leaving the player feeling lost or confused. There is clarity to the way that these games construct player pleasure.

Crafting this degree of pleasure is extremely challenging. Pleasure is difficult to design because it is an open-ended, multifaceted, and exceedingly complex concept. But that is also why it is such a fertile avenue for exploration by game designers. There are multitudes of game pleasures for you to create: pleasures that go deep into the hearts of your players; pleasures that transform your players and the ways that they understand the world; pleasures that expand the very medium of games. The process of discovering and inventing these pleasures is itself a unique form of bliss: the boundless joy of game design.

Against “Addiction”

The great damnation of the game [of Chess] has come from those who have been plagued by it. None has expressed so convincingly his sad and resigned self-denial as a minister who in 1680 wrote a letter, giving ten reasons why he refused to play the game. Among them is one of the most beautiful lines in English literature: “It hath not done with me when I have done with it.” Truly this one sentence could be the motto for all addictions.—Norman Reider, “Chess Oedipus, and the Mater Dolorosa”

The play of pleasure in games is immensely complex, but we have done our best to trace some of its contours. Before ending this chapter, there is an additional issue we must address: addiction. Addiction and addictive play can mean many things. But by and large, among game designers, addiction is considered a positive trait, the mark of compelling play. In business terms, lots of addicted players mean that a game has a greater chance of being a commercial success.
Meaningful play can become addictive. If a player enjoyed the play of a game, he or she will probably want to play it again. If you create a space of possibility that rewards players for exploration, then you are likely to have players that want to see more permutations of how the rules play out. The same-but-different quality intrinsic to all games is at the core of a game's ability to engross players and bring them back into the magic circle again and again. "Addiction" in this sense is merely shorthand for a game experience that can support this depth of meaningful play.

As a game designer, it is flattering to find that players are addicted to your game. It might be that they use your game regularly to relax or unwind. Maybe they find your game a great way to interact with friends. Or perhaps they write fiction around your game's storyline and participate in the fan culture your game has spawned. All of these forms of so-called "addiction" are the mark of dedicated players, of meaningfully engaged people experiencing the play of pleasure provided by a game.

At the same time, there are negative connotations to the word "addiction" as well. Medically speaking, addiction is a genuine disorder, whether the addiction is to substances like alcohol or drugs, to negative behaviors like bullying or shoplifting, or to behaviors that are generally considered positive, like exercise or reading. Suffice it to say that the use of term "addiction," when used by professionals in the game industry, does not describe medically pathological behavior. Instead, it refers to engaged and repeated play, to players that enjoy a game and therefore play it more than one time.

Because of the negative connotations of the term, the repeatable play of games is sometimes naively compared to a genuine medical disorder. But the word "addiction" is a misnomer, as play scholar Brian Sutton-Smith points out:

The persistent concentration we are talking about is sometimes mistaken for addiction. But its compulsive quality is the same experience by those who have fallen in love, or are taken by some hobby or sport, ... It is not an addiction where what occurs is a surrender to outside forces over which one has no control. We must distinguish such compulsive avocations from addiction. Video games are of this first kind. Our proposal, then, is that video games, like all other forms of exciting play, lead to a compulsive and persistent attendance on the games themselves. In this, they are like all games and all play which has long been noted for holding children's attention when they should be coming inside for their supper, or leaving the playground to go into school.

Play is intrinsically engaging. But that doesn't mean that it is negatively addictive. It is true that some forms of play can become pathological. People can become compulsive gamblers, or they can spend so many hours in an online MUD that they neglect aspects of their life outside the game. These rare cases, often highly publicized, are the exceptions that prove the rule. The overwhelming majority of play phenomena are not destructively addictive. This is true even for forms of play most commonly associated with pathological addiction, such as gambling. In The Ambiguity of Play, Sutton-Smith presents extensive research on gambling with the conclusion that "the majority of players gamble moderately and with positive results for family life and pleasure." The existence of addictive play disorders doesn't mean that all play is bad for you. Eating disorders and addictions abound. But that doesn't mean that you should avoid the pleasure of dining.

To play is to find free movement within a more rigid structure. When a game activity becomes pathologically addictive, this movement is censured: free movement is shut down, the sense of free choice evaporates, and meaningful play abates. In this experiential sense, when a player becomes medically addicted to some form of play, play as we have defined it no longer exists. In other words, addictive play, in the negative sense used by the medical community, is not really play at all.
Case Study: The L Game

An Exception to Every Rule

One game that flies in the face of some of our ideas about meaningful play and pleasure is the L Game, designed by Edward de Bono, a writer and researcher who focuses on lateral thinking and creative problem-solving. The rules are summarized below.  

How to play the L Game

- **Pieces.** The Board is made up of 16 squares. Each player (only 2 can play) has an L piece that he must move when it is his turn. There are also two neutral pieces that either player can move.

- **Object.** The object of the game is to maneuver the other player into a position on the board where he cannot move his L piece.

- **Starting Position.** Proceeding from the starting position, the first player (and each player on each move thereafter) must move the L piece first. When moving, a player may slide, turn or pick up and flip the L piece into any open position other than the one it occupied prior to the move. When the L piece has been moved, a player may move either one (but only one) of the neutral square pieces to any open square on the board. It is not required that the neutral piece be moved, this is up to the player! A player wins the game when his opponent cannot move his L piece.
Further Reading

"Designing Interactive Theme Park Rides: Lessons From Disney’s Battle for the Buccaneer Gold," by Jesse Schell and Joe Shochet (see page 68)

Flow: The Psychology of Optimal Experience,
by Mihaly Csikszentmihalyi
One of the great pleasures of games is the feeling of being in flow—a state of deep and all-encompassing absorption with the activity at hand. Csikszentmihalyi studies the qualities and conditions that allow for flow, which have many connections to the qualities and conditions of game play.

Recommended:
Chapter 3: Enjoyment and the Quality of Life
Chapter 4: The Conditions of Flow
Chapter 7: Work as Flow

Notes
1. Wordnet, Princeton University.
3. Ibid. p. 548
11. Csikszentmihalyi, Flow, p. 3.
12. Ibid. p. 49.
19. Herz, Joystick Nation, p. 64.
The L Game

It is difficult to get a sense of the L game without actually playing it, but the game essentially consists of players trying to place their pieces in such a way so as to keep their opponents from making a legal move. The two players end up shuffling pieces on the small but crowded board, taking turn after turn, until one of them hits upon a winning move.

The challenge of the L Game is that it does not provide clear feedback for players as they progress towards the goal. In Checkers, even a beginner can get a sense of the game’s progress: if white has lost most of its pieces and black’s pieces are all still on the board, then black is clearly progressing toward victory. But in the L Game, because the pieces are never removed and do not progress step by step toward a victory condition, it is very difficult to tell which player is gaining or losing ground. Playing the L Game can feel more like taking on an arbitrary, frustrating puzzle than playing a game.

Jeff Fedderson, a student in a class taught by Eric Zimmerman and Frank Lantz, analyzed the game. He found the game strategy so opaque that he wrote a program to play the L Game against itself. He experimented with strategies for his computer players and identified several patterns: for example, programs that attempted to occupy the center four squares of the board and avoided putting the long edge of an L shape on the edge of the board were much better at playing the game.

But even armed with these very short-term strategic goals, the L Game remains stubbornly resistant to providing meaningful feedback about progress toward the end goal. However, this strange feature of the L Game is actually what makes it so distinctive and compelling as a play experience. Despite the fact that the L Game seems to violate some of our most basic ideas about meaningful play, it still provides pleasure. Sometimes, when design rules are broken in a very original way, whole new modes of play can be invented.
Pleasure is intrinsic to games in many ways. The act of playing a game, submitting to a set of rules, is itself a form of pleasure. The restraint that limiting game behavior affords heightens the player's sense of pleasure.

Games provide autotelic pleasures, experiences that are pursued for their own sake. Although it is true that games provide extrinsic pleasures that affect a player's life outside the game, all games also provide intrinsic, autotelic pleasures that are significant only within the artificial meanings that the game creates.

Games must provide a double seduction for players. First, players must be seduced into entering the magic circle. Second, players must be continually seduced into remaining inside the circle of play.

There are many established typologies that address the forms of pleasure provided by games. Typologies of game pleasure are generally less useful for theorizing pleasure and more useful for organizing observations about game experience.

Psychologist Mihaly Csikszentmihalyi describes optimal experience as flow. Flow is the exhilarating pleasure that occurs when someone is engaged with an activity and feels in control of his or her actions. Although flow is not unique to games, it is a useful way of thinking about the creation of game pleasure.

Csikszentmihalyi names eight characteristics of flow, each of which has a strong connection to games. Four of the eight characteristics describe the effects of flow:

- the merging of action and awareness
- concentration
- the loss of self-consciousness
- the transformation of time

The other four characteristics describe the prerequisites of the kind of activity that will result in flow:

- a challenging activity
- clear goals
- clear feedback
- the paradox of having control in an uncertain situation
- Games possess a quality we call same-but-different. Every time one plays a game, the formal structure remains the same, but the way the rules play out are different. This quality of games makes it pleasurable for players to explore the space of possibility. Because play is often transformative, the continued exploration of a game can change the game structure itself, leading to a potentially endless sequence of same-but-different pleasures.

- Entrainment means both to carry along and to trap. Entrainment is the process of falling into the rhythmic patterns of pleasure that games can provide.

- The goal of a game is a key component in shaping the experience of pleasure. The goal is the object of desire held out to entice players to continue playing.

- Suspended between moment-to-moment core mechanics and the ultimate end goal are short-term goals. Short-term goals help players make plans in a game as well as provide moments of satisfaction when they reach them.

- Behavior theory is a branch of psychology that studies observable behavior. Conditioning is the acquisition of learned behaviors through rewards and punishments. Rewards and punishments can be used to teach players how to behave in a game from moment-to-moment, as well as create an experience that rewards players for their participation over time.

- Reinforcement schedules refer to the rate at which players receive rewards and punishments. Schedules can be either fixed or variable in regard to an interval of time or the ratio of action to outcome. Generally, variable reinforcement shapes behavior more powerfully than fixed reinforcement.

- Challenge is an important way to shape player pleasure. If the challenge of a game is too high for a player's skills, anxiety results. If there is not enough challenge, boredom results. Ideally, games provide a balanced challenge at all moments.

- A game designer can only indirectly design the pleasurable experience of a game through the creation of a game's rules. To provide meaningful play and pleasure, the actions a player takes must be discernable and integrated.

- The word addiction has different meanings in the medical community and the game community. For game designers, addiction is a positive quality that signifies players' meaningful interaction with a game. When pathologically addictive behavior emerges from a game, play is no longer possible.
The Lessons of Lucasfilm’s Habitat

F. Randall Farmer and Chip Morningstar

Context
This paper was presented at The First Annual International Conference on Cyberspace in 1990. It was published in Cyberspace: First Steps, Michael Benedikt (ed.) [MIT Press, 1990].

Gaming the Game
The Game Design Process
Game Communities

For more than 30 years, F. Randall Farmer has been connecting people with each other using computers as the mediating technology. He has co-created the following: one of the first online forums, the first Trek MUD, the first graphical MMOG with the first avatars, the first virtual MMOG currency, the first virtual information marketplace, the first fully distributed virtual world platform, the first no-plugin web session platform, and more. He continues to publish on these topics.

Chip Morningstar was one of the founders of Electric Communities, a cyberspace design and development company. He was also heavily involved in the initial development of Fujitsu’s WorldsAway service, for which Electric Communities provided creative and technical oversight. Morningstar worked at Lucasfilm Ltd where he was the designer and project leader for Lucasfilm’s Habitat, the world’s first large scale commercial multiperson online graphical virtual world.
Introduction
Lucasfilm’s Habitat was created by Lucasfilm Games, a division of LucasArts Entertainment Company, in association with Quantum Computer Services, Inc. It was arguably one of the first attempts to create a very large scale commercial multi-user virtual environment. A far cry from many laboratory research efforts based on sophisticated interface hardware and tens of thousands of dollars per user of dedicated computer power, Habitat is built on top of an ordinary commercial online service and uses an inexpensive—some would say “toy”—home computer to support user interaction. In spite of these somewhat plebeian underpinnings, Habitat is ambitious in its scope. The system we developed can support a population of thousands of users in a single shared cyberspace. Habitat presents its users with a real-time animated view into an online simulated world in which users can communicate, play games, go on adventures, fall in love, get married, get divorced, start businesses, found religions, wage wars, protest against them, and experiment with self-government.

The Habitat project proved to be a rich source of insights into the nitty-gritty reality of actually implementing a serious, commercially viable cyberspace environment. Our experiences developing the Habitat system, and managing the virtual world that resulted, offer a number of interesting and important lessons for prospective cyberspace architects. The purpose of this paper is to discuss some of these lessons. We hope that the next generation of builders of virtual worlds can benefit from our experiences and (especially) from our mistakes.

Due to space limitations, we won’t be able to go into as much technical detail as we might like; this will have to be left to a future publication. Similarly, we will only be able to touch briefly upon some of the history of the project as a business venture, which is a fascinating subject of its own. Although we will conclude with a brief discussion of some of the future directions for this technology, a more detailed exposition on this topic will also have to wait for a future article.

The essential lesson that we have abstracted from our experiences with Habitat is that a cyberspace is defined more by the interactions among the actors within it than by the technology with which it is implemented. While we find much of the work presently being done on elaborate interface technologies—Data Gloves, head-mounted displays, special-purpose rendering engines, and so on—both exciting and promising, the almost mystical euphoria that currently seems to surround all this hardware is, in our opinion, both excessive and some-
what misplaced. We can’t help having a nagging sense that it’s all a bit of a distraction from
the really pressing issues. At the core of our vision is the idea that cyberspace is necessarily
a multiple-participant environment. It seems to us that the things that are important to the
inhabitants of such an environment are the capabilities available to them, the characteristics of
the other people they encounter there, and the ways these various participants can affect one
another. Beyond a foundation set of communications capabilities, the technology used to present
this environment to its participants, while sexy and interesting, is a peripheral concern.

What Is Habitat?

Habitat is a “multi-player online virtual environment” [its purpose is to be an entertainment
medium; consequently, the users are called “players”]. Each player uses his or her home
computer as a front end, communicating over a commercial packet-switching data network
to a centralized backend system. The front end provides the user interface, generating a real-
time animated display of what is going on and translating input from the player into requests
to the backend. The backend maintains the world model, enforcing the rules and keeping
each player’s front end informed about the constantly changing state of the universe. The
backend enables the players to interact not only with the world but also with each other.

Habitat was inspired by a long tradition of “computer hacker science fiction”, notably
Vernor Vinge’s novel, True Names¹, as well as many fond childhood memories of games of
make-believe, more recent memories of role-playing games and the like, and numerous other
influences too thoroughly blended to pinpoint. To this we add a dash of silliness, a touch of
cyberpunk,²,³ and a predilection for object-oriented programming.⁴

The initial incarnation of Habitat uses a Commodore 64 for the frontend. One of the
questions we are asked most frequently is, “Why the Commodore 64?” Many people somehow
get the impression that this was technical decision, but the real explanation has to do with
business, not technology. Habitat was initially developed by Lucasfilm as commercial product
for QuantumLink, an online service [then] exclusively for owners of the Commodore 64. At the
time we started (1985), the Commodore 64 was the mainstay of the recreational computing
market. Since then it has declined dramatically in both its commercial and technical signifi-
cance. However, when we began the project, we didn’t get a choice of platforms. The nature of
the deal was such that both the Commodore 64 for the front end and the existing Quantum Link
host system [a brace of Stratus fault-tolerant minicomputers] for the backend were givens.
A typical Habitat scene

The largest part of the screen is devoted to the graphics display. This is an animated view of the player’s current location in the Habitat world. The scene consists of various objects arrayed on the screen, such as the houses and tree you see here. The players are represented by animated figures that we call “Avatars.” Avatars are usually, though not exclusively, humanoid in appearance. In this scene you can see two of them, carrying on a conversation.

Avatars can move around, pick up, put down and manipulate objects, talk to each other, and gesture, each under the control of an individual player. Control is through the joystick, which enables the player to point at things and issue commands. Talking is accomplished by typing on the keyboard. The text that a player types is displayed over his or her Avatar’s head in a cartoon-style “word balloon”.

The Habitat world is made up of a large number of discrete locations that we call “regions.” In its prime, the prototype Habitat world consisted of around 20,000 of them. Each region can adjoin up to four other regions, which can be reached simply by walking your Avatar to one or another edge of the screen. Doorways and other passages can connect to additional regions. Each region contains a set of objects which define the things that an Avatar can do there and the scene that the player sees on the computer screen.

Some of the objects are structural, such as the ground or the sky. Many are just scenic, such as the tree or the mailbox. Most objects, however, have some function that they perform.
For example, doors transport Avatars from one region to another and may be opened, closed, locked and unlocked. ATMs (Automatic Token Machines) enable access to an Avatar’s bank account. Vending machines dispense useful goods in exchange for Habitat money. Habitat contained its own fully-fledged economy, with money, banks, and so on. Habitat’s unit of currency is the Token, owing to the fact that it is a token economy and to acknowledge the long and honorable association between tokens and video games.

<table>
<thead>
<tr>
<th>Object Class</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATM</td>
<td>Automatic Token Machine; access to an Avatar’s bank account</td>
</tr>
<tr>
<td>Avatar</td>
<td>Represents the player in the Habitat world</td>
</tr>
<tr>
<td>Bag, Box</td>
<td>Containers in which things may be carried</td>
</tr>
<tr>
<td>Book</td>
<td>Document for Avatars to read [e.g., the daily newspaper]</td>
</tr>
<tr>
<td>Bureaucrat-in-a-box</td>
<td>Communication with system operators</td>
</tr>
<tr>
<td>Change-o-matic</td>
<td>Device to change Avatar gender</td>
</tr>
<tr>
<td>Chest, Safe</td>
<td>Containers in which things can be stored</td>
</tr>
<tr>
<td>Club, Gun, Knife</td>
<td>Various weapons</td>
</tr>
<tr>
<td>Compass</td>
<td>Points direction to West Pole</td>
</tr>
<tr>
<td>Door</td>
<td>Passage from one region to another; can be locked</td>
</tr>
<tr>
<td>Drugs</td>
<td>Various types; changes Avatar body state, e.g., cure wounds</td>
</tr>
<tr>
<td>Elevator</td>
<td>Transportation from one floor of a tall building to another</td>
</tr>
<tr>
<td>Flashlight</td>
<td>Provides light in dark places</td>
</tr>
<tr>
<td>Fountain</td>
<td>Scenic highlight; provides communication to system designers</td>
</tr>
<tr>
<td>Game piece</td>
<td>Enables various board games: backgammon, checkers, chess, etc.</td>
</tr>
<tr>
<td>Garbage can</td>
<td>Disposes of unwanted objects</td>
</tr>
<tr>
<td>Glue</td>
<td>System building tool; attaches objects together</td>
</tr>
<tr>
<td>Ground, Sky</td>
<td>The underpinnings of the world</td>
</tr>
<tr>
<td>Head</td>
<td>An Avatar’s head; comes in many styles; for customization</td>
</tr>
<tr>
<td>Key</td>
<td>Unlocks doors and other containers</td>
</tr>
<tr>
<td>Knick-knack</td>
<td>Generic inert object; for decorative purposes</td>
</tr>
<tr>
<td>Magic wand</td>
<td>Various types, can do almost anything</td>
</tr>
<tr>
<td>Paper</td>
<td>For writing notes, making maps, etc.; used in mail system</td>
</tr>
<tr>
<td>Pawn machine</td>
<td>Buys back previously purchased objects</td>
</tr>
<tr>
<td>Plant, Rock, Tree</td>
<td>Generic scenic objects</td>
</tr>
<tr>
<td>Region</td>
<td>The foundation of reality</td>
</tr>
<tr>
<td>Sensor</td>
<td>Various types, detects otherwise invisible conditions in the world</td>
</tr>
<tr>
<td>Sign</td>
<td>Allows attachment of text to other objects</td>
</tr>
<tr>
<td>Stun gun</td>
<td>Non-lethal weapon</td>
</tr>
<tr>
<td>Teleport booth</td>
<td>Means of quick long-distance transport; analogous to phone booth</td>
</tr>
<tr>
<td>Tokens</td>
<td>Habitat money</td>
</tr>
<tr>
<td>Vendroid</td>
<td>Vending machine; sells things</td>
</tr>
</tbody>
</table>
Many objects are portable and may be carried around in an Avatar’s hands or pockets. These include various kinds of containers, money, weapons, tools, and exotic magical implements. Listed here are some of the most important types of objects and their functions. The complete list of object types numbers in the hundreds.

**Implementation**

The following, along with several programmer-years of tedious and expensive detail that we won’t cover here, is how the system works:

At the heart of the *Habitat* implementation is an object-oriented model of the universe.

The front end consists of a system kernel and a collection of objects. The kernel handles memory management, display generation, disk I/O, telecommunications, and other “operating system” functions. The objects implement the semantics of the world itself. Each type of *Habitat* object has a definition consisting of a set of resources, including animation cels to drive the display, audio data, and executable code. An object’s executable code implements a series of standard behaviors, each of which is invoked by a different player command or system event. The model is similar to that found in an object-oriented programming system such as Smalltalk, with its classes, methods and messages. These resources consume significant amounts of scarce front end memory, so we can’t keep them all in core at the same time. Fortunately, their definitions are invariant, so we simply swap them in from disk as we need them, discarding less recently used resources to make room.

When an object is instantiated, we allocate a block of memory to contain the object’s state. The first several bytes of an object’s state information take the same form in all objects, and include such things as the object’s screen location and display attributes. This standard information is interpreted by the system kernel as it generates the display and manages the run-time environment. The remainder of the state information varies with the object type and is accessed only by the object’s behavior code.

Object behaviors are invoked by the kernel in response to player input. Each object responds to a set of standard verbs that map directly onto the commands available to the player. Each behavior is simply a subroutine that executes the indicated action; to do this it may invoke the behaviors of other objects or send request messages to the backend. Besides the standard verb behaviors, objects may have additional behaviors which are invoked by messages that arrive synchronously from the backend.
The backend also maintains an object-oriented representation of the world. As in
the front end, objects on the backend possess executable behaviors and in-memory state
information. In addition, since the backend maintains a persistent global state for the entire
Habitat world, the objects are also represented by database records that may be stored on
disk when not “in use”. Backend object behaviors are invoked by messages from the front
end. Each of these backend behaviors works in roughly the same way: a message is received
from a player’s front end requesting some action; the action is taken and some state changes
to the world result; the backend behavior sends a response message back to the front end
informing it of the results of its request and notification messages to the front ends of any
other players who are in the same region, informing them of what has taken place.

The Lessons
In order to say as much as we can in the limited space available, we will describe what we
think we learned via a series of principles or assertions surrounded by supporting reason-
ing and illustrative anecdotes. A more formal and thorough exposition will have to come later
in some other forum where we might have the space to present a more comprehensive and
detailed model.

We mentioned our primary principle above:

A multi-user environment is central to the idea of cyberspace.

It is our deep conviction that a definitive characteristic of a cyberspace system is that
it represents a multi-user environment. This stems from the fact that what [in our opinion]
people seek in such a system is richness, complexity and depth. Nobody knows how to produce
an automaton that even approaches the complexity of a real human being, let alone a society.
Our approach, then, is not even to attempt this, but instead to use the computational medium
to augment the communications channels between real people.

If what we are constructing is a multi-user environment, it naturally follows that
some sort of communications capability must be fundamental to our system. However, we
must take into account an observation that is the second of our principles:

Communications bandwidth is a scarce resource.

This point was rammed home to us by one of Habitat’s nastier externally imposed de-
sign constraints, namely that it provide a satisfactory experience to the player over a 300 baud
serial telephone connection [one, moreover, routed through commercial packet-switching
networks that impose an additional, uncontrollable latency of 100 to 5000 milliseconds on each packet transmitted).

Even in a more technically advanced network, however, bandwidth remains scarce in the sense that economists use the term: available carrying capacity is not unlimited. The law of supply and demand suggests that no matter how much capacity is available, you always want more. When communications technology advances to the point were we all have multi-gigabaud fiber optic connections into our homes, computational technology will have advanced to match. Our processors’ expanding appetite for data will mean that the search for ever more sophisticated data compression techniques will still be a hot research area (though what we are compressing may at that point be high-resolution volumetric time-series or something even more esoteric).  

Computer scientists tend to be reductionists who like to organize systems in terms of primitive elements that can be easily manipulated within the context of a simple formal model. Typically, you adopt a small variety of very simple primitives which are then used in large numbers. For a graphics-oriented cyberspace system, the temptation is to build upon bit-mapped images or polygons or some other graphic primitive. These sorts of representations, however, are invitations to disaster. They arise from an inappropriate fixation on display technology, rather than on the underlying purpose of the system.

However, the most significant part of what we wish to be communicating is human behaviors. These, fortunately, can be represented quite compactly, provided we adopt a relatively abstract, high-level description that deals with behavioral concepts directly. This leads to our third principle:

An object-oriented data representation is essential.

Taken at its face value, this assertion is unlikely to be controversial, as object-oriented programming is currently the methodology of choice among the software engineering cognoscenti. However, what we mean here is not only that you should adopt an object-oriented approach, but that the basic objects from which you build the system should correspond more-or-less to the objects in the user’s conceptual model of the virtual world, that is, people, places, and artifacts. You could, of course, use object-oriented programming techniques to build a system based on, say, polygons, but that would not help to cope with the fundamental problem.
The goal is to enable the communications between machines to take place primarily at the behavioral level [what people and things are doing] rather than at the presentation level [how the scene is changing]. The description of a place in the virtual world should be in terms of what is there rather than what it looks like. Interactions between objects should be described by functional models rather than by physical ones. The computation necessary to translate between these higher-level representations and the lower-level representations required for direct user interaction is an essentially local function. At the local processor, display-rendering techniques may be arbitrarily elaborate and physical models arbitrarily sophisticated. The data channel capacities required for such computations, however, need not and should not be squeezed into the limited bandwidth available between the local processor and remote ones. Attempting to do so just leads to disasters such as NAPLPS.7,8

Once we begin working at the conceptual rather than the presentation level, we are struck by the following observation:

*The implementation platform is relatively unimportant.*

The presentation level and the conceptual level cannot (and should not) be *totally* isolated from each other. However, defining a virtual environment in terms of the configuration and behavior of objects, rather than their presentation, enables us to span a vast range of computational and display capabilities among the participants in a system. This range extends both upward and downward. As an extreme example, a typical scenic object, such as a tree, can be represented by a handful of parameter values. At the lowest conceivable end of things might be an ancient Altair 8800 with a 300 baud ASCII dumb terminal, where the interface is reduced to fragments of text and the user sees the humble string so familiar to the players of text adventure games, "There is a tree here." At the high end, you might have a powerful processor that generates the image of the tree by growing a fractal model and rendering it three dimensions at high resolution, the finest details ray-traced in real-time, complete with branches waving in the breeze and the sound of wind in the leaves coming through your headphones in high-fidelity digital stereo. And these two users might be looking at the same tree in same the place in the same world and talking to each other as they do so. Both of these scenarios are implausible at the moment, the first because nobody would suffer with such a crude interface when better ones are so readily available, the second because the computational hardware does not yet exist. The point, however, is that this approach covers
the ground between systems already obsolete and ones that are as yet gleams in their designers' eyes. Two consequences of this are significant. The first is that we can build effective cyberspace systems today. Habitat exists as ample proof of this principle. The second is that it is conceivable that with a modicum of cleverness and foresight you could start building a system with today's technology that could evolve smoothly as the tomorrow's technology develops. The availability of pathways for growth is important in the real world, especially if cyberspace is to become a significant communications medium [as we obviously think it should].

Given that we see cyberspace as fundamentally a communications medium rather than simply a user interface model, and given the style of object-oriented approach that we advocate, another point becomes clear:

*Data communications standards are vital.*

However, our concerns about cyberspace data communications standards center less upon data transport protocols than upon the definition of the data being transported. The mechanisms required for reliably getting bits from point A to point B are not terribly interesting to us. This is not because these mechanisms are not essential (they obviously are) nor because they do not pose significant research and engineering challenges (they clearly do). It is because we are focused on the unique communications needs of an object-based cyberspace. We are concerned with the protocols for sending messages between objects, that is, for communicating behavior rather than presentation, and for communicating object definitions from one system to another.

Communicating object definitions seems to us to be an especially important problem, and one that we really didn't have an opportunity to address in Habitat. It will be necessary to address this problem if we are to have a dynamic system. The ability to add new classes of objects over time is crucial if the system is to be able to evolve.

While we are on the subject of communications standards, we would like to make some remarks about the ISO Reference Model of Open System Interconnection. This multi-layered model has become a centerpiece of most discussions about data communications standards these days. Unfortunately, while the bottom 4 or 5 layers of this model provide a more or less sound framework for considering data transport issues, we feel that the model's Presentation and Application layers are not so helpful when considering cyberspace data communications.
We have two main quarrels with the ISO model: first, it partitions the general data communications problem in a way that is a poor match for the needs of a cyberspace system; second, and more importantly, we think it is an active source of confusion because it focuses the attention of system designers on the wrong set of issues and thus leads them to spend their time solving the wrong set of problems. We know because this happened to us. "Presentation" and "Application" are simply the wrong abstractions for the higher levels of a cyberspace communications protocol. A "Presentation" protocol presumes characteristics of the display are embedded in the protocol. The discussions above should give some indication why we feel such a presumption is both unnecessary and unwise. An "Application" protocol presumes a degree of foreknowledge of the message environment that is incompatible with the sort of dynamically evolving object system we envision.

A better model would be to substitute a different pair of top layers: a Message layer, which defines the means by which objects can address one another and standard methods of encapsulating structured data and encoding low-level data types (e.g., numbers); and a Definition layer built on top of the Message layer, which defines a standard representation for object definitions so that object classes can migrate from machine to machine. One might argue that these are simply Presentation and Application with different labels, but we don't think the differences are so easily reconciled. In particular, we think the ISO model has, however unintentionally, systematically deflected workers in the field from considering many of the issues that concern us.

**World Building**

There were two sorts of implementation challenges that Habitat posed. The first was the problem of creating a working piece of technology—developing the animation engine, the object-oriented virtual memory, the message-passing pseudo operating system, and squeezing them all into the ludicrous Commodore 64 (the backend system also posed interesting technical problems, but its constraints were not as vicious). The second challenge was the creation and management of the Habitat world itself. It is the experiences from the latter exercise that we think will be most relevant to future cyberspace designers.

We were initially our own worst enemies in this undertaking, victims of a way of thinking to which we engineers are dangerously susceptible. This way of thinking is characterized by the conceit that all things may be planned in advance and then directly implemented according to the plan's detailed specification. For persons schooled in the design and construction
of systems based on simple, well-defined and well-understood foundation principles, this is a natural attitude to have. Moreover, it is entirely appropriate when undertaking most engineering projects. It is a frame of mind that is an essential part of a good engineer’s conceptual tool kit. Alas, in keeping with Maslow’s assertion that, “to the person who has only a hammer, all the world looks like a nail”, it is a tool that is easy to carry beyond its range of applicability. This happens when a system exceeds the threshold of complexity above which the human mind loses its ability to maintain a complete and coherent model.

One generally hears about systems crossing the complexity threshold when they become very large. For example, the Space Shuttle and the B-2 bomber are both systems above this threshold, necessitating extraordinarily involved, cumbersome and time-consuming procedures to keep the design under control—procedures that are at once vastly expensive and only partially successful. To a degree, the complexity problem can be solved by throwing money at it. However, such capital intensive management techniques are a luxury not available to most projects. Furthermore, although these dubious “solutions” to the complexity problem are out of reach of most projects, alas the complexity threshold itself is not. Smaller systems can suffer from the same sorts of problems. It is possible to push much smaller and less elaborate systems over the complexity threshold simply by introducing chaotic elements that are outside the designers’ sphere of control or understanding. The most significant such chaotic elements are autonomous computational agents (e.g., other computers). This is why, for example, debugging even very simple communications protocols often proves surprisingly difficult. Furthermore, a special circle of living Hell awaits the implementors of systems involving that most important category of autonomous computational agents of all, groups of interacting human beings. This leads directly to our next (and possibly most controversial) assertion:

*Detailed central planning is impossible; don’t even try.*

The constructivist prejudice that leads engineers into the kinds of problems just mentioned has received more study from economists and sociologists\textsuperscript{10-15} than from researchers in the software engineering community. Game and simulation designers are experienced in creating virtual worlds for individuals and small groups. However, they have had no reason to learn to deal with large populations of simultaneous users. Since each user or group is unrelated to the others, the same world can be used over and over again. If you are playing
an adventure game, the fact that thousands of other people elsewhere in the [real] world are playing the same game has no effect on your experience. It is reasonable for the creator of such a world to spend tens or even hundreds of hours crafting the environment for each hour that a user will spend interacting with it, since that user’s hour of experience will be duplicated tens of thousands of times by tens of thousands of other individual users.

Builders of online services and communications networks are experienced in dealing with large user populations, but they do not, in general, create elaborate environments. Furthermore, in a system designed to deliver information or communications services, large numbers of users are simply a load problem rather than a complexity problem. All the users get the same information or services; the comments in the previous paragraph regarding duplication of experience apply here as well. It is not necessary to match the size and complexity of the information space to the size of the user population. While it may turn out that the quantity of information available on a service is a function of the size of the user population, this information can generally be organized into a systematic structure that can still be maintained by a few people. The bulk, wherein the complexity lies, is the product of the users themselves, rather than the system designers—the operators of the system do not have to create all this material. [This observation is the first clue to the solution to our problem.]

Our original specification for Habitat called for us to create a world capable of supporting a population of 20,000 Avatars, with expansion plans for up to 50,000. By any reckoning this is a large undertaking and complexity problems would certainly be expected. However, in practice we exceeded the complexity threshold very early in development. By the time the population of our online community had reached around 50 we were in over our heads [and these 50 were “insiders” who were prepared to be tolerant of holes and rough edges].

Moreover, a virtual world such as Habitat needs to scale with its population. For 20,000 Avatars we needed 20,000 “houses”, organized into towns and cities with associated traffic arteries and shopping and recreational areas. We needed wilderness areas between the towns so that everyone would not be jammed together into the same place. Most of all, we needed things for 20,000 people to do. They needed interesting places to visit—and since they can’t all be in the same place at the same time, they needed a lot of interesting places to visit—and things to do in those places. Each of those houses, towns, roads, shops, forests, theaters, arenas, and other places is a distinct entity that someone needs to design and create. We, attempting to play the role of omniscient central planners, were swamped.
Automated tools may be created to aid the generation of areas that naturally possess a high degree of regularity and structure, such as apartment buildings and road networks. We created a number of such tools, whose spiritual descendents will no doubt be found in the standard bag of tricks of future cyberspace architects. However, the very properties which make some parts of the world amenable to such techniques also make those same parts of the world among the least important. It is really not a problem if every apartment building looks pretty much like every other. It is a big problem if every enchanted forest is the same. Places whose value lies in their uniqueness, or at least in their differentiation from the places around them, need to be crafted by hand. This is an incredibly labor intensive and time consuming process. Furthermore, even very imaginative people are limited in the range of variation that they can produce, especially if they are working in a virgin environment uninfluenced by the works and reactions of other designers.

Running the World
The world design problem might still be tractable, however, if all players had the same goals, interests, motivations and types of behavior. Real people, however, are all different. For the designer of an ordinary game or simulation, human diversity is not a major problem, since he or she gets to establish the goals and motivations on the participants’ behalf, and to specify the activities available to them in order to channel events in the preferred direction. Habitat, however, was deliberately open-ended and pluralistic. The idea behind our world was precisely that it did not come with a fixed set of objectives for its inhabitants, but rather provided a broad palette of possible activities from which the players could choose, driven by their own internal inclinations. It was our intent to provide a variety of possible experiences, ranging from events with established rules and goals [a treasure hunt, for example] to activities propelled by the players’ personal motivations [starting a business, running the newspaper] to completely free-form, purely existential activities [hanging out with friends and conversing]. Most activities, however, involved some degree of pre-planning and setup on our part—we were to be like the cruise director on an ocean voyage, but we were still thinking like game designers.

The first goal-directed event planned for Habitat was a rather involved treasure hunt called the "D'nalsi Island Adventure". It took us hours to design, weeks to build [including a 100-region island], and days to coordinate the actors involved. It was designed much like the puzzles in an adventure game. We thought it would occupy our players for days. In fact, the puzzle was solved in about 8 hours by a person who had figured out the critical clue in the first
15 minutes. Many of the players hadn’t even had a chance to get into the game. The result was that one person had had a wonderful experience, dozens of others were left bewildered, and a huge investment in design and setup time had been consumed in an eye blink. We expected that there would be a wide range of “adventuring” skills in the Habitat audience. What wasn’t so obvious until afterward was that this meant that most people didn’t have a very good time, if for no other reason than that they never really got to participate. It would clearly be foolish and impractical for us to do things like this on a regular basis.

Again and again we found that activities based on often unconscious assumptions about player behavior had completely unexpected outcomes (when they were not simply outright failures). It was clear that we were not in control. The more people we involved in something, the less in control we were. We could influence things, we could set up interesting situations, we could provide opportunities for things to happen, but we could not dictate the outcome. Social engineering is, at best, an inexact science [or, as some wag once said, “In the most carefully constructed experiment under the most carefully controlled conditions, the organism will do whatever it damn well pleases”].

Propelled by these experiences, we shifted into a style of operations in which we let the players themselves drive the direction of the design. This proved far more effective. Instead of trying to push the community in the direction we thought it should go, an exercise rather like herding mice, we tried to observe what people were doing and aid them in it. We became facilitators as much as we were designers and implementers. This often meant adding new features and new regions to the system at a frantic pace, but almost all of what we added was used and appreciated, since it was well matched to people’s needs and desires. We, as the experts on how the system worked, could often suggest new activities for people to try or ways of doing things that people might not have thought of. In this way we were able to have considerable influence on the system’s development in spite of the fact that we didn’t really hold the steering wheel—more influence, in fact, than we had had when we were operating under the illusion that we controlled everything.

Indeed, the challenges posed by large systems are prompting some researchers to question the centralized, planning dominated attitude that we have criticized here, and to propose alternative approaches based on evolutionary and market principles. These principles appear applicable to complex systems of all types, not merely those involving interacting human beings.
The Great Debate

Among the objects we made available to Avatars in Habitat were guns and various other sorts of weapons. We included these because we felt that players should be able to materially affect each other in ways that went beyond simply talking, ways that required real moral choices to be made by the participants. We recognized the age old story-teller's dictum that conflict is the essence of drama. Death in Habitat was, of course, not like death in the real world! When an Avatar is killed, he or she is teleported back home, head in hands (literally), pockets empty, and any object in hand at the time dropped on the ground at the scene of the crime. Any possessions carried at the time are lost. It was more like a setback in a game of "Chutes and Ladders" than real mortality. Nevertheless, the death metaphor had a profound effect on people's perceptions. This potential for murder, assault and other mayhem in Habitat was, to put it mildly, controversial. The controversy was further fueled by the potential for lesser crimes. For instance, one Avatar could steal something from another Avatar simply by snatching the object out its owner's hands and running off with it.

We had imposed very few rules on the world at the start. There was much debate among the players as to the form that Habitat society should take. At the core of much of the debate was an unresolved philosophical question: is an Avatar an extension of a human being (thus entitled to be treated as you would treat a real person) or a Pac-Man-like critter destined to die a thousand deaths or something else entirely? Is Habitat murder a crime? Should all weapons be banned? Or is it all "just a game"? To make a point, one of the players took to randomly shooting people as they roamed around. The debate was sufficiently vigorous that we took a systematic poll of the players. The result was ambiguous: 50% said that Habitat murder was a crime and shouldn't be a part of the world, while the other 50% said it was an important part of the fun.

We compromised by changing the system to allow thievery and gunplay only outside the city limits. The wilderness would be wild and dangerous while civilization would be orderly and safe. This did not resolve the debate, however. One of the outstanding proponents of the anti-violence point of view was motivated to open the first Habitat church, the Order of the Holy Walnut (in real life he was a Greek Orthodox priest). His canons forbid his disciples to carry weapons, steal, or participate in violence of any kind. His church became quite popular and he became a very highly respected member of the Habitat community.
Furthermore, while we had made direct theft impossible, one could still engage in indirect theft by stealing things set on the ground momentarily or otherwise left unattended. And the violence still possible in the outlands continued to bother some players. Many people thought that such crimes ought to be prevented or at least punished somehow, but they had no idea how to do so. They were used to a world in which law and justice were always things provided by somebody else. Somebody eventually made the suggestion that there ought to be a Sheriff. We quickly figured out how to create a voting mechanism and rounded up some volunteers to hold an election. A public debate in the town meeting hall was heavily attended, with the three Avatars who had chosen to run making statements and fielding questions. The election was held, and the town of Populopolis acquired a Sheriff.

For weeks the Sheriff was nothing but a figurehead, though he was a respected figure and commanded a certain amount of moral authority. We were stumped about what powers to give him. Should he have the right to shoot anyone anywhere? Give him a more powerful gun? A magic wand to zap people off to jail? What about courts? Laws? Lawyers? Again we surveyed the players, eventually settling on a set of questions that could be answered via a referendum. Unfortunately, we were unable to act on the results before the pilot operations ended and the system was shut down. It was clear, however, that there are two basic camps: anarchy and government. This is an issue that will need to be addressed by future cyberspace architects. However, our view is that a virtual world need not be set up with a "default" government, but can instead evolve one as needed.

**A Warning**

Given the above exhortation that control should be released to the users, we need to inject a note of caution and present our next assertion:

*You can't trust anyone.*

This may seem like a contradiction of much of the preceding, but it really is not. Designers and operators of a cyberspace system must inhabit two levels of virtual world at once. The first we call the "infrastructure level", which is the implementation, where the laws that govern "reality" have their genesis. The second we call the "perceptual level", which is what the users see and experience. It is important that there not be "leakage" between these two levels. The first level defines the physics of the world. If its integrity is breached, the consequences can range from aesthetic unpleasantness (the audience catches a glimpse of
the scaffolding behind the false front) to psychological disruption (somebody does something "impossible", thereby violating users’ expectations and damaging their fantasy) to catastrophic failure (somebody crashes the system). When we exhort you to give control to the users, we mean control at the percipient level. When we say that you can’t trust anyone, we mean that you can’t trust them with access to the infrastructure level. Some stories from Habitat will illustrate this.

When designing a piece of software, you generally assume that it is the sole intermediary between the user and the underlying data being manipulated (possibly multiple applications will work with the same data, but the principle remains the same). In general, the user need not be aware of how data are encoded and structured inside the application. Indeed, the very purpose of a good application is to shield the user from the ugly technical details. It is conceivable that a technically astute person who is willing to invest the time and effort could decipher the internal structure of things, but this would be an unusual thing to do as there is rarely much advantage to be gained. The purpose of the application itself is, after all, to make access to and manipulation of the data easier than digging around at the level of bits and bytes. There are exceptions to this, however. For example, most game programs deliberately impose obstacles on their players in order for play to be challenging. By tinkering around with the insides of such a program—dumping the data files and studying them, disassembling the program itself and possibly modifying it—it may be possible to "cheat." However, this sort of cheating has the flavor of cheating at solitaire: the consequences adhere to the cheater alone. There is a difference, in that disassembling a game program is a puzzle-solving exercise in its own right, whereas cheating at solitaire is pointless, but the satisfactions to be gained from it, if any, are entirely personal.

If, however, a computer game involves multiple players, delving into the program’s internals can enable one to truly cheat, in the sense that one gains an unfair advantage over the other players of which they may be unaware. Habitat is such a multi-player game. When we were designing the software, our “prime directive” was, “The backend shall not assume the validity of anything a player computer tells it.” This is because we needed to protect ourselves against the possibility that a clever user had hacked around with his copy of the front end program to add “custom features.” For example, we could not implement any of the sort of “skill and action” elements found in traditional video games wherein dexterity with the joystick determines the outcome of, say, armed combat, because you couldn’t guard against
someone modifying their copy of the program to tell the backend that they had "hit," whether they actually had or not. Indeed, our partners at QuantumLink warned us of this very eventuality before we even started—they already had users who did this sort of thing with their regular system. Would anyone actually go to the trouble of disassembling and studying 100K or so of incredibly tight and bizarrely threaded 6502 machine code just to tinker? As it turns out, the answer is yes. People have. We were not 100% rigorous in following our own rule. It turned out that there were a few features whose implementation was greatly eased by breaking the rule in situations where, in our judgment, the consequences would not be material if people "cheated" by hacking their own systems. Darned if people didn't hack their systems to cheat in exactly these ways.

Care must be taken in the design of the world as well. One incident that occurred during our pilot test involved a small group of players exploiting a bug in our world database which they interpreted as a feature. First, some background. Avatars are hatched with 2000 Tokens in their bank account, and each day that they login they receive another 100T. Avatars may acquire additional funds by engaging in business, winning contests, finding buried treasure, and so on. They can spend their Tokens on, among other things, various items that are for sale in vending machines called Vendroids. There are also Pawn Machines, which will buy objects back (at a discount, of course).

In order to make this automated economy a little more interesting, each Vendroid had its own prices for the items in it. This was so that we could have local price variation [i.e., a widget would cost a little less if you bought it at Jack's Place instead of The Emporium]. It turned out that in two Vendroids across town from each other were two items for sale whose prices we had inadvertently set lower than what a Pawn Machine would buy them back for: Dolls (for sale at 75T, hock for 100T) and Crystal Balls (for sale at 18,000T, hock at 30,000T!). Naturally, a couple of people discovered this. One night they took all their money, walked to the Doll Vendroid, bought as many Dolls as they could, then took them across town and pawned them. By shuttling back and forth between the Doll Vendroid and the Pawn Shop for hours, they amassed sufficient funds to buy a Crystal Ball, whereupon they continued the process with Crystal Balls and a couple orders of magnitude higher cash flow. The final result was at least three Avatars with hundreds of thousands of Tokens each. We only discovered this the next morning when our daily database status report said that the money supply had quintupled overnight.
We assumed that the precipitous increase in "T1" was due to some sort of bug in the software. We were puzzled that no bug report had been submitted. By poking around a bit we discovered that a few people had suddenly acquired enormous bank balances. We sent Habitat mail to the two richest, inquiring as to where they had gotten all that money overnight. Their reply was, "We got it fair and square! And we're not going to tell you how!" After much abject pleading on our part they eventually did tell us, and we fixed the erroneous pricing. Fortunately, the whole scam turned out well, as the nouveau rich Avatars used their bulging bankrolls to underwrite a series of treasure hunt games which they conducted on their own initiative, much to the enjoyment of many other players on the system.

**Keeping “Reality” Consistent**

The urge to breach the boundary between the infrastructure level and the percipient level is not confined to the players. The system operators are also subject to this temptation, though their motivation is expediency in accomplishing their legitimate purposes rather than the gaining of illegitimate advantage. However, to the degree to which it is possible, we vigorously endorse the following principle:

*Work within the system.*

Wherever possible, things that can be done within the framework of the percipient level should be. The result will be smoother operation and greater harmony among the user community. This admonition applies to both the technical and the sociological aspects of the system.

For example, with the players in control, the *Habitat* world would have grown much larger and more diverse than it did had we ourselves not been a technical bottleneck. All new region generation and feature implementation had to go through us, since there was no means for players to create new parts of the world on their own. Region creation was an esoteric technical specialty, requiring a plethora of obscure tools and a good working knowledge of the treacherous minefield of limitations imposed by the Commodore 64. It also required a lot of behind-the-scenes activity that would probably spoil the illusion for many. One of the goals of a next generation *Habitat*-like system ought to be to permit far greater creative involvement by the participants without requiring them to ascend to full-fledged guru-hood to do so.

A further example of working within the system, this time in a social sense, is illustrated by the following experience. One of the more popular events in *Habitat* took place late in the
test, the brainchild of one of the more active players who had recently become a QuantumLink employee. It was called the “Dungeon of Death”.

For weeks, ads appeared in Habitat’s newspaper, The Rant, announcing that that Duo of Dread, DEATH and THE SHADOW, were challenging all comers to enter their lair. Soon, on the outskirts of town, the entrance to a dungeon appeared. Out front was a sign reading, “Danger! Enter at your own risk!” Two system operators were logged in as DEATH and THE SHADOW, armed with specially concocted guns that could kill in one shot, rather than the usual.12 These two characters roamed the dungeon blasting away at anyone they encountered. They were also equipped with special magic wands that cured any damage done to them by other Avatars, so that they wouldn’t themselves be killed. To make things worse, the place was littered with dead ends, pathological connections between regions, and various other nasty and usually fatal features. It was clear that any explorer had better be prepared to “die” several times before mastering the dungeon. The rewards were pretty good: 1000 Tokens minimum and access to a special Vendroid that sold magic teleportation wands. Furthermore, given clear notice, players took the precaution of emptying their pockets before entering, so that the actual cost of getting “killed” was minimal.

One evening, one of us was given the chance to play the role of DEATH. When we logged in, we found him in one of the dead ends with four other Avatars who were trapped there. We started shooting, as did they. However, the last operator to run DEATH had not bothered to use his special wand to heal any accumulated damage, so the character of DEATH was suddenly and unexpectedly “killed” in the encounter. As we mentioned earlier, when an Avatar is killed, any object in his hands is dropped on the ground. In this case, said object was the special kill-in-one-shot gun, which was immediately picked up by one of the regular players who then made off with it. This gun was not something that regular players were supposed to have. What should we do?

It turned out that this was not the first time this had happened. During the previous night’s mayhem the special gun was similarly absconded with. In this case, the person playing DEATH was one of the regular system operators, who, used to operating the regular Q-Link service, simply ordered the player to give the gun back. The player considered that he had obtained the weapon as part of the normal course of the game and balked at this, whereupon the operator threatened to cancel the player’s account and kick him off the system if he did not
comply. The player gave the gun back, but was quite upset about the whole affair, as were many of his friends and associates on the system. Their world model had been painfully violated.

When it happened to us, we played the whole incident within the role of DEATH. We sent a message to the Avatar who had the gun, threatening to come and kill her if she didn’t give it back. She replied that all she had to do was stay in town and DEATH couldn’t touch her (which was true, if we stayed within the system). OK, we figured, she’s smart. We negotiated a deal whereby DEATH would ransom the gun for 10,000 Tokens. An elaborate arrangement was made to meet in the center of town to make the exchange, with a neutral third Avatar acting as an intermediary to ensure that neither party cheated. Of course, word got around and by the time of the exchange there were numerous spectators. We played the role of DEATH to the hilt, with lots of hokey melodramatic shtick. The event was a sensation. It was written up in the newspaper the next morning and was the talk of the town for days. The Avatar involved was left with a wonderful story about having cheated DEATH, we got the gun back, and everybody went away happy.

These two very different responses to an ordinary operational problem illustrate our point. Operating within the participants’ world model produced a very satisfactory result. On the other hand, what seemed like the expedient course, which involved violating this model, provoked upset and dismay. Working within the system was clearly the preferred course in this case.

**Current Status**

As of this writing, the North American incarnation of Lucasfilm’s *Habitat*, QuantumLink’s “Club Caribe,” has been operating for almost two years. It uses our original Commodore 64 front end and a somewhat stripped-down version of our original Stratus backend software. Club Caribe now sustains a population of some 15,000 participants.

A technically more advanced version, called *Fujitsu Habitat*, has recently started pilot operations in Japan, available on Nifty Serve. The initial front end for this version is the new Fujitsu FM Towns personal computer, though ports to several other popular Japanese machines are anticipated. This version of the system benefits from the additional computational power and graphics capabilities of a newer platform, as well as the Towns’ built-in CD-ROM for object imagery and sounds. However, the virtuality of the system is essentially unchanged and Fujitsu has not made significant alterations to the user interface or to any of the underlying concepts.
Future Directions

There are several directions in which this work can be extended. Most obvious is to implement the system on more advanced hardware, enabling a more sophisticated display. A number of extensions to the user interface also suggest themselves. However, the line of development most interesting to us is to expand on the idea of making the development and expansion of the world itself part of the users' sphere of control. There are two major research areas in this. Unfortunately, we can only touch on them briefly here.

The first area to investigate involves the elimination of the centralized backend. The backend is a communications and processing bottleneck that will not withstand growth above too large a size. While we can support tens of thousands of users with this model, it is not really feasible to support millions. Making the system fully distributed, however, requires solving a number of difficult problems. The most significant of these is the prevention of cheating. Obviously, the owner of the network node that implements some part of the world has an incentive to tilt things in his favor there. We think that this problem can be addressed by secure operating system technologies based on public-key cryptographic techniques.19,20

The second fertile area of investigation involves user configuration of the world itself. This requires finding ways to represent the design and creation of regions and objects as part of the underlying fantasy. Doing this will require changes to our conception of the world. In particular, we don't think it will be possible to conceal all of the underpinnings to those who work with them. However, all we really need to do is find abstractions for those underpinnings that fit into the fantasy itself. Though challenging, this is, in our opinion, eminently feasible.

Conclusions

We feel that the defining characteristic of cyberspace is the shared virtual environment, not the display technology used to transport users into that environment. Such a cyberspace is feasible today, if you can live without head-mounted displays and other expensive graphics hardware. Habitat serves as an existence proof of this contention.

It seems clear to us that an object-oriented world model is a key ingredient in any cyberspace implementation. We feel we have gained some insight into the data representation and communications needs of such a system. While we think that it may be premature to start establishing detailed technical standards for these things, it is time to begin the discussions that will lead to such standards in the future.
Finally, we have come to believe that the most significant challenge for cyberspace developers is to come to grips with the problems of world creation and management. While we have only made the first inroads onto these problems, a few things have become clear. The most important of these is that managing a cyberspace world is not like managing the world inside a single-user application or even a conventional online service. Instead, it is more like governing an actual nation. Cyberspace architects will benefit from study of the principles of sociology and economics as much as from the principles of computer science. We advocate an agoric, evolutionary approach to worldbuilding rather than a centralized, socialistic one.

We would like to conclude with a final admonition, one that we hope will not be seen as overly contentious:

Get real.

In a discussion of cyberspace on Usenet, one worker in the field dismissed Club Caribe (Habitat’s current incarnation) as uninteresting, with a comment to the effect that most of the activity consisted of inane and trivial conversation. Indeed, the observation was largely correct. However, we hope some of the anecdotes recounted above will give some indication that more is going on than those inane and trivial conversations might indicate. Further, to dismiss the system on this basis is to dismiss the users themselves. They are paying money for this service. They don’t view what they do as inane and trivial, or they wouldn’t do it. To insist this presumes that one knows better than they what they should be doing. Such presumption is another manifestation of the omniscient central planner who dictates all that happens, a role that this entire article is trying to deflect you from seeking. In a real system that is going to be used by real people, it is a mistake to assume that the users will all undertake the sorts of noble and sublime activities which you created the system to enable. Most of them will not. Cyberspace may indeed change humanity, but only if it begins with humanity as it really is.

Notes

Note 1: One of the questions we are asked most frequently is, “Why the Commodore 64?” Many people somehow get the impression that this was a technical decision, but the real explanation has to do with business, not technology. Habitat was initially developed by Lucasfilm as a commercial product for QuantumLink, an online service (then) exclusively for owners of the Commodore 64. At the time we started (1985), the Commodore 64 was the mainstay of the recreational computing market. Since then it has declined dramatically in both its commercial and technical significance. However, when we began the
project, we didn't get a choice of platforms. The nature of the deal was such that both the Commodore 64 for the frontend and the existing QuantumLink host system (a brace of Stratus fault-tolerant minicomputers) for the backend were given.

Note 2: Habitat contains its own fully-fledged economy, with money, banks, and so on. Habitat's unit of currency is the Token, reflecting the fact that it is a token economy and to acknowledge the long and honorable association between tokens and video games. Incidentally, the Habitat Token is a 23-sided plastic coin slightly larger than an American quarter, with a portrait of Vernor Vinge and the motto "Fiat Lucre" on its face, and the text "Good for one fare" on the back; these details are difficult to make out on the Commodore 64 screen.

Acknowledgements
We would like to acknowledge the contributions of some of the many people who helped make Habitat possible. At Lucasfilm, Aric Wilmunder wrote much of the Commodore 64 frontend software; Ron Gilbert, Charlie Kelner, and Noah Falstein also provided invaluable programming and design support; Gary Winnick and Ken Macklin were responsible for all the artwork; Chris Grigg did the sounds; Steve Arnold provided outstanding management support; and George Lucas gave us the freedom to undertake a project that for all he knew was both impossible and insane. At Quantum, Janet Hunter wrote the guts of the backend; Ken Huntsman and Mike Ficco provided valuable assistance with communications protocols. Kazuo Fukuda and his crew at Fujitsu have carried our vision of Habitat to Japan and made it their own. Phil Salin, our boss at AMIX, let us steal the time to write this paper and even paid for us to attend the First Conference on Cyberspace, even though its immediate relevance to our present business may have seemed a bit obscure at the time. We'd also like to thank Michael Benedikt, Don Fussell and their cohorts for organizing the Conference and thereby prompting us to start putting our thoughts and experiences in writing.

References


Goldberg, Adele, and Robson, David, Smalltalk-80: The Language and Its Implementation [Addison-Wesley, Reading, Mass, 1983].


Unwritten Rules

Stephen Sniderman

Context
"Unwritten Rules" is the feature article in the first issue of The Life of Games (October, 1999), the online journal at www.gamepuzzles.com/tlog/tlog.htm, which I co-edit with Kate Jones, founder of Kadon Gamepuzzles. This article is an attempt to clear the ground for discussing games by challenging the widely held assumption that a game is fundamentally different from other human activities (such as the law or business) because we can know all its rules. Games are thus invoked as a model of a fully describable closed system, but I try to show that a game played by humans cannot be a closed system and therefore cannot be fully described.

The Rules of a Game
Gaming the Game
Speaking of Games

Stephen Sniderman has been teaching American literature and creative writing at Youngstown State University since 1969. He has published two books (Language Lovers' Word Puzzles; Stanley Newman Presents Grid Play), a game system (Flying Colors) with Kadon Gamepuzzles, and dozens of puzzles and games in GAMES Magazine and English Journal.
Regardless of what game you’re playing, you cannot know all the rules.
Whether the "game" is tic-tac-toe, chess, baseball, language, etiquette, education, science, religion, law, business, politics or war, the entire set of rules governing the system cannot be spelled out. No matter how hard we try to indicate what is required, allowed, and proscribed, some of the most fundamental principles of playing the game will always elude us. And yet, paradoxically, we can still play the game—some deeper rules are always operating [i.e., affecting the players’ behavior] without the players’ being aware of them.

What do we mean by a game?
A game is a play activity that consists of an object [a goal or goals that the players are trying to accomplish] and constraints on the players’ behavior [what they must do and/or what they may not do in attempting to achieve the game’s object]. To play a game is to pursue that game’s object while adhering [more or less] to its constraints. Some of these constraints [the “recorded rules”] are explicitly spelled out and are what we generally understand to be “the rules of the game,” but every game is also governed by constraints that are rarely if ever made explicit. Some of these “unrecorded rules” are literally unstateable.

An example with tic-tac-toe
Suppose I challenge you to a game of tic-tac-toe. Could anything be more straightforward? But just to be sure, we review the rules. We’ll play on a 3x3 grid, we’ll alternate turns, we’ll play only in empty squares, I’ll play X’s, you play O’s, I’ll play first, and the first player to get three of his/her symbol in a row, column, or diagonal wins the game. Aren’t these all the rules of tic-tac-toe?

Well, for one thing, nothing has been said about time. Is there a time limit between moves? Normally, we both “understand” that there is, and we both “know” that our moves should be made within a “reasonable” time, say 20 seconds. If one of us takes longer, the other starts to fidget or act bored, may even make not-so-subtle comments, and eventually threaten to quit. Without having stated it, we have accepted a tacit time limit. And because we haven’t stated it, it is fairly flexible and very functional.

Is it a rule, or isn’t it?
Suppose it is my turn and, no matter what I do, you will win on your next move. Couldn’t I prevent that from happening, within the rules stated, by simply refusing to play? Nothing in the
rules forces me to move within a particular amount of time, so I simply do not make my next move. Haven’t I followed the rules and avoided losing? And yet, if you’ve ever played a game, you know that this strategy is almost never employed and would be completely unacceptable. Anybody who seriously resorted to such a tactic would be considered childish, unsportsmanlike, or socially undesirable and would probably not be asked to play in the future. This behavior seems to violate some fundamental but rarely stated principle of the game without any of us ever having to discuss it.

**Self-defeating rules**

But can’t we state the principle it violates? Can’t we just make that principle one of the rules of tic-tac-toe and other games? The answer is—yes, of course we can, but we will not eliminate the problem. Suppose we add the following rule: *Players will make their moves within a reasonable amount of time.* Have we solved anything? What is a “reasonable” amount of time? One minute? Five? 30? A million? And who determines what is reasonable—the player whose turn it is or the other player?

Such a rule is actually self-defeating because it calls attention to the fact that we cannot spell out what “reasonable” means.

So why not just specify a time limit for each move? Because we would just create even more perplexing problems for ourselves. For one thing, we would have to indicate when a player’s time is running and when it is not. If one player had to answer the phone, for example, would we count that time or wouldn’t we? To state the rule fully, we would have to list every life situation that could possibly interrupt a player’s turn and state whether it should count against that player’s time limit. Obviously, we could never complete such a list.

**Practical solutions**

A far more practical “solution,” the one most of us have used all our lives in “friendly” games, is to say nothing about time limits and rely on our opponent’s intuitive understanding of a “reasonable” time for a move and his/her desire to keep the game moving and therefore enjoyable. In other words, we depend on unstated—and probably unstateable—“rules” (really just expectations) when we play a game for fun.

In tournament or professional games, of course, we cannot leave things so loose, and various methods have been employed to solve the time dilemma. Generally, specific time limits are spelled out, as are specific penalties, including forfeit, for exceeding them. Official
devices are employed for timing moves—chess clocks, the shot-clock in college and pro basketball, stopwatches in baseball games and tennis tournaments, and so on. But once we move beyond “reasonable” to “official” time, we create a whole new set of problems, problems that can no longer be solved with a simple agreement between or among players.

**Rulings versus rules**

As any sports fan knows, the difficulties that arise with “official” rules and “officials” to interpret them are often more intractable than those we face in friendly games. Since no set of rules can list every possible situation that might come up during tournament play, someone in charge, rather than the players themselves, must decide if a player has violated a rule (such as exceeding a time limit) and what penalty should be invoked.

Suppose, for instance, that a fire alarm sounds during a chess tournament and players are forced to evacuate the room. Someone in charge of the tournament must determine whether or not the time spent out of the room should be counted against the players whose clocks were running. It is doubtful that the tournament rules will help them. Or suppose the shot clock in an NBA game stops functioning temporarily. When it is fixed, the officials must decide how much time to put on the clock. How could any rule specify the precise amount of time that would be appropriate? Or suppose a professional tennis player complains of cramps. A human being, not a rulebook, must determine whether the player’s complaint is legitimate and decide whether to grant the player additional time to recover.

Presumably, the officials’ decisions in these situations would be based on the notions of fairness, sportsmanship, and practicality, notions that have never been—and almost certainly cannot be—fully codified and agreed upon. Therefore, no matter how exhaustive and specific we try to make the rules about time limits (or anything else) in a game, we will always have to rely on other people’s acceptance of a set of principles that neither they nor we can put into words. That’s the nature of any human system—the most important aspects of it are unstatable and unknowable.

In *The Celebrant*, Eric Rolfe Greenberg cogently illustrates this little-recognized truth. He depicts the famous incident in baseball lore that got Bonehead Merkle his nickname. With two outs in the bottom of the ninth inning and the score tied, Merkle is on first and a teammate is on third. The next batter hits the ball cleanly into right field for a single, which drives in the apparent game-winning run. Fans pour out onto the field in celebration. Merkle, afraid for his safety, heads directly to the dugout without touching second base. The fielding
team calls for an appeal play at second and attempts to retrieve the ball and touch second for
the third out, ending the inning and negating the tie-breaking run.

But where is the ball? No one is sure because the field is swarming with fans. Nev-
evertheless, one of the fielders, holding a ball, touches second base and claims that Merkle has
been forced out and that the game is still tied. The question arises, is the ball he retrieved the
one that was actually hit? By this time, the umpire has left the field and must be summoned
from his dressing room, which he adamantly refuses to leave—until his life is threatened.
When he does finally stick his head out, he refuses to change his ruling. Naturally, the losing
team appeals to the commissioner of baseball to settle the matter. This worthy stalls as long
as he can and finally declares the game null and void and orders that it be replayed.

Greenberg makes it clear that the commissioner’s decision is influenced by political
and social considerations that have little to do with any rulebook. The game of baseball has
spilled over into real life and the depth of the “rules” governing the sport can be glimpsed.

No game is an island

As this example reminds us, no game or sport is played in a vacuum. All play activities exist
in a “real-world” context, so to play the game is to immerse yourself in that context, whether
you want to or not. In fact, it is impossible to determine where the “game” ends and “real life”
begins. As a result, knowing only the recorded rules of a game is never enough to allow you
to play the game.

Think of the constraints that do not ordinarily get included as part of the recorded
rules of tic-tac-toe but which nevertheless influence the behavior of almost all players. Some
of these involve the conventions, “etiquette,” or “ethos” of this particular game and may vary
from region to region or even family to family.

For example, I would guess that few tic-tac-toe players talk trash to each other [an
acceptable and even expected behavior in some games and sports, like basketball].

Similarly, I’m willing to bet that few people play tic-tac-toe for money [in contrast to
Poker] or prizes [as is sometimes true with Scrabble] or masters points [as with Tournament
Bridge] or glory [as in Central Park chess].

Also, most people, I suspect, would probably allow their opponent, especially an inex-
perienced player or a young child, the opportunity to “take back” an obviously unwise move.
Playing fair
Other unwritten rules are associated with being “a good sport” and would apply to virtually all games in our culture. For example, you may not attempt to coerce your opponent, through physical force or threats or bribery or blackmail, into putting a symbol on a particular square. You may not attempt to cause your opponent physical, mental, or emotional harm to keep him or her from competing effectively. You may not attempt to distract your opponent while he or she is contemplating the next move. On the other hand, you must make your moves in a “reasonable” time. You must take the game seriously and attempt to win. You must play “fair” at all times.

To understand the difficulty—or, more accurately, the impossibility—of spelling out every rule governing the behavior of tic-tac-toe players, try to imagine programming a computer to “understand” what is meant by the sentences in the previous paragraph. For instance, think about the notion of “distracting” an opponent. What counts and what doesn’t? Suppose you are chewing gum or smoking or wearing perfume and your opponent claims to be bothered by the sounds or aromas you are producing. What would we tell Deep Blue about this situation? Can we really list every behavior that qualifies as distracting?

The human factor
Or for that matter, can we ever be sure (in the sense that we could program a computer to determine) that a player is “really” distracted? In his famous match with Boris Spassky in Reykjavik, Iceland, in 1972, Bobby Fischer claimed to be “distracted” by negative vibes that were emanating from his opponent’s camp. Officials could hardly appeal to the recorded rules, as “complete” as those might have been, to determine how to handle Fischer’s complaints. They had to use their experience with people, including Fischer and Spassky, their understanding of human psychology, their awareness of the political and social implications of the situation, and their diplomatic skills to arrive at a satisfactory compromise. Which of these notions is programmable?

Even Deep Blue, the most sophisticated chess program ever devised, cannot distinguish between a game played for blood (or money) and one played for fun; cannot recognize when a move should count and when politeness or common sense or common courtesy or compassion or medical emergency dictates that it shouldn’t; cannot take into account the emotional needs of its opponent; cannot know when it’s appropriate to abandon the game
or suspend play; cannot, in short, understand the social, political, moral, psychological, and philosophical context in which the game occurs.

**Unspoken basics**

Obviously, our ability to participate in a particular game is dependent on our knowledge of many “rules” which no one has ever spelled out to us. Yet it is easy to overlook this simple fact. In *When Elephants Weep*, the authors tell about a group of scientists who attempted to teach dolphins to play water polo. Although the dolphins were able to learn how to put the ball in the net (and seemed to derive pleasure from doing so) when the trainers tried to get them to stop the other team from “scoring,” the dolphins launched an all-out war on the other team’s players, using methods that no person steeped in the concepts of sportspeopleship would ever use.

After this experience, the trainers gave up their effort, apparently concluding that their task was hopeless, that dolphins couldn’t be taught to play the sport. My guess is that they assumed that all the dolphins needed to be taught were the recorded rules of water polo and the creatures would be able to play the game like adult human beings. These scientists evidently did not realize how much of our knowledge of proper game behavior precedes the learning of the statable constraints of a particular sport.

But suppose these trainers had recognized, after their initial failure, that they had to provide their trainees with some more fundamental “rules” of game playing. Would they ever have been able to teach dolphins all they need to know to play a single “human” game? Are dolphins capable of understanding fairness and sportscreatureship, “time in” vs. “time out,” practice vs. competition, winning and losing? And even if they were, how would we go about teaching these concepts to them? Wouldn’t we have to teach them much of our culture in order for them to play the game as we do?

**Sportsmanship 101**

To grasp the immensity of the trainers’ task, let us look more closely at what we must know and do to play the simplest game in our culture. We must:

1. intuitively understand what is meant by *play* in our culture, recognize how it differs from other activities, and be able to tell when someone is involved in the behaviors associated with play in general and games in particular;
2. intuitively understand what game/sport is being played, which behaviors constitute part of that activity and which do not, when the activity is underway, when it is in suspension, and when it is concluded;

3. consciously understand and pursue the object(s) of the game (i.e., what we must accomplish to be "successful");

4. consciously understand and follow all [or at least a large majority of] the defining prescriptions and proscriptions of the game, the "written," statable rules—i.e., what we must and must not do in the course of pursuing the object or objects;

5. consciously understand and follow the etiquette of the game—i.e., the unwritten but sometimes stated traditions associated with the game that do not necessarily affect the play itself (e.g., appropriateness of talking, gloating, taunting, celebrating, stalling, replaying a point, giving advice to your opponent or teammates, letting players take back moves, etc.);

6. intuitively understand and follow the ethos of that particular game—i.e., the unwritten and rarely expressed assumptions about how to interpret and enforce the "written" rules (e.g., palming in basketball; the strike zone in American and National Leagues; the foot-fault in tennis);

7. intuitively understand and follow the conventions of playing any game according to the culture of the participants—i.e., the unwritten and generally unstatable customs related to playing, competing, winning/losing, etc. (e.g., taking the game with the appropriate seriousness, knowing what takes priority over winning and over playing, not taking injury or personal obligation to avoid losing; playing "hard" regardless of the score; not claiming that previous points didn’t "count");

8. intuitively understand and respond to the "real-life" context in which the game is being played—i.e. the social, cultural, economic, political, and moral consequences of the result (e.g., whether someone’s livelihood or self-esteem depends on the outcome).

**Going through the motions**

Obviously, we are never merely playing a game. Or, to say it another way, we are never playing only one game. We are always conscious of the game’s relation to the world in which we live, the world in which that game is one small part.
How much of this context could a non-human “understand?” Is a racehorse “playing
the game” of horse racing or merely responding to the urgings of the jockey? Is Deep Blue
“playing” chess or merely making moves on a chessboard according to a particular algorithm?
Is either trying to win?

If not, they are not playing the game in any meaningful sense. As I see it, to perform
the skills and behaviors associated with the game without consciously pursuing the object(s)
of the game is not equivalent to playing the game. We might be practicing the game, pre-
tending to play (as with pro wrestlers or actors in a movie about a sport), or exercising our
muscles, but there is no game without the attempt, on the part of at least one of the players,
to achieve the stable object of that game. [Could dolphins ever be taught to pursue such an
object, or would they merely go through the motions of play? And how would we know?]

In addition, it is not possible to pursue the object of the game independent of the key
prescriptions and proscriptions. Built into the object(s) of any game is the manner in which
it/they must and must not be pursued.

The primary object of a football game, for example, is not to cross your opponents’
goal line while carrying a football; it is to score a touchdown. An equipment manager carrying
a bag of footballs through the end zone of a football field has not scored a touchdown. These
are profoundly different events, and perceiving the difference between them is a key to un-
derstanding the game. Thus, not understanding the difference between them is tantamount
to not understanding the game of football. Could any non-human ever make this distinction?

“Time in”
Perhaps the single most important “rules” that are literally unstatable, then, are those that
define the context of the game and answer the question, “When is the game being played?”
None of us can say how we know that we are in fact playing a particular game (rather than,
say, just practicing), but we generally have no trouble knowing that we are. That suggests that
there are many subtle cues we give and receive about what play activity we are engaged in,
what “counts,” when time is “in,” when the game has started, when play is suspended, and
when the game has ended.

Let me offer a personal example. When my buddies and I play tennis, we meet each
other at the court at a prearranged time, take out our tennis racquets and some balls, warm
up for 15-20 minutes (hitting ground strokes, volleys, overheads, and serves), and eventually
someone asks, “Ready?” or perhaps “Ready to play?” If anyone says no, we continue to warm up. If everybody says Yes (or nobody says No), we toss away all but three balls. At this point, I (and presumably the others) understand that the actual game is going to begin with the next serve. There is never a formal announcement that play is about to begin. At most, the server will hold up a ball and the others will nod or wave.

None of us has ever acknowledged that this is our practice, none of us has stated any of these behaviors, as “rules,” none of us would be able to say how we arrived at these customs, yet none of us, I assume, would have any doubt when the game has started.

Could I program a computer or teach a dolphin to operate with the same certainty? Could I specify all the variations in our ritual so that non-humans (or non-sports fans) could identify the boundary line between warm-up and play?

**On your mark... get set...**

Players, fans, and officials of any game or sport develop an acute awareness of the game’s “frame” or context, but we would be hard pressed to explain in writing, even after careful thought, exactly what the signs are. After all, even an umpire’s yelling of “Play Ball” is not the exact moment the game starts. (And think how confused a new fan of baseball would be when some dignitary threw out “the first pitch”!) We must rely on our intuition, based on our experience with a particular culture, to recognize when a game has begun.

We cannot, in other words, program a computer to understand all the conditions that must be satisfied for humans in a particular culture to say that a game is underway. If the computer is turned on and the software for that game booted up, the computer is, by necessity, playing the game, even if its “opponent” is a two-year-old, a monkey, or an accidental jiggling of the keyboard.

In addition, the computer will go on “playing” until it is turned off, even if its opponent moves on to other activities or drops dead. This phenomenon is the premise of the movie, Wargames, in which a supercomputer, WOPR, cannot distinguish between a “game” of Therm-onuclear War and the real thing. When told it is involved in an actual battle, not a simulation, WOPR’s reply is, “What’s the difference?”

By contrast, a human being is constantly noticing if the conditions for playing the game are still being met, continuously monitoring the “frame,” the circumstances surrounding play, to determine that the game is still in progress, always aware (if only unconsciously) that the other participants are acting as if the game is “on.”
For example, in our tennis game, a player will occasionally say, after failing to return a serve, "I wasn’t ready." If the others decide that the player is serious in that announcement, the point is usually replayed. How we determine whether or not the player is joking is beyond my understanding (although I'm perfectly capable of making such a determination) and certainly not in my power to express in words.

"Time out"

But there are other reasons, still more difficult to explain, why a particular serve in our game does not "count," i.e., is (usually) replayed. If the players on the receiving team decide that the server's concentration has been "unfairly" disrupted after serving a fault (because, for example, someone from another court has asked us to retrieve their ball or something else has caused "too much" time to elapse), they generally tell the server to "take two," that is, to try his/her first serve again. In effect, they have made a ruling that the server has been inappropriately distracted between the first and second serve and "deserves" a second chance at two serves for that point.

But what exactly is an "unfair" disruption of play according to the etiquette of our game? Can any one of us spell out precisely what situations warrant a second chance and which do not? After all, we are making no effort here to follow the practices of some official tennis game, so we have no rulebook to appeal to, even if we wished to. (Actually, I would feel silly consulting one for such a petty matter.)

I assume that we are all just following a tradition of hackers' tennis that has been passed down over the generations, almost always by imitation rather than by any explicit explanation.

I also assume that our behavior is based on our own notions of "fairness," not on something we could explain in detail.

As a result, I'm not even certain that the other players in my game have the same reasons for telling someone to "take two" as I do, but I have noticed a reasonable consistency over the years.

Occasionally, we facetiously [I assume] debate about whether we should give the opposing player another first serve, but our discussion itself is usually seen as a sufficient distraction to settle the matter in the server's favor. Incidentally, I have never heard the server request a second chance, except in jest [I have assumed], regardless of the circumstances, and some servers will not accept the receiving team's ruling unless it is insisted upon.
Below the surface: Who's the best sport?
A kind of sub-game is going on “underneath” the more obvious one called tennis. Many hackers, myself included, try to one-up each other in politeness and thoughtfulness, so this aspect of our tennis matches can be thought of as a kind of game-within-the-game in which the object is to come off as the best sport.

Of course, no one ever acknowledges this game and no winner is ever announced. My guess is that this practice gives us hackers a chance to feel successful on some level, regardless of the outcome of the match.

Keep in mind that I have never discussed any of these customs with my tennis buddies and probably never will, but I can say that almost every hacker I’ve ever played tennis with (including those who are fierce competitors and those who are impolite and inconsiderate in other ways) has practiced this non-professional courtesy, and I’m confident that if I played in a friendly game in Oklahoma or Maine or Florida or Arizona, I would see this same tradition being followed.

Yet what chance does a computer, a dolphin, a non-native speaker, or even a non-player have of understanding this game of “Who’s the best sport?” It’s the kind of thing you have to learn from experience, observation, and inference, not from a set of statable rules.

How can you tell?
Distinguishing between counting and not counting, between “time in” and “time out,” is probably the single most basic skill a game player, fan, or official must possess. Without it, a participant or observer could not tell the difference between the preliminaries (such as a warm-up), the breaks in the action (such as a time-out), the aftermath (such as a handshake or a victory lap) and the game itself, could not know when to expend energy and when to relax, could not keep score accurately, could not determine what behavior was affecting the outcome, and so on.

Obviously, we learn to make these distinctions, but we learn them without being aware, for the most part, that we are learning anything. As a result, the process by which we decide that a game is being played is generally hidden from us and therefore seems perfectly natural, not something that has to be learned.

We forget that children, people from other cultures, and adults in our own culture who are unfamiliar with the game cannot automatically tell which actions are part of the game and which are not.
But even if someone understands the notions of play (#1 in our list above), recognizes when a particular game/sport is being played (#2) and is familiar with its object and "official" (written) rules (#3 and #4), such a person would have difficulty participating in the game/sport at any level without a great deal of additional information (or "rules") about the activity.

The outsider
To illustrate this notion, let us imagine a person named Leslie who has taken extensive tennis lessons, memorized an official USTA rule book, and watched professional tennis on television but never actually played a match at any level and never played or watched or read about any other games (which presumably share some of the unstated rules of tennis).

One day, let's suppose further, someone invites Leslie to substitute in one of our doubles games. Even assuming his skills were similar to ours, I would venture to say that Leslie would not have much fun and would make the rest of us very unhappy. He would almost certainly get very confused and frustrated at the way my friends and I play "tennis."

In fact, Leslie might not even recognize it as tennis at all and might conclude that we are playing some bastardized form of the game.

And in a sense, he would be absolutely right.

By the book
For one thing, as Leslie would be dismayed to discover, none of our rules are "official," in the sense that they are written down or formally agreed upon.

We all seem to assume that we are following the most important rules of professional tennis, except where that is not possible. So, for example, when the ATP adopted a tiebreaker rule for deciding a set, most (but not all) of the games I was involved with also adopted that practice.

In general, the only rules we discuss are those we are uncertain about, such as whether it is legal to touch the net during a point or hit the ball before it crosses the net. Otherwise, we have never spelled out the "rules" we are using, have never stated which set of "official" laws we will abide by, have never established an authority to settle disputes, and have never ever consulted a rule book (at least not at the court) to determine the "correct" way to play. When we disagree about the rules, which rarely happens, we use our knowledge of pro tennis to defend our position.
Not by the book
But we certainly don’t do everything as they do on the ATP tour. As I have already indicated, we give people a second chance at a first serve according to our own lights, not what we see happening at Wimbledon.

To save money, we do not open a new can of balls every seven games, and when we play indoors [where we have to pay for court time], we switch ends of the court after each set, not after every odd game.

In addition, we never assess penalty points for swearing, racquet abuse, exceeding time limits, or foot-faulting. We might grumble about these violations, especially if we think a player is getting an unfair advantage, but we tolerate them, apparently because we perceive them as too trivial to worry about.

Yet some of the people I play with are fanatic about the height of the net. They use a tape measure to make sure the middle of the net is exactly 36 inches high and raise or lower it as needed. They even bring “doubles sticks” to raise the net to the appropriate height at the sides. Wouldn’t our “inconsistency” drive Leslie crazy?

Obviously, one of the most crucial [and rarely stated] meta-rules of games that someone like Leslie (or a computer or a dolphin) would not understand is that we can play them any way we wish, as long as we have [apparent] agreement among the participants. If we want to play tennis with a racquetball or without a net, what’s to stop us?

Tradition-bound
And yet, in my experience, few people choose to play games or sports in innovative ways. Although they are willing to eliminate “trivial” or inessential rules, most people evidently want to feel connected to the tradition of “real” games (i.e., professional sports), even when the rules of the pro version are inappropriate for the local circumstances.

So, for instance, almost all junior high school basketball hoops are 10 feet high, just as they are for the Chicago Bulls, even though the kids are two or three feet shorter than players in the NBA. I guess we like to create the illusion for ourselves that these youngsters are playing the same game as Michael Jordan.

House rules
Even if Leslie finally figured out exactly how our “rules” differed from the ATP’s, he would undoubtedly still be very uncomfortable in our doubles game. For one thing, we play a relatively “casual” game.
We often talk to each other between points, jokingly insult one another, compliment a particularly good shot, ask what the score is, predict what is going to happen next, and so on. Between games, we might exchange personal information or tell jokes.

None of this, of course, happens in professional level tennis, at least not the matches shown on television.

My guess is that Leslie would be disconcerted by our apparent lack of decorum. He would probably perceive us as being remarkably uninterested in the outcome of the game, when in fact we play to win almost as “seriously” as the pros. If he was used to silence between points and games, his concentration might be seriously upset.

**Banter protocols**

Perhaps he would eventually be able to shrug off our casualness as a trivial idiosyncrasy that doesn’t affect the game in any significant way, but it is doubtful that he would be able to participate in the banter. In that case, our “rules” would accommodate his silence. No one is required by our etiquette to talk if s/he doesn’t want to, although we (at least I) tend to prefer those with “personality.” The game is just not as much fun (for me) with duds or robots.

If Leslie did start to talk, though, he might find himself violating other aspects of our etiquette. Certain subjects are taboo, or at least frowned upon or rarely mentioned. Business, for example, is almost never discussed between points and rarely between games. (Perhaps this is merely because the people I play with don’t share work experience.)

More significantly, politics and religion are strictly avoided. At most, someone will make a passing comment about the president or some interesting current event, but I can’t remember a single remark about abortion or gun control or any other such controversial topic, even when I have played with other academics. It’s as if we do not want to acknowledge that we might have serious disagreements outside the tennis court.

Would Leslie recognize that we are limiting our comments to certain topics? Until I wrote these last sentences, I had never articulated this “rule” even to myself (though I’ve been playing for over 40 years).

**Our own language**

Leslie would almost certainly have more difficulty getting used to our line-calling practices. Since we don’t have officials, we (like most hackers, I assume) have devised a fairly elaborate system for deciding if a ball is in or out.
Keep in mind that we have never discussed this system, never written it down, never spelled it out in any way, yet our entire game depends on each player’s following a fairly rigid, if unstatable, set of behaviors. (I’m willing to bet that is generally the case with most amateurs, including those in tournaments, which rarely have official line-callers.)

First, we sometimes use hand signals to indicate “in” (a palm down) or “out” (a finger point), and sometimes, when we think the call is obvious, we say nothing at all. As far as I can tell, we use hand signals only when the ball is not returnable and say “out” when a player has hit the ball back and we wish to indicate that the point is over.

Second, we have a set of “rules” governing which player makes which call. Generally, players on the team about to hit the ball are expected to call the lines, even if a player from the opposing team is closer to the ball when it hits near the line. For example, on a serve, the partner of the player receiving the ball is supposed to announce an out ball.

Of course, there are exceptions (which I can only hint at). Sometimes, for example, the player that hit the ball (or his/her partner) has an unobstructed view of the situation and makes the call. Sometimes, more than one player makes the call. Occasionally two players disagree and a discussion ensues.

To settle a disputed line call, some players like to look for the impression (called a “spot”) the ball has left on the playing surface. If they cannot find a spot, they generally assume the ball hit the line (and the point is awarded to the hitter).

**Fuzzy boundaries**

For the most part, in keeping with the game of “Who’s the best sport?”, players try to appear calm, rational, polite, and objective about line-calling, but occasionally someone will get upset over another’s call, and a new game, whose rules are even harder to describe, breaks out. In this game (“I’m Right and You’re Wrong”), the object is to get the other player to back down and agree with your perception.

What players under these circumstances are allowed or not allowed to say depends partly on the social rules that are in operation—the power relations among the players off the court—so once again we see the fuzziness of the boundary between game and non-game.

In most cases, the desire to continue play or to win the sportsmanship game ends an argument fairly quickly (but I remember once when a player and his grandson argued for over 15 minutes about a particular line call). Usually, when an impasse is reached, players will agree to take the point over.
As should be clear by now, I would never get all our practices down on paper, no matter how long I stayed at it. In fact, I haven’t even finished explaining our system for calling lines, or the “rules” related to the length of time it’s appropriate to debate a particular line call.

In addition, in my attempt to codify our game for “outsiders” (those who have never seen us, or other hackers, play), I have found myself distorting the reality for the sake of convenience. In many cases, I ignored what I knew to be clear exceptions to avoid getting bogged down in impossibly complicated nuances that I’m only dimly aware of.

For instance, one friend, John, and I always discussed controversial issues when we played singles but never when we played doubles! I also ignored the fact that the various groups I play tennis with do not play by identical rules (e.g., normally we spin a racket to determine which team serves first, but when we play at Nazim’s house, the player who opens a can of balls serves the first game); only hinted at the effect a change in circumstances (outdoor vs. indoor, free vs. fee) can have on our game; and oversimplified the modifications in our game over the years.

Thus, as I’ve tried to show, the “casual” game of tennis that my buddies and I play is really based on an enormously complex set of “rules”—assumptions, traditions, and conventions—that govern our behavior on the court (whether we are consciously aware of it or not).

My contention is that no one could ever “fully” describe those rules or those governing the players of any other game.

The infinite-regress trap
It is time to see exactly why a complete listing of a game’s rules is impossible. There are several reasons:

1. Game rules, like any rules, must be stated in some language, and all language is subject to interpretation. But the rules for interpreting any language would also have to be stated in some language, and these rules would likewise have to be interpreted. We are trapped in an infinite regress. Thus, the question “What are the rules?” can never be answered fully.

2. Each individual player could have his or her own personal conception of a game which would differ (if only slightly) from all other players’ versions, and each player’s understanding of that game’s rules could change over time. No finite list of rules could include an infinite number of possible variations.
3. Since any two players could be playing the same game with different interpretations, there would have to be a set of meta-rules for reconciling these differences when they surface.

Of course, these meta-rules are, in effect, the rules to another game and are therefore subject to the same interpretive variations as the rules of any other game. Again, we run into an infinite regression. There is no bottom line, no point when we can accurately say, "These are the ultimate meta-rules for settling disputes."

Thus, the questions, "How do we settle disputes about the rules themselves, about whether a player has violated a rule, and about the appropriate penalties for a rule violation?" can never have a final answer.

4. Even if two players agree on certain rules and how to interpret them, disputes about what actually occurred (such as whether a ball landed on the back line or just beyond it) can still arise, and the players will need to abide by meta-rules in settling these disputes as well. These meta-rules, like those in #3 above, are also part of an infinite regression, so the question "How do we settle disputes about what really happened?" has no ultimate resolution either.

5. Since there are various "levels" of rules, "higher" rules (such as a real-world crisis) might have to take precedence over "lower" rules (such as time constraints); there must be a set of meta-rules for determining when this is appropriate. As with the other meta-rules we've looked at, there is no "final" set for ending disputes, so the question, "When is it appropriate to suspend certain rules?" cannot be given a full answer.

6. Since all games begin and end and may be interrupted by "outside" events (such as a TV ad), we must have a set of meta-rules for determining when the constraints apply and when they don't. Again, these meta-rules are susceptible to interpretation and dispute, leading to yet another unending regression.

"Simons" often take advantage of this fact by tricking players into thinking play hasn't begun and then saying something like, "Before we start, say hi to your neighbor. Ah, I didn't say 'Simon says.'" Therefore, the question, "When do the rules apply?" cannot be fully answered.
We can see now why it is impossible to spell out a complete set of rules for any game. Now we need to ask why we have no trouble playing a wide variety of games.

**If we can't know all the rules, how can we play any game at all?**

*Is it because participants rarely have to deal with "meta-rules" and so the infinite-regress problem almost never comes up?*

To me, this is not a plausible explanation. There are simply too many occasions we can name—in virtually every game ever played—in which meta-rule questions arise. When a player accidentally rolls the dice off the table, argues a call, gives (or refuses to give) an opponent a handicap, calls for a do-over, takes a mulligan, asks for a director’s ruling, warns an opponent about an unwise move, or encourages the other team to play faster, the players are facing situations that are not [and could not be] completely covered in the recorded rules. Meta-rules (and even meta-meta-rules) are an integral part of all rule-governed activities.

*Is it because players don’t take games seriously so it doesn’t matter that they can’t know all the rules?*

Again, this doesn’t work for me because it is clearly not true in all cases. Obviously, some players (myself included) care deeply about the game and the outcome. Many of us are playing for high stakes—money, prestige, a trophy, pride, self-esteem, ego satisfaction, a feeling of control, etc. In fact, it’s probably pretty rare for players to have no emotional involvement in the game they are playing. After all, why play unless the results “matter” in some important way?

My guess is that almost all players almost all of the time take almost all games very “seriously.”

*Is it because players mistakenly believe that there is a “bottom line,” that the rules are clear, complete, and “final,” and that somebody somewhere knows all of them?*
This is getting closer to sounding right, but is still a half-truth at best. Having the misconception that a game’s rules are solid and statable can provide a player with a sense of confidence in the “reality” of a game, but my realization that no one can know (let alone state) the rules of our doubles game has not dampened my enthusiasm for tennis one iota. In fact, my recognition that games, like languages, can exist only because of an unspoken, almost mystical, agreement among the participants actually enhances my appreciation of them.

Although my attitude may be idiosyncratic, I seriously doubt that anyone else’s enjoyment of a game (or willingness or ability to play it) would be diminished by realizing that we can’t list all its rules.

“It’s only a game”

I believe we can go on playing games wholeheartedly even when we are aware of the incompleteness of their rules. Why? Because, on a gut level, we cannot distinguish between something fanciful—like a movie or a joke or a dream or a game—and something “real.”

Games feel like any life-event, so we can be immersed in them even though we may know intellectually that they are artificial constructions. Therefore, it makes no difference to us (emotionally) that a list of rules governing them cannot be completed, just as we can be profoundly affected by a joke or piece of fiction or nightmare that is not logical, realistic, or “complete.”

We can suspend disbelief and rationality (even when some part of our brain is telling us it’s only a story or it’s only a dream) and respond deeply to creations of the imagination—our own or others’.

We can do this because we have the wonderful (and perhaps unique) capacity to operate on the “as if” level; we can play a game as if we know all its rules, as if there is an ultimate set of meta-rules, as if all potential disputes can be settled. We can imagine a game in the abstract and in a vacuum and can project that Platonic ideal onto the one that must be played in the world of social and political reality.

In other words, we can operate on (at least) two distinct levels of cognition at once. We can play any game as if it had an autonomous existence, even though we know perfectly well that the players create the game each time they agree to play and that any player at any time can destroy the game by quitting, by arguing, by stalling, or by any number of other spoilsport tactics.
Similarly, we can play any game as if it is important (and genuinely feel that it is), even though we know that it is not very high on our list of life priorities. We can play any game as if it transcended our culture, even though we recognize that players can have "unfair" (dis)advantages as a result of their upbringing. We can play any game as if it transcended morality (so we might intentionally and unashamedly foul or fool an opponent) even though we know that players can cheat or violate the rules in inappropriate ways.

Suspension of disbelief
Without this ability to operate in the conditional universe of "Suppose..." and "What if ...," game playing would be impossible, as would drama and fiction and, I suspect, language itself. We must be able to behave as if a game were not "merely" play, even though we are fully aware it is nothing else.

Like an actor, we must be able to take on a role but never give up our sense of self. We must be "in" the game to enjoy it but never so far in that we forget who we are. It is a delicate balance fraught with danger, which is perhaps why so many people (especially adults) shy away from games.

Non-human game players?
It is also, I believe, one more reason that computers (at least as they are today) will never play a game in the same sense that humans do. Computers have no conditional, no ability to create temporary self-delusions, no play mode, no sense of "as if." To a computer (we must assume), a chess move is just another calculation, no different from finding the square root of pi.

To a human, a chess move is (usually) part of a carefully designed pretense, a system of orchestrated assumptions, an artificial structure that can bring stimulation, competition, camaraderie, fun, and a variety of other good feelings. In general, the chess-playing human voluntarily accepts a particular challenge that involves a specific goal and specific constraints and which s/he can abandon at any time. The chess-playing computer, on the other hand, does not choose to start and cannot stop on its own. The human is aware of the voluntary and "non-serious," conditional nature of the activity, but the machine is not (and probably can never be).

What about animals? Does any non-human creature have the ability to suppose, to imagine something that doesn't exist except as an agreement among participants? If not, they
will never play a game as we do. They will either take it too seriously or not seriously enough and, therefore, like any spoilsport, undermine the enjoyment of the game for any human participants or observers (as was the case with the water-polo-playing dolphins).

But even if animals (or computers) could think in the conditional, they still might not be able to play games as we do. They would also have to be able to trust other players to function in basically the same way. To play a game (or use a system) meaningfully without knowing all the rules requires the faith that others understand the game/system as you do or at least will behave in ways that seem consistent with such an understanding. Without that faith, a player would inevitably end up being the spoilsport.

**Meta-rules in other arenas**

By way of analogy, consider our (or any other) monetary system. Most people recognize that the currency we use has no inherent worth and that it gains its value from mutual (if tacit) agreement among its users, which means its value is subjective, symbolic, and subject to change.

Few people believe that there is an objective, stable method for determining how much milk a dollar should buy. Most of us understand that there are no “rules” or meta-rules we can refer to that would settle a dispute about the value of a dollar bill and that its purchasing power is dependent on consensus, on other people’s willingness to give us this much milk for this many dollars. And yet we can still use the coin of the realm and, for the most part, get our money’s worth (by our own standards).

The system works even though no one can explain it fully and even though we all know it could collapse at any moment if people stopped trusting each other or the system itself.

The same is true with another currency—language. Even though words have no inherent meaning and no one has been able to list all the rules governing the construction of sentences, we can still communicate reasonably effectively for most purposes.

We all know that anyone at any time can choose to destroy the process by acting on Humpty Dumpty’s premise that words can mean whatever we want them to mean. We know that there is no rulebook, no authority, no indisputable arbiter we can appeal to in such a case (since they would all have to use more words to settle the dispute).

Like any game, communication is dependent on the participants’ willingness to operate as if there were universal agreement about meanings and grammatical rules.
We need to remember, though, that games are not analogous to these two currencies in at least one crucial way. Both money and language, after all, serve obvious, vital functions in the world, whereas the value of games is not nearly as apparent. We can easily understand why people would almost always try to go along with a monetary or linguistic system, since they believe that both can benefit them and the community significantly. In addition, most people recognize that destroying either system could ultimately threaten their own well-being.

**Rule-preserving meta-rules**

But games? The common perception is that no one gets hurt if a game is spoiled. So why would anyone continue to submit to an arbitrary (and incomplete) set of rules that was causing him or her to lose face, patience, and/or money? Why do people continue to play "by the rules" when they are losing the game?

Since losing is undesirable, we need to explain why so few players take advantage of the fact that the rules are incomplete and therefore infinitely challenging. We need to understand why people almost always play as if the rules were not only complete but knowable and statable, and rarely allow themselves to play the meta-game of arguing about the rules and the meta-rules, ad infinitum.

One possible answer, of course, is that players don't realize that this "strategy" exists, but I think that all of us have witnessed many examples of the kind of behavior I'm talking about. Almost everyone has seen images of managers and players, nose to nose with an umpire, arguing a call or an interpretation of the rules, and even non-sports fans have probably seen TV ads based on John McEnroe's antics on the court, so I have to assume that virtually everyone realizes that this option is theoretically available to any player.

So what are the real "meta-rules" that keep most of us from playing this particular meta-game? Here are a few of them:

1. A game is supposed to be for fun, and, playing the game itself is more fun than playing the meta-game of arguing. Except for young boys in the frontyards of America (who will argue endlessly about a single play), most players have learned that the meta-game is boring, repetitive, and fruitless, often ending in a stand-off;

2. A game is supposed to test certain skills, and these do not usually include the skills of debate, sophistry, and intimidation tested by the meta-game;
3. A game is supposed to be for camaraderie, and arguing about the rules leads to antagonism rather than a spirit of friendly competition;

4. Players are supposed to be good sports (whatever that is), and rule challengers are perceived as poor sports or even spoilsports;

5. The "ideal" game, the game we all want to play, works fine as it is and does not include a discussion of rules or meta-rules;

6. A set of rules that has been tested is better than one that has not, so if it's not broken, don't fix it;

7. Doing things as others have done them in the past allows us to feel connected to our ancestors, our culture, and our traditions;

8. Following the rules that others follow allows us to compare ourselves to a wide spectrum of players, not just our immediate opponent(s);

9. Challenging long-standing traditions is inherently unwise because it creates the impression that nothing is sacred and could, if carried far enough, lead to anarchy.

For all these reasons, a player who argues about rules risks disapproval, sanctions, and even ostracism, so the vast majority of us choose to "leave well enough alone." Most people avoid and frown on the meta-game of arguing with rules and meta-rules because, without necessarily being aware of their reasons, they perceive it as a threat to pleasure, continuity, and stability. Thus, most games continue to be played "as they always have been." For the same reasons, many people are suspicious of new games.

To return to our central question, then, we can play a game even though we can't know all its rules because, for a variety of reasons, we tacitly conspire with our fellow players to act as if we know them all.

The big picture
In this way, games are no different from every system we use. In an important sense, all rule-governed systems—including law, politics, war, morality, education, economics, and language—are games, as many people have noted. Therefore, virtually all of the lessons we learn from "non-serious" games are directly transferable to the "real" world. What are those lessons? What follows from the acknowledgment that no human system has a completable set of "rules?" Let us spell out some of the implications.
1. Power and authority are arbitrary, not inevitable, depend on consensus (or at least acquiescence), and have no "divine" right to exist.

2. Rules for any system are not handed down from above, can exist only through the agreement of the participants, are always open to negotiation among the "players," and are continually evolving. As Robert McConville reminds us, if a game survives, "the rules for playing the game are constantly being changed as they are passed from tribe to tribe and generation to generation" [The History of Board Games, p. 8].

3. The most powerful rules, the ones least likely to be violated, are those that are not stated explicitly, those that people have to infer or intuit. To state a rule is to invite players to break it, but to leave a rule unstated is to make its violation almost literally "unthinkable."

4. We cannot accurately predict how any rule, stated or unstated, will be interpreted or enforced, so no rule, simply by its existence, will necessarily produce or prevent a desired behavior.

5. We cannot accurately predict or control what customs, norms, conventions, traditions, or expectations will evolve for a particular game or system of rules.

6. No set of rules is inherently superior to any other. In order to judge a set of rules, we must employ a set of meta-rules, which themselves would have to be judged by a set of meta-meta-rules, and so on ad infinitum.

7. An infinite number of sets of rules will "work," will allow us, individually or collectively, to function successfully (or at least to our own satisfaction).

8. The longer a system is followed and the more people who attempt to follow it, the more complex the recorded rules will become, and the more sets of meta-rules and meta-meta-rules, etc., will be recorded. Consider any legal system, religion, or professional sport as prime examples.

9. Every person operates according to an unlimited number of sets of rules, so it is almost inevitable that some of these sets (such as religion and business) will come in conflict with each other, which means that every person is also operating according to an unlimited number of sets of meta-rules for reconciling such conflicts, and an unlimited number of sets of meta-meta-rules and so on.
10. As humans, we have little choice but to act as though some of these sets of rules were absolute and indisputable. Otherwise, we would be trapped in an infinite regression and utterly unable to make meaningful choices.

11. Paradoxically, we cannot live according to any set of rules (because we can never know them all and because they will inevitably conflict with other sets we are trying to live by), so in order to continue to perceive ourselves as faithfully following a “complete” set of rules, we must learn to rationalize our deviations from it (or feel a great deal of guilt).

12. It is reasonable to say we are playing a game/living by a system even though we are not following all its rules. For this reason, following some of the rules in a system creates the expectation [in ourselves and others] that we will follow all the rules, including the unstated and the unstatable ones.

13. No one can tell for sure if someone [including oneself] is “really” playing a game/living by a system because it is not possible for anyone to follow all the rules in a game or system. Therefore, we can pretend to be playing any game/living by any system without others being able to detect that we are pretending. We can also pretend to be pretending and so on, and no one will be able to tell the difference.

14. No two people can possibly follow the same set of rules in exactly the same way.

Obviously, the recognition that we cannot know all the rules in a system can have a profound effect on how we approach the world. It can make us want to curl up in a corner with our thumb in our mouth or to go out and make sweeping changes in our most important institutions. It can destroy us or free us, depending on how we feel about a world in which there are no absolutes, no bottom lines, no final list of rules, a world in which all systems are “equal” and all meaning relational. Some [including myself] are comfortable with, even invigorated by, this notion, but others [perhaps a large majority] are enormously disturbed by it.

**Today Parcheesi, tomorrow the world**

Of course, there is nothing new about the relativist claim, but, to my knowledge, no one has applied the concept to games, those obviously artificial constructs. The argument has raged about more “important” human systems, like law and religion and language, so emotions, desires, and values always tend to cloud the issues. People understandably want to believe
that their beloved institutions are sacred, unchanging, and right, but [almost] no one feels that way about games.

So I have chosen to examine the reality of rules and meta-rules in this non-volatile, "safe" context of games, hoping I would not scare away those who tend to shun a relativistic argument. My goal has been to show convincingly that we cannot know all the rules but we can still play the game, so that I could suggest, through analogy, that

\[
\begin{center}
\text{We can go on using (and revering) any system even if we acknowledge that it is as artificial, arbitrary, challengeable, and "incomplete" as any game.}
\end{center}
\]

\[
\begin{center}
\text{Any system, no matter how long it has been around and no matter how complex its list of rules and meta-rules, is viable only as long as there are individuals who support it.}
\end{center}
\]

**Conclusions**

If my efforts have been successful, if people take away valuable lessons about "life" from this analysis of games, it will demonstrate, ironically, that games can indeed serve at least one vital social function: as abstractions of "real-world" situations, they can provide an analog to other, more "important" and more complicated, aspects of life and thus can help us see what otherwise might be invisible. If for no other reason, games should not be dismissed as trivial forms of entertainment. If we remember to use them wisely, they could be a profoundly important aspect of our culture. As the young would say, GAMES RULE!
constituting an instance of that game may require a very different descriptive and analytic framework. Gary Fine [1980] has used the term "gaming" to distinguish between the game itself and what real players do in constructing a particular instance of that game. In this paper we will explore, at least in a preliminary way, something of the differences between game rules and the rules of gaming. In the process we hope to put some "play" back into games.

**Games and Gaming**

For some time it has been the vogue to adopt games as metaphors for social interaction (Bateson, 1972; Goffman, 1961; Harré, 1974), but only rarely have we approached games from a social communicational interactional perspective. Goldstein [1971] provides a notable exception, as does Gary Fine's [in press] recent analysis of fantasy role-play gaming. In both we find a good deal of "ambiguity, spontaneity and flexibility." Fine, for example, observes that

In fantasy role-play gaming, rules and outcomes do not have the inevitability that they possess in most formal games, rather both features are negotiated and rules are adjusted by the referee and his group. Thus, ironically, fantasy games are in some ways more like "life" than like "games," despite their position as games. [p. 4]

It is not the particular game Fine studied that sets it off from "most formal games," but the type of analysis he has chosen. His game looks more like "life" than like "games" because he is not describing the game for us here, but the social and interactional activity of gaming. Looked at from a similar perspective, even professional sports, which we tend to regard as the most formal of "formal games," also display much more flexibility than might be suggested by their rules.

To the NHL's thirteen referees, life is a constant struggle to maintain the game's flow, keep the coaches off their backs and the players off each other's. Most important, referees must have an instinct for which violations to call and which to ignore. They themselves talk of "good" penalties [flagrant violations such as tripping the player with the puck] and "bad" ones [minor offenses such as hooking a player who doesn't have the puck late in a tight game]. "You could call a penalty a minute," says referee Ron Fournier. "But that's not what we're supposed to do. You call a guy for a minor infraction, and even though you cite the rule number, he just looks at you and says, 'What's that?' It doesn't earn you respect. [Newsweek, January 5, 1981]
This referee is well aware he is supposed to be providing something more than a literal reading of the rule book. Maintaining the game’s “flow” and earning the respect of other participants depends upon knowledge of both game rules and gaming rules. One must not only know how the rule book defines and penalizes “hooking.” One must also know that this “hook” is not the same as that “hook.”

If we adopt, for the moment, Goldstein’s distinction, we might distinguish between an “ideal” game rule, which refers to an action (a “hook”) and to a game-prescribed outcome (a penalty), and a “real” rule, which refers to a “hook” as something other than a contextless action and to shared, socially-prescribed, negotiated outcomes (interruptions in game flow, loss of respect and such). The term “respect” makes reference to social relationships, to something of a very different order than game rules. It derives from the status of games as a social process, not from the status of the activity as a game.

When we encounter words like “respect,” we move from an analysis of games into the realm of gaming. The quotation at the beginning of this chapter (from the author’s observations of a group of young Foursquare players) similarly moves us out of the game and into gaming: “Why Rooie Rules?” “Because Rooie Rules are nice.” We will be exploring Rooie Rules as game rules and as “nice” rules. In the process, we will be asking whether it is possible to understand these rules without also understanding why they are “nice.”

The World of Foursquare

For almost three years, children were observed playing the game of Foursquare on the lower school playground of a private Friends school in the suburbs of Philadelphia, Pennsylvania. More recently extensive interviews were conducted with many of the players. The following is drawn from a larger research effort focused on the negotiational aspects of gaming, and especially on identifying some of the largely tacit understandings that allow these children to construct this kind of complex communal activity.

The groups of children observed come generally from white, middle to upper socio-economic groups. Less than a third come from Quaker families. The game of Foursquare tends to be dominated by girls, especially those in third to fifth grades, but boys have also been observed in the game, with and without girls. Older and younger children regularly join in. At the time of the observations cited below, the “regulars” consisted largely of third and fourth grade girls.
Children in this school have played Foursquare during recess for more than twenty years. The game is played on a square court painted on a portion of paved playground. The court is further divided into four equal squares, each of which constitutes the play space for one of the four active players. The player occupying one of the squares is called the "king." Before each round of play, the "king" calls a set of rules for that round (in our opening example, "I call Rookie Rules. Duckfeet."). This player then serves a large rubber ball to one of the other players. The ball is bounced among the players until it bounces more than once in a player's square, or until a player fails to hit it into the square of one of the other players. That player is out. He or she leaves the court and goes to the end of a line of players waiting to enter the game. The remaining players rotate toward the "king" square, filling in the vacated square, and the first player in line enters the game through the square farthest from the "king."

Foursquare is a simple ball-bouncing game. One of the ways it is made before each round of play. [sic] Such calls can be used for a wide variety of purposes, including increasing game excitement, adjusting the level of difficulty, and assisting or scapegoating other players. At one level the "king's" rules prescribe or prohibit certain actions and specify their consequences. In this sense they function as game rules. At another level, they may display or set the general tone for a particular round of play or for the overall gaming occasion.

It is this latter quality that allows the call of "Rookie Rules" in this game. A girl named Rookie was one of the regular players when the current observations were made. The king's call makes reference to her preferred set of rules, which include the following: "no holding" ([the ball must be hit, not caught and thrown]; "no slams" [bounces high over a player's head]; "duckfeet" [being hit on the legs] is out [rather than a "takeover"]; "spins" are allowed; and so on. Each of these individual game "rules" can be called by a "king." The call of "Rookie Rules" is a kind of shorthand, covering a long list of individual calls.

Despite the fact that play regularly proceeds after a call of "Rookie Rules," no player, including Rookie and the "king" who calls them, can supply a complete list of rules encompassed by this call. In fact, this call is very regularly used by very inexperienced players to avoid having to specify a particular list of rules before they have learned what calls are possible. What allows the game to proceed within such apparent ambiguity concerning the precise rules of the game is the tacit understanding that Rookie Rules are "nice," and "nice" is perhaps the paramount concern among these players. It is far more important to understand "nice" play than to understand the rules.
What Makes Rooie Rules “Nice”?

How are Rooie Rules “nice”? First, even though they are not explicitly mentioned and most of the players could not list all of them, Rooie Rules are understood to prohibit all kinds of what the kids call “rough stuff.” This includes all moves like “slams” or “wings” (hard, low shots to the corner of a player’s square) which are difficult for the receiving player to return. “Nice” players are supposed to give others a fair chance of returning the ball and, even more basic, they are not supposed to try to get other players out of the game deliberately.

Second, Rooie Rules are “nice” because they prohibit “rough stuff” in a “nice” way. Even when prohibited by a call such as Rooie Rules, players may still “slam” and “wing” and “hold” without being called out for doing so. When observations of this game began, this seeming lack of direct relationship between the rules as called and being “out” was particularly puzzling. Why bother to call the rules if no one was ever out for violating them?

Things became clearer one day when one of the players explained to another the rule really was “please don’t hold the ball unless you really have to.” This interpretation allows more experienced players to be “nice” to “little kids,” who can do little more than catch and throw the ball. It also reflects a shared sense that it is unfair to penalize players, who, in the heat of play, lack sufficient control to avoid holding the ball briefly before returning it. Fairness, it seems, is an important component of “nice.”

Later conversations with players confirmed that actions such as “wings,” “slams” and “holds” were not really prohibited by Rooie Rules though one should try not to do them. What was prohibited is what the kids call “purpose stuff.” A common cited example of the latter is “holding” the ball while deciding which player to get out of the game. Perceived intentionality joins fairness as another component of “nice.”

Third, Rooie Rules are “nice” because they avoid direct confrontations over player actions. Even the most blatant “purpose stuff” is rarely directly challenged. Certainly, no player would ever be called out for such violations. They may be simply ignored, especially when directed toward a player scapegoated by a dominant group of players. If not ignored, they will be handled in a less direct—and “nicer”—way. It is very common to observe rather elaborate performances—exaggerated leaning, grunts, cries of “whew!” and dramatic mopping of brows—around rather easy shots that just happen to land as a “slam” in the offending player’s square. This latter observation underscores perceived intentionality as a major component of
“nice” among these gamers. One is reminded of Goffman’s [1959] concept of “demeanor.” It is apparently less important here that one be “nice” than one make an appropriate display of being “nice.”

Highly ritualized “yes you are” / “no I’m not” exchanges are very common. Again, from fieldnotes:

Angie: “Sally! You’re playing rough!”

Sally: “So are you!”

Angie: “No I’m not! I’m being nice!”

Smiles, sideways glances, and the glints in players’ eyes when they engage in such exchanges belie their seriousness. There is a quality of collusive “play” around ways of making deliberate actions look accidental, a shared delight in a virtuoso performance or a comment on a performance that just misses the mark.

It is only when “purpose stuff” does not have this playful, among friends quality that one is likely to see sanctions applied for violations of the “rules.” Players who consistently “slam” or “wing” without this playful collusion quite often find themselves on the receiving end of a wild shot they cannot return. Such occasions are always followed, of course, by profuse apologies from the hitter, who “just couldn’t help it,” and who puts on quite a performance to support a claim of innocence in the whole matter.

This use of indirect sanctions, or at least “accidental” sanctions, bypassing a simple appeal to the game rules, is also part of being “nice”. If the “real” rules of this game tend to revolve around perceived intentions of the player (as in “purpose stuff”), then it becomes rather awkward to invoke those rules and still be “nice.” Invoking a “rule” is not merely a statement of fact about a player’s actions, but an accusation of having violated something of the social order, a much more serious charge. Among these gamers, invocation of such rules would involve not only explicit recognition that all players are not equal under the rules, but also an implicit accusation of being purposefully vindictive or nasty—of not being “nice.” It is not “nice” to violate “nice” rules, but it is also not “nice” to accuse someone directly of doing so. Instead, sanctions are imposed in a way which allows everyone to act as though they were accidental, accompanied by an appropriate “I couldn’t help it” performance.
At least among the gamers observed here, the call of “Rooie Rules” invokes less a list of individual rules’ calls than a general framework for player interaction. The latter rests upon shared standards for fairness, perceived intentionality and appropriate demeanor within the group. The term “nice,” and its contrast “rough,” are employed among this particular set of gamers to refer to a rather complex matrix of social rights and obligations. It appears that understanding and accepting these standards is even more critical to sustaining the game activity than understanding the set of rules currently in force. This became particularly clear one day when a group of boys joined the regular female players. The boys clearly understood the “real” rules of the game, as the “regulars” played it, but they actively challenged the implicit demand that they play “nice.” In the process, what was usually implicit became more and more explicit. Their behavior triggered a rather active discourse concerning the “real” and “ideal” rules of this game. Some very interesting things began to happen.

When Players Won’t Be “Nice”
As might be predicted among boys and girls of this age, the boys almost immediately drove the girls crazy by very overtly using “rough stuff” (“slams” and “wings”) to get the girls out of the game. This does not mean the girls were not also using such moves. What enraged them was the boys’ failure to disguise “purpose stuff” in the kinds of “I couldn’t help it” performances demanded by “nice” play. The boys would, for example, call, “Rough square, Getting out on serves,” and then slam the ball high over one of the girls’ heads on the serve.

Totally outraged, the girls would counter, when one of their number was “king,” with a call of “Rooie Rules.” But, as we might expect, calling “nice” rules had little effect. The boys blatantly continued to “slam” and “wing” the ball past them. Since the girls were still bound by their “nice” rules, which prohibited direct confrontation over such actions, there was little they could do. As play proceeded, however, the girls gradually abandoned some of the trappings of “nice” play. They began handling violations quite differently. The following are excerpts from fieldnotes. We begin with three girls and one boy on the court.

Angie [the “king”]: “Rooie Rules, Rooie Rules.”

Angie pauses, looks around, and then walks over to the players waiting in line to get into the game.

Angie [to Rooie, who is waiting in line]: “Rooie, tell them your rules.”
As Angie returns to her square, she glares rather pointedly at Hoover, the boy who has just entered the game, while Rooie lists her rules.

[It should be noted that another understanding among these gamers is that players are only responsible for violating a rule they know about. Only if they know, and then violate, a rule can they be denied a takeover of the last round. This attempt to list very explicitly the rules in effect is highly unusual. It functions as a kind of warning to the offending players.]

A little later, Cindy [who is now the “king”] calls: “Rooie Rules.”

But Andy continues to “wing” and “slam” the ball consistently. After several such hits, Rooie, who is waiting in line, walks over to Andy’s square.

Rooie [to Andy]: “You’re out! Wings are out!”

Cindy steps forward to back Rooie up.

Cindy [to Andy]: “I called Rooie Rules and there’s no wings! You’re out!”

As Andy leaves the court he mumbles something about being a “fish.”

The term “fish” refers to a scapegoated player. In over six months of observing this game, this was the first time the author had observed anyone being called out for “wings.” The exchange above is a very significant departure from the usual patterns of play. Andy is well aware of this. He knows he’s been had.

The girls’ revenge was shortlived, however. In reacting to the boys’ refusal to play “nice” by becoming more explicit in their calls of the rules, and by applying direct sanctions for violations, the girls began digging themselves into a rather deep hole. They expanded a call of “Rooie Rules,” for example, to “Rooie rules. No slams. No wings. No rough stuff.” They tried explicitly to prohibit each of the boys’ offending actions. Naturally, the boys could always find actions the girls had not specifically prohibited. One particularly exasperated “king” recognized the problem when she tagged her call of the rules with, “And nothing you guys do!”

Of course, on the other side, the girls could not completely avoid violating their own rules, now differently defined. The boys were only too happy to point this out. Again, from fieldnotes, we start with four girls as active players on the court.
Angie ("the king"): "Fair square. Rooie Rules. Fair square."

(The call of "fair square" means no one should get anyone else out. It is a reference here to a desire to keep the girls in, and the boys out of the game.)

The boys waiting in line can be heard mumbling something about "holding."

In an unusual move, Angie stops the ball and turns to Sandy, a young and inexperienced player.

Angie: "No holding, Sandy."

She immediately puts the ball back into play.

More grumblings can be heard from the boys in line.

A little later in this round of play Andrea does a "double tap" before hitting the ball to the next player.

("Double taps" is a very common move, especially during a low pressure round like that framed by "fair square." It is a type of "fancy" move, in which the ball is tapped twice in the air before being hit to the next player.)

One of the boys in line immediately steps forward and shouts: "Holding!"

The other boys chime in, accusing Andrea and the other players of holding the ball, a move prohibited by Rooie Rules.

The girls on the court try to ignore the ruckus. They continue to play, as Angie (the "king") protests: "They can't help it!"

(The appeal here, of course, is to the "real" rule which makes reference to "purpose stuff," not to the act of "holding.")

The boys continue to complain. Angie finally turns to them and says: "Okay, then. I call 'holding.'"

A little later Angie again calls: "Rooie Rules." But now she appends: "And there's holding."

Mike, one of the boys who at this time was playing regularly with the girls, shouts sarcastically: "Ha! Ha! Ha! Ha!"

Andy (rather pointedly): "Did you call holds?" Angie directs a withering glance at the boys, and starts the next round of play without comment.
The girls are trapped, and both they and the boys know it. The cornerstone of "Rooie Rules" is "no holding." The call of "Rooie Rules—and there's holding" is totally contradictory and unthinkably within the normal course of play. We have reached a point where the "ideal" logic of the game has been invoked, where rules have become rigidly pegged to player actions and game outcomes. In the process, the logic of the "real" rules has become paradoxical. The players have lost the leeway which allows the usual understanding that this "hold" is not the same as that "hold." The flow of the game, as these particular players constitute it, cannot survive such a rigid linking of rule and action.

**Rules and Paradox**

The kind of pickle the girls find themselves in looks suspiciously like the stuff of classical logical paradox (Bateson, 1972). They are left holding a bag very neatly labelled "no holding," and simultaneously tagged "this is holding" and "this is not holding." This is not normally paradoxical at all within the everyday logic of the game as these players constitute it, but it has become so in the confusion of action and rule. The philosophers might have helped them out by demonstrating that the problem is really one of propositions differing in order of abstraction, a simple error in logical typing. "Holding" and "nice" are not of the same logical order.

It's unlikely, of course, the players would find such an argument very persuasive or very relevant to their predicament, because they don't confuse the two in the normal course of events. All conversations with these children indicate a clear recognition of the difference between the way it's supposed to be in an ideal sense, and the way it is. The two are intricately and elaborately interwoven into a consistent and coherent framework which sustains the complex communal activity of Foursquare.

Some recognition that game rules and player actions, and the interpretive gaming scheme which binds them together, are of different orders may be more useful to us, as researchers, than to these players. In our thinking about games, we create our own paradox when we take the "play" out of games, the "fun" out of gaming, by treating a description of the game rules as descriptive of the activity or experience. Something of the same phenomenon seems to be at work, too, when we speak of games as frames, as though games frame activity, rather than gamers.

Games aren't much "fun" when rules, rather than relationships, dominate the activity, when there is no attention to "flow," "fairness," "respect" and "nice." We need the leeway to
be playful in these relationships, to share and enjoy the performance that sneaks nastiness
by as nice, that displays knowledge of the "ideal" rule and plays with the boundaries of the
"real" rule. Taking the leeway out by treating all rules as rigidly prescriptive and tied to actions
subordinates "fun," "flow," "spontaneous involvement" (Goffman, 1961) to the activity. In actual
gaming, as Goldstein (1971) and Fine (in press) remind us, quite the opposite may be true.

Studies of this type suggest a number of assumptions about game rules that require
active reconsideration. First, our thinking about rules in the games literature has been rather
simplistic and monolithic. In other social science and philosophical traditions, the whole notion
of a "rule" is considered to be highly problematic. Rules are assumed to be of many different
types, multi-layered and hierarchical, referencing very different antecedents and outcomes.
They are assumed to be subject to constant negotiation and reinterpretation in the course of
everyday life. The current study, for example, suggests all rules are not equal and all players
are not equal under the rules.

Second, because the activity is a game, we have tended to assume game rules are
at all times explicit and foregrounded. We have taken for granted all debates and disputes
concern the game, not the relationships among players. Observations of players, however,
suggest the degree of rule explicitness may constantly shift, and such phenomena are highly
contexted and indicative of social relationships among gamers (Erikson and Shultz, 1976,
Shultz, 1976).

Finally, we might propose that the apparent paradoxes and transformations of
"play" derive less from the logic of players than from the logic of our descriptions. Contradic-
tion seems almost inevitable when we confuse the logic of the game and the logic of contexted
gaming.

Epilogue
Just to close the story begun above, this particular gaming occasion broke down into one of
the few complete stalemates observed over a period of several years. Whether a prohibited
move had occurred or not simply could not be resolved by appeal. The girls refused to be
"nice" and give an offending player, a boy, the benefit of the doubt, as they usually do by allowing
him a "takeover." Just before the bell ending the recess period, we find Angie turning to one of
the boys and suggesting the only acceptable mode of resolution: "Okay, we'll give you guys
another chance. But only if you promise to be nice."
Note
An earlier version of this paper was presented to the Annual Meetings of The Association for the Anthropological Study of Play [TAASP], Fort Worth, Texas, April, 1981. The author wishes to thank Brian Sutton-Smith and Alexandra Hepburn for their very helpful comments on early drafts of the paper.

References
What are the rules of a game? Ask almost anyone and they'll probably mention at least one of the following:

1. They're what's printed on the rule-sheet of a published game - or, in the case of a traditional game, in a book said to be written by some character called Hoyle. Either way, they're essentially some sort of written description of how the game goes.
2. They're the "official" rules which you must follow if you don't want to be thrown out of the club. They are essentially prescriptive and proscriptive (they tell you what you must and must not do).
3. They're instructions designed to teach the uninitiated how to play the game.

Already the concept of rules starts to cover quite a lot of ground. They may be descriptive, prescriptive, proscriptive, or instructive. But they can be more than that. For instance:

4. There are rules of appropriate behaviour, like not starting to pick your cards up before they've all been dealt. These are sometimes called the Proprieties, or Etiquette.
5. There are rules of sanction and correction, which only operate when players contravene the basic rules of play. These are often known as the Laws.
6. Then there are the rules (so called) of strategy. These, indeed, are the only rules covered by Hoyle himself in his earliest book, A Treatise on the Game of Whist (1742), which offers "some Rules whereby a Beginner may, with due attention to them, attain to the playing it well..."
7. Consider also games that require or at least invite you to play a part. In Chess, you are a commander in charge of an army; in the Chinese Game of Promotions, you are a prospective Mandarin; in Monopoly, you are a property developer and business tycoon. If these are rules at all, they may be designated dramatic or liturgical. You may decline to call them rules, yet they are a valid part of the
meaningful experience of playing the game and can't be entirely omitted from any comprehensive account of it.

So let's start again and try imposing some structure on formulating an answer to the question "What are the rules of a game?"

**Explicit rules**

It is widely assumed that all games have official rules that are recorded in writing. But it is mistaken. For one thing, most games are not book games but folk games, being transmitted by word of mouth, example, and practice. For another, even where written rules do exist, probably no folk games and certainly very few book games can lay claim to a widely recognised governing body responsible for authorising them.

Nevertheless, all games must have rules of some sort, otherwise they cease to be formal entities and become merely undefined periods of unstructured play. The most basic rules of a game are not a form of words but a set of operational procedures you apply to the gaming equipment in order to play the game. Following Salen and Zimmerman, I refer to these as the operational rules. Operational rules are what you apply to the hardware of gaming equipment to produce an instance of play. I will put this into diagrammatic form as the basis of a model of a game that I intend to build up during the course of this paper:

**Operational rules (software) > Equipment (hardware) > Play (application)**

Operational rules are explicit, in the sense that they can be verbalised, even if not necessarily recorded in writing. The realisation or embodiment of the rules in a physical set of gaming equipment may be likened to the hardware of the game, or, by another analogy, to its skeleton. When we unearth the remains of a game in an ancient tomb, we find only the skeleton: we cannot know (or be sure we know) how the game was actually played. The application of the rules to the equipment is what produces the actual play of the game. A game must be played to be fulfilled, or to have meaning. Every time you play a game you are breathing life into its clay in accordance with its operational rules, and it becomes what Salen and Zimmerman describe as a "meaningful experience of
The most basic level of experience suggests that the rules of a game are something inherent in the game itself - or, more accurately (since a game is essentially a mode of behaviour), an abstraction existing in the minds of all its players. They are expressed in words every time someone describes a game or explains how to play it. Not everyone will have exactly the same understanding or grasp of the game, so they're unlikely to transmit their knowledge in exactly the same form of words. These rules are therefore not a known quantity but an average of all the understandings of all the players. As such, they may contain inconsistencies. The totality of rules of all but the simplest games are not exactly a cloud of unknowing, but could be described as a cloud of fuzzy knowing. In telling you how the game is played, they serve to establish its formal identity. Huizinga, in Homo Ludens, says "Every game has its rules"; but I would go further and say "Every game is its rules, for they are what define it".

For example, Noughts & Crosses (Tic-Tac-Toe) is defined in the Oxford English Dictionary as:

```
a children's game with a figure containing nine spaces [in three rows of three], which are filled up by two players alternately with ciphers and crosses, the object of each being to place three of one kind in a line.
```

This definition clearly expresses the rules of Noughts and Crosses, but it is equally clearly a description rather than a prescription. It says nothing about authority or obligations but merely asserts the identity of the game. It relates the name of the game to a set of procedures, and a set of procedures to the name of the game. It doesn't say you have to play it this way. All it says, in effect, is that, if you play a game that follows this set of procedures, then the game you are playing is Noughts & Crosses; otherwise it isn't.

Cosimo Cardellicchio has likened the abstract body of operational rules to a virus inhabiting the mental software of the gaming community (by which term I mean all the players of a given game). He also pointed out that, like any other virus, it is prone to mutation, thus causing variations and evolution of one game to another. Variable comprehension and faulty transmission are two ways in which rules change and so cause games to evolve.
Implicit rules

The operational rules of a game are known consciously, and are made explicit whenever they are addressed and verbalised. But there are also at least two types of rules that are unconsciously known or implicit. I'll call these respectively (a) foundational rules and (b) behavioural rules, and add them to the model as follows:

```
Foundational rules > operational rules < behavioural rules
  | equipment
  | play
```

Foundational rules

As Salen and Zimmerman observe, the operational rules of Snakes & Ladders include rolling dice, moving counters towards a goal 100 squares distant, and taking short cuts backwards or forwards when landing on a snake or a ladder respectively. But you could play essentially the same game by any means of randomising a number from one to six, adding it mentally to a running total, and automatically reducing or increasing your total when it reaches a certain predetermined level, the winner being the first person to total 100 exactly. They describe these as the deep, underlying mathematical rules shorn of any reference to a board, dice, counters, or snakes, ladders, and everything else that turns the game into a meaningful experience of play. Salen and Zimmerman call these the constitutive rules; but no such word appears in the Oxford English Dictionary, and I think a more sensible term is foundational rules.

Behavioural rules

Not all the rules of a game are expressible in words. Salen and Zimmerman refer to certain rules of behaviour as the "Unwritten Rules", this being the title of a substantial essay by Steven Sniderman. Sniderman also uses Noughts and Crosses to introduce the concept of unwritten rules of behaviour:
Is there a time limit between moves? Normally, we both "understand" that there is, and we both "know" that our moves should be made within a "reasonable" time, say 20 seconds. If one of us takes longer, the other starts to fidget or act bored, may even make not-so-subtle comments, and eventually threaten to quit. Without having stated it, we have accepted a tacit time limit [...] Is it a rule, or isn't it? Suppose it is my turn and, no matter what I do, you will win on your next move. Couldn't I prevent that from happening, within the rules stated, by simply refusing to play? Nothing in the rules forces me to move within a particular amount of time, so I simply do not make my next move. Haven't I followed the rules and avoided losing? And yet, if you've ever played a game, you know that this strategy is almost never employed and would be completely unacceptable. Anybody who seriously resorted to such a tactic would be [ostracized] in the future. This behavior seems to violate some fundamental but rarely stated principle of the game without any of us ever having to discuss it.

Sniderman also discusses implicit rules about whether verbal banter is acceptable, and, ranging even more widely, observes:

Since there are various "levels" of rules, "higher" rules (such as a real-world crisis) might have to take precedence over "lower" rules (such as time constraints), there must be a set of meta-rules for determining when this is appropriate. As with the other meta-rules we've looked at, there is no "final" set for ending disputes, so the question, "When is it appropriate to suspend certain rules?" cannot be given a full answer.

Not all behavioural rules are implicit, however. The more widely and seriously a game is played, the more explicit do rules of behaviour become. In the International Laws of Contract Bridge, for example, they are referred to, reasonably enough, as the Proprieties and are made quite explicit.

The written rules

Now let's consider the written rules, by which I mean any set of rules that has been formulated and recorded in writing. I asserted above that most of the world's games, especially taking variations into account, are not recorded in writing. Is this because the vast majority of the world's population is illiterate? No. When I looked into this question I was surprised to learn that nearly 80 per cent of the global population aged 15 and over is officially literate, and that the figure is expected to rise to 85 per cent by 2015. So am I wrong in my assertion about folk games and book games, or is something else going on here? I think the latter.

It must be true that some countries are more or less literate than others. But
even a largely literate population, such as that of Britain or the United States, may include a large number broadly describable as non-literate for one reason or another. For example:

- A significant proportion will be children, who may be described as pre-literate. They play games, of course, but children's games are essentially folk games rather than book games.
- Some people will be, to all intents and purposes, functionally illiterate. But it is patently obvious that being illiterate is no bar to learning and playing games.
- Probably a larger number are nowadays described as aliterate - that is, they can read and write, but avoid doing so whenever possible. So they too will pick their games up in the traditional non-literate way.
- If we describe children as pre-literate, I suppose we can also invent a category for the post-literate - that is, people whose literacy has become compromised by age or failing eyesight.

I don't know what proportion of the remaining truly literate society remains after deducting the various types of non-literates. But even if it is over 50 per cent there still remains the cultural fact that following the written rules of a game demands a peculiar degree of concentration and interest that many people would prefer to avoid if they can. The plain fact of the matter is that learning games is, by long tradition, of itself not a literate activity. After all, games-playing is a form of cultural behaviour, and people learn most of their behaviour from other people, not from books. Players are only likely to call on books and printed rule-sheets when they are passionately concerned with what they like to call the "official" rules - that is, those authorised as prescriptive and proscriptive. One consequence of this tradition is the fact that book games tend to remain relatively static over long periods of time, and are less likely to vary and evolve than most folk games.

An inevitable consequence of verbal transmission is the growth of variation in the operational rules of play. Each ad hoc exposition of the rules is unique to the person explaining the game. Variation can occur because the learner may not fully understand the rules, perhaps because certain situations that occur but rarely are not mentioned by the teacher, and the newer player will have to find some way of dealing with them when they do occur by inventing rules to cover them. In course of time, rules transmitted verbally undergo changes like those encouraged in the game of Chinese Whispers. Book rules, on the other hand, tend to retain an element of fixity. Indeed, a notorious feature of the transmission of book rules is that they often remain transmitted in a form that no longer corresponds to the way a game is actually played in real life. A classic
example is that of Brag, the English equivalent of Poker. The form of the game described in Hoyles up to the mid 1970s was one that seems to have died out by the 1790s.

**A template for written rules**

Who are the rule-writers, and how should they do it? Obvious candidates include games inventors, the authors of Hoyle-type gaming manuals and anthologies, and anthropologists who are sufficiently interested in the subject to describe hitherto unknown folk games to the wider world.

As to the order of descriptive events, I favour the following:

1. **The name of the game.** This is more than interesting: you could say it is vital. The name of a game is the portal through which you pass from the real world into that special world in which games are played. The subject of how games get their names and what they mean would make an interesting paper for a future colloquium.
2. **Its classification.** This relates it to other games and should include a brief account of its distinctive points.
3. **Authority.** State the provenance of the game, sources of authority, and authority of the person drafting the description.
4. **Number of players, and how disposed.** (For example, Solo or partnership?)
5. **Social status.** (Played by men, or women, or mixed? Regarded as childish, intellectual, disreputable?)
6. **The gaming equipment and a brief description of how it is to be manipulated.**
7. **The aim of the game.** (I never cease to be amazed at how many game books say nothing about how you win the game until they have gone into excruciating detail of how you play it. At toy fairs, when I ask the inventor or publisher of a new game what it’s all about, I usually have to interrupt them within the first half-minute to say "Yes, that’s all very well, but what are you aiming to do? How do you decide who has won?")
8. **Detailed rules of play in the normal course of events, with specifications of what you may and may not do.**
9. **Special rules governing exceptional cases and occurrences.**
10. **Penalties and corrections for irregularities.**
11. **Ending, winning, scoring, continuation (of a multi-part game).**
12. **Pay-off.** This may be money, title, prestige, or something else that you carry away with you into the real world. Just as the name of the game is the portal from which you pass out of the real world and into the play world, so the pay-off is the portal through which you pass back out of the play world and into the real world.

These are the main elements of what I consider the basic essentials of a game description, other than rules of behaviour, strategic guidelines, and any so-
called rules relating to a game's thematic, representational or allegorical aspects. (Note 5.)

The "official" rules

When you sit down to play a game, you all normally expect to be playing the same game - in other words, following the same set of operational rules. Why? Because a fundamental assumption of play is that everyone should be playing on a basis of absolute equality - as they metaphorically say, on a level playing-field. There are apparent exceptions to this, of course. For example, some games are asymmetrical, like Fox & Geese or Entropy, where equality is secured by alternating positions in a series of games. Some games, such as Chess and Go, offset inequalities by handicapping the stronger player. Either way, the object is the same: it is to provide a basic element of fairness. But again - why should a desire for equality and fairness be regarded as fundamental to the play of games?

This brings us on to perhaps the most fascinating and paradoxical characteristic of games - namely, that they are at the same time both competitive and cooperative. They are competitive in that both players want to win and will do everything in their power to do so. In fact, anyone who looks to be uninterested in winning is likely to be considered something of a spoilsport. At the same time, however, they are cooperative, in that both players will be operating within an agreed and equal set of constraints on their freedom to act.

This paradoxical behaviour appears to be a peculiarly human characteristic, as suggested by the following paragraph from Sniderman's essay The Unwritten Rules:

[We read] about a group of scientists who attempted to teach dolphins to play water polo. Although the dolphins were able to learn how to put the ball in the net (and seemed to derive pleasure from doing so), when the trainers tried to get them to stop the other team from "scoring," the dolphins launched an all-out war on the other team's players, using methods that no person steeped in the concepts of sportspeopleship would ever use. After this experience, the trainers gave up their effort, apparently concluding that their task was hopeless, that dolphins couldn't be taught to play the sport. My guess is that they assumed that all the dolphins needed to be taught were the recorded rules of water polo and the creatures would be able to play the game like adult human beings. These scientists evidently did not realize how much of our knowledge of proper game behavior precedes the learning of the statable
constraints of a particular sport.

The problem of authority arises whenever there is a disparity between the sets of operational rules encoded in competing players' minds, thus producing situations in which they disagree on the rules. If they're unable or unwilling to resolve the dispute themselves, they will have to consult an external authority. This raises the question as to what constitutes a valid external authority. I will approach this by reference to what I call levels of association.

Levels of association and authority

The smaller the degree of association among the players, the narrower the scope of authority needs to be. For example, if you're playing a card solitaire, you constitute the smallest possible degree of association of players - that is, none at all - so if you encounter a situation where you're not sure what the rule is, you don't have to consult an authority: you can simply invent a rule to cover it. This probably explains why there are so many different games of Patience, each with so many different variations.

The next level is (what I call) the ad-hoc association of two or more players playing the same game at the same table at the same time. Theoretically, they can resolve a dispute by acting corporately like a single player and agreeing on a rule to cover it. Or they can agree to acknowledge one of themselves as an authority and abide by that player's decision. This may sound a bit far-fetched, but listen to this amazing assertion from Aquarius, a 19th-century writer on card games:

[As to the] Spaniards of Europe and America [...] Their way of referring to the dealer to settle every doubtful point or dispute is very marked. A book is never mentioned. The decision of the dealer, right or wrong, settles everything or anything, without a murmur, during outplay. Thus little hitches are readily disposed of, and any game can continue. A dealer can ask advice or consult with others, but his decision is his own, and must be immediate. Players come in and leave a game with a substitute very suddenly, and agree to anything done for them. The coolness, courtesy, and skill of the Spaniard at card playing renders him in such things superior to card players of other nations.

I can quote another example of this from my own domestic circle of players. In word games, which we play a lot, one of the commonest queries encountered is whether or not a claimed word is valid. For reasons that I won't go into, no
dictionary is entirely adequate for this purpose. The house rule followed by my group is that a given word is acceptable if at least one opponent accepts it. This may not work for every group of players, but it does with mine, if only because I insist on it!

A group of people who regularly play together can build up a quite a set of "house rules" by following this sort of procedure. Alternatively, of course, they can go to a higher level of authority, such as a book by an acknowledged expert or Hoyle-figure, or the printed rules of a national or international governing body whose rules are widely accepted as authoritative.

At the next higher level comes (what might be called) the group association, which is a group of people who regularly play together, usually at more than one table at a time, and who eventually all get to know one another, such as a local club. Here it is not practicable for each table to make up its own rules, so it is more desirable for the club itself either to devise its own house rules, or to declare itself subject to those of a national or international corporate authority.

The highest possible level is the total association of the whole gaming community - that is, all the players of a given game. Games and tournaments take place beyond the boundaries of a single club, and players will regularly encounter opponents they have never met before. This level of association requires the highest level of authority - such as Japanese Go Association, or the German Skat Federation, or the Fédération Internationale des Échecs - which may be exercised through the recognition of qualified referees, umpires and arbiters.

Despite all these external authorities, it is important to remember that they are authoritative only to the extent that the players agree to observe them. No book should be held to bear an intrinsic authority beyond that of the author's own competence and experience, nor should the rules drawn up by any official body be regarded as the official rules of the game in question, but only as the official rules of the body concerned. For example, the rule about having to play a certain number of moves within a given period of time is not an intrinsic rule of Chess. And such rules are only authoritative to the extent that players agree to abide by them: they can only be invested with an authority which, by their very submission to it, actually derives from that of the players themselves. By making reference to an external authority they are, in effect, harmoniously and
cooperatively legitimising those rules as their own.
As James Carse puts it:

The agreement of the players to the applicable rules constitutes the ultimate validation of those rules. [...] There are no rules that require us to obey rules. If there were, there would have to be a rule for those rules, and so on.

Given that players at the lowest levels of association can assert their own authority and regulate themselves, it is remarkable how often people contact writers of games book, such as me, with a specific query about procedure and ask how it is covered by the official rules. Often enough I reply that there are no official rules for the game in question but that I can either make the following recommendation from personal experience or refer them to someone else whom I think better qualified to pass an opinion. Why, in short, do regular players prefer to seek an external authority?

The obvious answer is to settle queries and irregularities without wasting time; but that is not all. As Steven Sniderman writes:

A game is supposed to be for fun, and, generally speaking, playing the game itself is more fun than playing the meta-game of arguing. [It] is supposed to be for camaraderie, and arguing about the rules leads to antagonism rather than a spirit of friendly competition. Doing things as others have done them in the past allows us to feel connected to our ancestors, our culture, and our traditions; [and] following the rules that others follow allows us to compare ourselves to a wide spectrum of players, not just our immediate opponent(s). Individual players vary in their degree of authoritarianism. There will always be conservatives who look up to a higher authority and are resistant to change, and radicals who prefer to adopt a creative and variable approach to the establishment of norms.

This brings us to the subject of...

**Authority and attitude**

Salen and Zimmerman categorise players according to their "lusory attitude" - that is, their attitude to the rules and practice of play - and present this in tabular form, which I here take the liberty of abbreviating and slightly rewording:

<table>
<thead>
<tr>
<th>Type of player</th>
<th>Lusory attitude</th>
<th>Relationship to rules</th>
<th>Interest in winning</th>
</tr>
</thead>
</table>

http://www.davpar.eu/gamester/rulesOK.html
<table>
<thead>
<tr>
<th>Standard</th>
<th>normal</th>
<th>normal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dedicated</td>
<td>over-zealous</td>
<td>intense</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unsportsmanlike</td>
<td>sometimes over-zealous, sometimes feigned</td>
<td>intense</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cheat</td>
<td>feigned</td>
<td>intense</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spoilsport</td>
<td>absent</td>
<td>absent</td>
</tr>
</tbody>
</table>

Enlightening as this is, I can't help thinking it to be significantly incomplete, for it omits the very category to which the authors themselves (presumably), and I (certainly), belong! To it must surely be added that of the games technician - that is, one who is mainly interested in the technical workings of a game, such as an inventor, critic, or experimental player, and who can often be too objective to feel more than an academic interest in winning. Let's try it out:

<table>
<thead>
<tr>
<th>Technical</th>
<th>inquisitive, exploratory, creative, experimental</th>
<th>insubordinate, openly critical</th>
<th>low</th>
</tr>
</thead>
</table>

This addition brings me to the final part of the model...

I said earlier that a major cause of change is variable comprehension and faulty transmission, and I think there is an interesting paper to be written on the subject of how games change when they pass from country to country or culture to culture. These might be described as accidental causes. But game rules are just as often modified deliberately. Those I describe as technical players - the inventors and critics - are the most obvious examples of change through critical encounter, or feedback from the play (application) or the board (realisation) to the rules (regulation). Inventing and developing games necessarily involves feedback: it is the *modus operandi* of the job.
Strong players of a strategy game may find it increasingly less of challenge, and seek ways of strengthening its appeal by introducing tighter constraints or more challenging opportunities. Weak players may find it appealing but altogether too much of a challenge, and seek ways of increasing its appeal by reducing its demands. And ordinary everyday players may introduce changes just for the sake of novelty, or to fill what is perceived as a gap - such as the literal gap represented by the Free Parking space in Monopoly. Indeed, the evolution of Monopoly from Lizzie Magie's original Landlord's Game furnishes an excellent example of how the inventor's original vision was virtually overturned by becoming a folk game and undergoing rapid evolution by collective modification. And ordinary players have not stopped modifying it ever since, as Spartaco Albertarelli documents in 1000 Ways of playing Monopoly.

I'd like to think the main reason why I don't play games very successfully is that I'm too much of a technician. It's a good thing there are not too many technicians about, otherwise few games would ever get played properly. In order to keep myself in check I need to remind myself of the following cautionary observations by Steven Sniderman:

> Doing things as others have done them in the past allows us to feel connected to our ancestors, our culture, and our traditions. Players are supposed to be good sports [...], and rule challengers are perceived as poor sports or even spoilsports. [...] A set of rules that has been tested is better than one that has not, so if it [ain't broke] don't fix it.

**Conclusion**

That said, we can finish up with an expanded diagram that also includes explicit rules of behaviour ("Laws"), advisory rules (rules of strategy), and the role of feedback in continuing modification of the operational rules - thus:
References

3. Cardelliccio, Cosimo: *A discussion on the concept of evolution in games*, a paper read at Board Games Studies VI (Marburg, April 2003). (Return)
6. Aquarius (= Louis d'Aguilar Jackson): *Spanish Games at Cards*, London 1890, pp.75-6). (Return)
Don’t follow these rules!
A Primer for Playtesting
Nathalie Pozzi and Eric Zimmerman

During our 2012 residency at the University of the Arts Berlin, we spent the summer with Graduate Fellows playtesting projects from theater, architecture, sound installation, games, philosophy, and more. This essay outlines the playtesting methodology we used by suggesting possible “rules” for structuring your own playtests.

What is playtesting?
Playtesting is a methodology borrowed from game design where unfinished projects are tested on an audience. A playtest happens when people come together to try out a work in progress. The next steps for changing the project are based on the results of the playtest.

Playtesting is also an attitude towards the creative process, an approach that emphasizes problem-solving through iteration and collaboration with members of your audience.

When is playtesting useful?
Playtesting can help develop any kind of work that involves interaction between a created experience and a participatory audience. Although many of the ideas of playtesting come from game design, they can be applied in any field.

What does playtesting look like?
Playtesting can look like any number of things. At the University of the Arts, we met as a group on a regular basis and shared works in progress. We would spend about 30-60 minutes interacting with and discussing one project – perhaps in a studio space, perhaps outdoors in a park or on the street - and then move on to the next.

Isn’t playtesting the same as user testing / editing / rehearsal / critique?
Yes and no. Playtesting is not discipline-specific and versions of it can be found in many practices. The style of playtesting we outline here comes from game design and is particularly relevant for projects that involve direct audience interaction.
THE “RULES”

.....before you playtest

A. Playtest before you think you are ready
You always playtest a work in progress, not a finished design. That means you
should playtest as early as you possibly can – usually much earlier than you think
you should. It is much much better to playtest your ugly prototype than to wait
and playtest a more polished project. A playtest is not a presentation. If you feel
ready and comfortable to present and playtest your design, you have waited too
long – it is probably too late to make substantial changes. Train yourself to
overcome your discomfort and playtest as early in the process as you possibly
can.
Is it too early for you to playtest? If the answer is yes, then playtest anyway.

B. Strategize for early playtesting
Figure out how to create a working prototype far in advance of any final deadline.
This is often a question of tactical implementation. Can you make a paper
prototype of a digital project? Can you scale down a work meant for 100
participants to something you can playtest with a dozen? Rather than plan your
entire project in advance, focus instead on what is needed to enable the next
playtest.
Simplify your project so that you can playtest today.

C. Know why you are playtesting
Enter into every playtest with a concrete idea about what you want to learn and
what questions you hope the playtest will answer. Narrowing what you want the
playtest to investigate can help you simplify your project and playtest sooner.
Generating research questions in advance will also help you structure the
playtest itself. If you are doing things right, your playtest will raise issues and
questions that you did not anticipate. However, you should still go into every
playtest with a clear agenda.
What is the one key question that you want your playtest to answer?

D. Prepare variations
Go into a playtest with different versions of your project to try out. This allows you
to make the most out of the playtest session and it also helps you to improvise
and try out new ideas during the playtest. Variations might mean different sets of
game rules to play, software settings to cycle through, or contexts for a
performance. Variations give you options if something breaks down, and they let
you do comparisons to see which variation works best. One tip: change as little
as possible each time (only one element) so that you can understand better the
exact effects of your change.
What can you change to try out different variations of your project?

E. Be grateful to your playtesters
Whoever is playtesting your project is doing you a big favor. They are donating their time and attention for the sole purpose of helping you with your unfinished project. Playtesting is hard. But no matter how much stress and uncertainty you might have about the project, try and maintain a feeling of gratitude towards your playtesters. Be happy they are there and be sure to let them know how thankful you are for their time.
Take a deep breath and say thanks.

F. Design the learning experience
Remember to design the way that people will learn about your project. If you are creating a complicated interactive system, the experience of learning how to understand and interact with the system is an important part of the overall design problem.
Does your playtest address the learning process?

G. Blame yourself, not your playtesters
Remember to warn your playtesters that they will be interacting with an unfinished, rough version of what will at some later point be a smoother experience. Be sure to tell them that if they are frustrated or confused, it is not their fault – it is your fault for not designing a better experience for them. It’s OK for them to be confused – after all, the most valuable part of the playtest is not what they do understand, but what they don’t.
Never make your playtesters feel foolish.

H. Know your testers
What do you need to know about your playtesters before the playtest begins? If you are meeting them for the first time or don’t know them very well, talk with each person and take notes that will help put their reaction to your project in context. Playtesters come in many varieties. For example, the learning curve of a hardcore gamer is very different than someone without deep experience in a particular game genre.
Do you know who your playtesters are?

I. Don’t explain
Put the project ahead of the theory. Resist the temptation to explain the ideas and intentions behind your project to your playtesters. Instead, let them interact with the LEAST possible explanation from you in advance. By explaining your ideas beforehand, you are ruining the chance to see the authentic reactions that your project provokes. It is hard to hold back and not explain. But by forcing your project to carry your ideas (rather than your explanation), you are challenging your work to be better.
Is it possible to not say anything before the playtest starts?

**J. Take notes**
In game design, we often prepare a sheet of paper for each playtester, with questions written out and room to take notes. The notes page is structured to facilitate what you need to know BEFORE, DURING, and AFTER each playtest. During a discussion, taking notes will help to elicit better feedback – if your testers see you taking notes they will be more likely to give you detailed and thoughtful answers.  
*Prepare a notes sheet and use it. It is worth the extra effort.*

**....during a playtest**

**K. Be selfish**
The purpose of your playtest is not for your playtesters to have fun. It is for you to learn what does and does not work about your project. If you try too hard to give playtesters a good time, you will lose the opportunity to get the hard truth from them. Don’t be afraid to show your playtesters something broken and half-finished. That is in fact the entire premise of the playtest.  
*Don’t worry about being entertaining.*

**L. Encourage your playtesters to talk aloud**
If it is possible for your project, ask your playtesters to talk out loud about their thoughts and feelings as they interact with your work. A “think-aloud” playtester can give you valuable insight into how they are perceiving and interpreting the details of your project. Let your playtesters tell you why they are doing what they are doing and what they think is happening as a result. This may require that you periodically remind them to vocalize.  
*Don’t be shy about reminding your playtesters to think aloud.*

**M. Notice everything**
Prepare on your notes sheet the categories of the main things you want to observe, such as when players seemed frustrated, what make them laugh, or how many times they tried and failed before they gave up. Keep track of how long it took to run the playtest, which variations your testers preferred, and any other important information. Try to take notes on everything that you can – otherwise, you will be at the mercy of your selective memory, which will cast everything in the best possible light.  
*Are you noticing everything – or just what you want to see?*
N. Shut up
While you are observing the playtest, say as little as possible. You will feel an overwhelming urge to help out your playtesters, to tell them what to do and what they are doing wrong. But you must do everything you can to not interfere. Their mistakes and misunderstandings are extremely useful: you must let them explore the project on their own. If they are completely confused, step in and assist them, but in general you should do everything you can to shut up. If you tell them what to do, you lose the main purpose of the playtest, which is to see how OTHER people react to your project. Learning to shut up during a playtest requires discipline.

*Can you shut up – not a just little, but really, completely, shut up?*

O. See the big picture
As your playtesters interact with your project, remember to not just focus on the workings of your designed system. Try to see the human element at play. What are the emotional responses of your playtesters, what is their body language, how are they interacting with each other? Seeing the bigger picture can help you understand when your audience is engaged and when they are bored. It is easy to focus too much on what you designed, rather than on the effect it is having.

*Stay focused on the impact of the project, not just the project itself.*

P. Don’t be afraid of data
One way to get objective about your playtest is to record data and put it in a spreadsheet. Every project has data to collect: At what moments did everyone fall silent? How many steps did each participant take as they walked through the space? If you are working in software, the program can record important user input, such as time spent in different areas of the experience. Otherwise, just remember to record the data in your notes. Too much data can be overwhelming to interpret, but tracking the right data can be incredibly valuable.

*What is the data that will answer your key questions?*

Q. Answer a question with a question
When playtesters ask you how something works, or what something means, it is probably because they are confused. Rather than explain it to them, you can answer their inquiry with a question of your own. Don’t tell them what the blue button does – instead, ask them what they think it does, or even better, what they think it SHOULD do. It’s more important to get them to speculate about your project than for you to explain it to them. Their opinions are more valuable than yours.

*Every time a playtester asks you something, ask them something back.*

R. Hunger for failure
One of the attitudes that helps with playtesting is to yeart for your project to fail. Of course we all want successful results, but unsuccessful moments are much
more useful. If you are only looking for the successes, you will remember the
smiles and laughter and think that your project is in perfect shape (we call this the
“happy face syndrome”). But you need to cultivate a desperate hunger to focus
on what is not working properly. Otherwise, your project will never get better.

Are you enjoying the successful moments too much and ignoring the failures?

...after a playtest

S. Discuss what happened
After the playtest, talk about the experience with your playtesters. Use your notes
sheet to structure the conversation. Begin with very specific questions, such as
what was most difficult for them to understand about the project, or why they
reacted to a particular aspect of the design. Finish with more general questions,
such as what they liked best about the experience or what they would change to
make it better.

The more concrete your questions, the more useful answers you will get.

T. Put feedback into context
It can be useful to distinguish between expert and non-expert testers. Experts are
familiar with what it means to make a project like yours. Non-experts aren’t.
When getting critical feedback from non-experts, remember that they are the
patient and you are the doctor – you can take their suggestions as symptoms of
what is and isn’t working in the project, rather than as directions for the next
steps in your design. If someone tells you to tear down a room and make it
bigger, they are really telling you that it feels small. Rather than take their advice,
perhaps just rearrange the furniture. Don’t expect your players to understand all
of the ramifications of every suggestion they make.

Ask for feedback, but don’t take suggestions literally.

U. Collaborate with your playtesters
One of the most thrilling moments of playtesting is collaborating with your
playtesters – brainstorming with them, trying out their ideas, and seeing how the
changes impact your project. Plan your playtest session so that you have time to
experiment with new ideas as they emerge through the playtesting itself. They
are seeing the project with fresh eyes and so their ideas are often better than
yours.

Embrace shared authorship with your playtesters.

V. The cruelly honest playtest
Playtests represent moments of truth – when your brilliant ideas may all come
crashing down. Playtests are truthful because they are a safe place to simulate
your final context. When your project is completed, you probably won’t be there
to explain away all of the problems and defend your intentions. In a playtest, you
get to cruelly see whether or not your ideas actually work in practice. Part of the playtest attitude is building up your pain tolerance and coming to enjoy the hard truth of the playtest. 
*Face the truth of your playtest, even if it hurts.*

**W. Embrace the unexpected**

Never forget that *play* is half of *playtest*. Being playful means being open to unexpected, happy accidents. Let go of the way you *want* your work to be used or interpreted. Be open to the strange new things people do with your project. Accidents are for those who are ready to take advantage of them. *

*If things don’t go as planned, you may be on to something better.*

**X. The playtest’s the thing**

The playtesting process is as important as the actual project you are making. If you can manage to get the process right, then you will find that the problems in your project begin to solve themselves. *

*Forget what you are making. Focus on how you make it.*

**Y and Z. Break these rules**

There is no single magic solution that will solve every problem you encounter. So you need to create the process that works for you. *Don’t follow these “rules.”* They are not meant to be followed – they are meant to be twisted, modified, broken, and refashioned into something new. The best playtest is the one you invent yourself.

- Nathalie & Eric
by Teneoan Woodruff

I've known Tey for a quarter century, and can say that she's one of the most pleasant people you'll meet. This masks one of the cruelest streaks I've ever seen. When it comes to playtesting, she is ruthless. Where most of us feel a need to help a confused player, or just explain one little rule they're missing, Tey will let them flounder in pain, merely noting their distress on a form. She is right to do so. From decades of experience, she knows how to extract the most useful information from a playtest group, and will share that with you now. As proof of my above statements, she will start by crushing your spirit.

What I'm about to say may offend some of you reading this essay—especially those of you who have just created a shinningly perfect game that is bound to set the gaming world on fire and make you a kabillion dollars. But, please, hear me out. Your game, and possibly your wallet, will thank you for it. So, here it goes: Your game's not as good as you think it is. At least not until you've had people who don't hold it—or you—near and dear to their hearts play it and agree with you.

How can you make your game the best it can be? One of the most important ways, and the one most often skipped by new designers, is by playtesting it. And playtesting does not end with having your friends and family members play the game you've told them how to play. Unless you plan on including yourself in every game box, that kind of playtesting has little value.

So welcome to the who, what, where, when, and why of playtesting. We'll work our way backwards through that list. By the time we're done, you'll have a better idea of how the playtesting process works and why it's really important that you do it.

Why should you playtest?
Who knows a game better than the person or people who created it? Nobody. A
game is like your child. You've created it, you've seen it through challenges, and you love it. But, also just like your children, your love can blind you to your game's faults if you're not careful. And even if your game truly is brilliant, fun, and engaging, if your rules are complicated, confusing, and wrong, your audience may never even get to the point of playing your game.

A playtest can help you learn a great deal about your game's strengths and weaknesses. When you create a game, you play it, you live with it, you work with it, you become very familiar with its ins and outs. In fact, you become so familiar with it that you often become blind to the stumbling blocks others might face when they first play it. For example, in one of the many playtests we conducted to figure out the best way to teach Magic: The Gathering to new players, we watched as players read the rule that says you should tap your land for mana to use it. We all knew what tapping for mana meant. Everyone knows what tapping for mana means, right? Right. It means touching the land card firmly with your index finger a couple of times. D'oh. It turns out a visual reference of someone turning a card to the right and getting one of the appropriate color of mana goes a long way in teaching the game term “tap.”

Playtesting is a crucial tool allowing you to step back from your game and see its flaws and strengths through new eyes—eyes of people who have never played the game before. In other words, the eyes of the consumer. Without this sort of objective playtesting, even experienced game designers can stumble on rules or gameplay elements that cause new players to give up on what is an otherwise excellent game.

**When should you playtest?**

There are several different times you should consider playtesting a new game. Each type of playtest has a different goal.

A developmental playtest is a playtest before the rules are finalized. That sort of playtest aims at understanding if the gameplay itself is what you want. These
playtests don’t worry about conveying how to play the game; instead we’re concerned only with if the game plays well. This first type of testing happens once you have game mechanics you’re happy with. The least formal way to test is to get some people you know, tell them how to play and see how it goes. You might start and stop several times, tinkering with the rules as you go. You can uncover some basic problems with rules and game mechanics through this method, but that’s about it. You’re too close to your game, and your friends and family are too close to you, to do more than this.

A **hand-taught playtest** is a better way to test how the game itself is playing. Here, you leave your friends and family behind, and recruit people in the target demographic. (Meaning the people your game is made for. And if you say “everyone,” you need to start playtesting **stat.**) Then you or someone else who knows how to play hand-teaches those people how to play. Even if you are the person teaching the game (and if you have little or no budget that will probably be the case), tell the playtesters you don’t have anything to do with the game’s creation. Why the lie? People are more comfortable giving critical feedback to someone if they aren’t worried about hurting their (meaning your) feelings.

Even if you do some good hand-taught playtesting—and I encourage you to do so—you’ll want to switch to the **blind playtest** when the game is very close to finished. A blind playtest involves a number of people in the target demographic for the game, with no association to the game or its creators, playing that game as if they had just bought it at a store (or as close to that as we can make it). You should have a set of rules laid out with graphics, a mocked-up set of game components, and mocked-up packaging to put those components in. This will get you as close as possible to the “real world” of someone buying and attempting to play your game. (There’s a more thorough type called a **double-blind playtest**, where even the person running the session doesn’t know how to play the game.)

A final form of playtest is a **focus group playtest**. A focus group is a carefully selected group of people in your target demographic. (This may involve parents of
your demographic depending on the main age of players you're going for.) A focus group can provide a lot of useful feedback on the look and feel of your product and its packaging. As with getting your rules right, getting the look and feel right will improve the odds of getting your game into the hands of the people who will love it.

Now, all these types of playtests cost money. You can spend as little as a couple pizzas and snacks, and putting your testers' names in the playtesting section of the credits. Or you can spend hundreds of thousands on complete testing including in-store product shelf-testing and the like. Since most readers of this essay are newer to game design, I won't get into how to conduct those larger-scale tests—although they give an advantage to the companies who can afford them. In the next section we'll talk about appropriate (or necessary) compensation for playtesters and other playtesting costs.

**Where should you playtest?**

From now on, we're going to concentrate on later-stage playtesting, because if you can only afford to do one type of playtest this is the type you should do. You have to see what reaction new people have to your game and if they can play it from the materials you're giving them.

So, where does this playtesting take place? For most games, the ideal setting for playtests is either the consumer's home or a focus group facility. People's homes are particularly good for family games or party games. Focus group facilities are great because they are neutral, the people who work at these places can recruit the playtesters using your criteria, and they provide a place for you to watch the playtests without interfering (behind a piece of one-way glass). The facility will also tape the sessions so you have records to look back on.

Of course, focus group facilities and recruitment grows costly quickly. You'll need to budget something between $5,000 and $50,000 to go this route. Getting professionals to neutrally recruit your participants and having neutral grounds to watch the playtest can reveal numerous flaws and strengths of your game, rules, and
packaging. Finding out a crucial mistake in a focus group facility can make the difference between success and failure. You might discover that one of the main terms you are using in your game is offensive to a certain segment of the populace—something we saw happen at Wizards with the TCG originally named 

To use a focus group facility and their recruiters, find a couple of places in your area and ask for references—particularly ones involving games and entertainment. Find out their menu of costs and determine which services you want to use. (For example, you could just choose to use their facility, or use both facility and recruiting.) For running the playtest, you should get facilitators who are familiar with the game and can ask questions after the testers play to elicit what they liked, disliked, and found hard. You should not be that person. You should be behind the mirror concentrating on the playtest. Companies usually use market researchers for this position, although focus group facilities can provide someone as well.

If you can't afford a focus group facility or your game is best tested in a home or other local environment, you'll have to go into the field to test. You can hire a market research firm to recruit testers, or you can try to do so on your own. If you recruit, get away from folks who know you. Place an ad on a college information board, ask around at after-school activity places, or get friends to post on Facebook to their friends (as long as they don't reveal they're friends with the person who made the game). The exact method depends on your pocket and who you need as playtesters.

Once you recruit playtesters and get to the location, if you are involved in the playtesting, make sure you tell the testers you had nothing to do with the game design and are just here to test it. That will help the testers feel freer to share negative as well as positive feedback.

What should you playtest?

Alright, let's say you're at some nice family's home. They've invited you in to playtest the new family game you've been perfecting. How do you go about getting
valuable and truthful information from them?

First off, be sure to bring enough playtest materials—mocked up as close to what you plan to sell as possible—for the test. If the game comes in a box, make the box. Have a set of backup materials just in case. You’ll also want to bring materials to take notes, a way of videotaping the session, and post-test forms for everyone involved to fill out. (Usually these forms involve rating the gameplay, ease of learning, rules, and so on. Basically, anything you’re trying to find out from the playtesters should go on that form.) Finally, you’ll need a release form which gives you legal rights to use the tapes and responses for your informational purposes and swear the playtesters to secrecy until the game comes out. You can find templates for these forms online, but you may want to consider a couple of hours of legal time to draft a basic form for your company.

Hand out the release form and secrecy form, and have the playtesters sign before you begin. Children need parents to sign for them. After that is done, the facilitator hands the playtesters the box and leaves or moves away from them. The facilitator should not answer any questions about the game or game materials while the testers play—or attempt to play. And neither should you or anyone else there. This is not the time to defend your game or correct mistakes. Seeing where people stumble over rules, mechanics, and visuals is what you’re here for. So be quiet! (If you’re conducting a developmental playtest, you will tell the players how to play and correct any play mistakes, but don’t comment on the game itself.)

Once the playtest is done (either in a given time frame or until the playtesters have completed the game), pass out the post-playtest sheets. Once testers fill out this sheet, the facilitator (and possibly others) should ask testers questions about what they liked and didn’t like about the game, as well as what they found problematic and anything that helped them understand the game better. If you say places the testers played wrong, make sure to tell them the correct way to play and find out what led them to the error. Some errors are matters of omission. For example, when playtesting a new soccer TCG in England, we found that simply
telling kids to draw a hand of cards wasn't enough—that led to players picking up their entire decks! We had to be more specific—shuffle your deck, put it down, and draw seven cards off the top. But many errors came down to poor formatting, poor—or no—graphics, poor or missing examples of play, and rambling text. Nobody likes to read rules, so your goal is to get the game across correctly with as few words as possible. If you can say it with a picture, do so.

Finally, pay the playtesters their consideration (often $25-$100 each) if you are doing so. And make sure you note their names correctly for inclusion in the credits.

After the playtest is over, go back and rework the game—then playtest again. Ideally, you'll have a set of very clean playtests, where the players enjoy themselves and learn from the rules, before shipping the game.

When testing, try to have three to twelve different groups play your game. Very often, you will find one or two groups that do something or play in a way that's very different from most people. Having a larger pool of playtesters helps you find out more consistently where the current strengths and weaknesses of your game lay.

Who should you playtest with?

You should playtest with people in the target demographic of your game. Is your game aimed at 6- to 8-year-old boys? Then get them in for playtests. Do you have an adult party game? Have a playtest with friends in one of their homes and see how it goes. Does your game contain an electronic device? Then be sure you have the devices you need with the game loaded for testers. (If the technical elements of the game aren't ready yet then you'll want a series of "screen" pages. The testers show the facilitator which "button" they are pushing and the facilitator hands out the next "screen").

If you don't playtest with the people your game is targeted at, the playtests won't help much. And remember, if you're playtesting with children, you'll want to get the parents' feedback too—after all, they'll probably be the ones buying the game.
So, you have your demographics and you've got materials for the testers ready to go. Time to recruit the testing troops! You can recruit from places your ideal demographic are likely to be. If girls 7 to 10 is your target, you may want to approach after school activity locations, sports games, the library, and the like. For college students who like sports, post information near the sports parts of the campus, try to get in their newspaper or online information page—and be sure to tell them who matches your criteria and how much you will pay them right off the start.

Getting playtesters for your game wouldn't be too difficult if all you needed was warm bodies. The issue is getting testers in the proper demographic who are outgoing, articulate, and analytically minded enough to express any problems they are having with game... Yeah, all that's true. But going beyond the niceties, you really have to determine if someone is open to directly criticizing the work—as harshly as they feel it merits. Also, you'll probably want some players who get pleasure from breaking the game. They play by searching for loopholes and amorphous rules they can use to their advantage—often pissing off the other players—but just as often showing you some major problems you need to fix before debuting your game.

In order to find a solid array of playtesters who have the characteristics you want, you'll have to conduct demographic surveys. These surveys ask basic information about the person, but for a game test, you'll likely want to know things like: What games do you play? How often do you play? What game is your favorite and why? What makes you pick up a game at a store? What can make you give up on a game?

In your post-playtest survey, in addition to quizzing the testers about specific aspects of your game and your rules, ask the testers to name their favorite games—again. Also, ask how likely they are to purchase the game (1/2/3/4/5 scale).

I could write a book on the ins and outs of playtesting, but probably the most important thing to remember is to stay neutral. If the testers know it's your game
or if you sit across from them rolling your eyes or making verbal comments, your data is tainted. And that won't do your “baby” any good when it's time to kick the game out on its own in the real world!

**Teeunwyn Woodruff** is a game, puzzle, and events designer in Sammamish, Washington. At Wizards of the Coast, Teeunwyn worked on games like Dungeons & Dragons, *Magic: The Gathering*, *Pokémon*, *Harrow*, *Betrayal at House on the Hill*, and *Duelmasters*. Her puzzles appear in magazines such as *Games* and *Wired*, as well as in alternate reality games. As creative director for Lone Shark Games, she has created games and immersive events for companies such as *Microsoft*, *Sony*, *Lucasfilm*, *Turbine*, *ArenaNet*, and *Southpeak Interactive*. She can’t even count the number of hours she’s spent watching playtests of all kinds.