6

On the Nature and Evaluation of Fidelity in Virtual Environments

Thomas A. Stoffregen*
University of Minnesota

Benoit G. Bardy
University of Paris

L. J. Smart
Miami University

Randy Pagulayan
Microsoft Game Studios

In this chapter, we discuss some issues relating to the evaluation of simulations. We concentrate on simulations that involve motion of the user, but our analysis may apply to simulation in general. We argue that in the strict sense of duplication, simulation is essentially impossible: With rare exceptions, sensory stimulation in a simulator cannot be identical to sensory stimulation that is available in the system being simulated. We suggest that these differences in stimulation provide information to the user about the nature of the simulation, as such. If this information

*Correspondence regarding this chapter should be addressed to Thomas A. Stoffregen, School of Kinesiology, University of Minnesota, 141A Mariucci Anena, 1901 4th St. S.E., Minneapolis, MN 55414.

is picked up, then it may be rare for users to perceive a simulation to be "the real thing." Researchers and developers commonly define the fidelity of simulations in terms of their tendency to give rise to illusory subjective experience in users. However, if simulation is perceived as such (that is, if perception is accurate rather than illusory), then there may be little practical value in metrics that rely on an illusory perception of reality. Metrics that assume that perception in simulations is illusory may be counterproductive if they focus the attention of researchers and developers on aspects of the simulation that are not directly relevant to issues of fidelity. We suggest that in many application domains, more useful metrics of fidelity can be developed from measurements of performance, rather than subjective experience.

Our use of *simulation* includes virtual environments (VEs) and virtual reality (VR). We discuss relations between simulations and the events or systems that they are intended to mimic. We begin by describing a scenario that we believe brings into focus issues that are central to the evaluation of simulator fidelity.

A Scenario: Simulation or Reality?

Imagine this. You awaken from a deep sleep to find yourself wearing a pilot's flight suit and helmet. You are strapped into a flight seat in a cockpit. The windscreen shows that you are at cruising altitude, flying through clouds. Quick, unpredictable motions of the cockpit indicate the presence of turbulence. Somehow in the night you have been spirited out of your bed and put into this situation. What, exactly, is this situation? It may be that you are in flight, that your sleeping body was inserted into the cockpit of an aircraft that then taxied, took off, and climbed under remote control. However, another possibility is that you are in a high-fidelity flight simulator; perhaps a 6-*DF* motion base in a visual dome, with 3-D sound imaging.

These two situations differ greatly in their meaning, that is, in their consequences for your behavior. If you are in flight, then you need to control the aircraft (or at least figure out how to turn on the autopilot), and a mistake may lead to a fatal crash. If you are in a simulator, then it does not much matter what you do (Shapiro & McDonald, 1992, p. 102, note that "real events sometimes demand action, while obviously fictional events can usually be enjoyed vicariously"). The difference between the flight and the simulator, then, is highly consequential; in this case it may be a matter of life and death.[1]

Will you be able to determine whether you are in an aircraft or a simulator?[2] If so, how? Will you be able to execute adaptive behaviors? If so, on the basis of what

---

[1]From the perspective of the ecological approach to perception and action (Gibson, 1979/1986), we would say that the simulator and the simulated have different affordances for behavior.

[2]This is the simulation version of the Turing Test: Can a user distinguish between a computer-generated reality and the real thing?

information? We see these questions as being the essence of the issue of simulator fidelity. Experiments at the level of our scenario would be required in order to determine the extent to which simulations are (or can be) differentiated from the systems that they simulate. Such experiments are very difficult to conduct, as they would depend on the absence, not just of informed consent but of any consent at all. However, it is only in such a situation that we could be certain that the subject's perceptions and behavior were not influenced by prior knowledge and expectations.

SIMULATION AND NATURAL LAW

In this section, we discuss relations between physical events and patterns of ambient energy that constitute potential sensory stimulation. We describe the hypothesis that physical reality structures ambient energy fields in such a way that the structure of these fields specifies, that is, is uniquely related to, the underlying reality. We then suggest that among real things are simulations (e.g., simulation devices and software). An implication of this is that simulations structure ambient energy fields differently than do the simulated systems and events, that is, that simulation is specified as such. Some aspects of our discussion may be obvious to many readers. However, we believe that the issues raised have important implications for the development of vehicular simulations, and that these implications have rarely been addressed in the simulation literature.

In this section, we also consider only patterns in energy (potential sensory stimulation), not the pickup of information by living systems. Accordingly, our discussion does not address psychological or physiological issues. It is logically prior to such issues. The activity and results of perception, and technological strategies that are designed to influence psychological processes (such as washout algorithms; Nahon & Reid, 1990), will be addressed in later sections.

Reality Is Specified

The spatiotemporal structure of light, sound, and other forms of ambient energy are influenced by events and situations in accordance with natural law (e.g., laws of the generation and propagation of optical, acoustic, and mechanical energy). Different laws apply to different forms of energy. For example, the differential absorption of optical energy at different frequencies is governed by certain properties of objects (e.g., their surface texture and pigmentation), whereas the absorption of acoustic energy at different frequencies is governed by other properties of objects (e.g., their rigidity and density). In addition, physical motion produces changes in ambient energy arrays, such as optical flow (Gibson, 1979/1986). The lawful relations that exist between reality and ambient energy arrays gives rise to the hypothesis of *specificity*. This is the idea that there exists a unique (that is, 1:1) relation between any given physical reality and the energy patterns to which it gives rise (Gibson, 1979/1986; Stoffregen & Bardy, 2001). If such a unique, lawful relation exists, then the energy patterns can be said to specify the corresponding reality. We stress that this is not a psychological hypothesis (Runeson & Vedeler, 1993). Specification exists prior to and independent of any psychological activity. The specificity hypothesis does not mandate any particular perceptual process, but it does make it possible for perception to be direct and veridical.

Simulation Is Specified, as Such

There are many differences between simulations and the things that they simulate, between the *simulator* and the *simulated*. After all, the purpose of simulation is to reproduce some of the characteristics of a system or situation (e.g., sensory stimulation and constraints on behavior) without reproducing others (e.g., the expense and/or danger of operating the actual system). Exactly which characteristics of a simulated system can be reproduced? Are there some characteristics that cannot be reproduced? In particular, can a simulator faithfully recreate the sensory stimulation that occurs in the simulated system?

A simulator is a real device that is different from the device or situation that it simulates. Both simulations and the systems that they imitate influence the structure of ambient energy arrays in accordance with physical law. If sensory stimulation produced by a simulator were exactly identical to that produced by the simulated system, then the user would have no perceptual basis for differentiating the simulator from the simulated. We refer to this as a state of *stimulus fidelity*. Researchers have not claimed that stimulus fidelity exists in current simulations or that it will exist in future ones. In fact, there is widespread agreement that technological improvement will not lead to 100% accuracy (Gibson, 1971; Hochberg, 1986; Stappers, Overbeeke, & Gaver, 2002), implying that the energy arrays in the simulator will never be identical to those in the simulated. We use *nonidentity* to refer to differences between the simulator and the simulated in the structure of ambient energy arrays.[3] Do nonidentities contain useful information, or are they noise? Are they picked up by users, or are they ignored? It is common to interpret non-identities as sources of ambiguity (noise), that is, as increasing the observer's uncertainty about reality (e.g., Edgar & Bex, 1995; Hochberg, 1986). An alternative interpretation, which we endorse, is that non-identities specify the

[3]It is important to note that our use of *nonidentity* is not equivalent to the concept of sensory conflict. The latter is an interpretation of discrepancies that exist in sensory stimulation (Stoffregen & Riccio, 1991). The discrepancies that are interpreted as sensory conflict exist in the sensory stimulation associated with a given system (either simulator or simulated). The nonidentities discussed in this chapter exist between simulation and simulated systems. Because we do not accept the sensory conflict interpretation (Riccio & Stoffregen, 1991), we use *discrepancy* to refer to intermodal nonidentities that exist within a given system.

simulations, as such, and that this information is picked up by users (Stoffregen, 1997; cf. J. J. Gibson, 1971).[4]

What follows is a partial list of ways in which the structure of energy in simulations is nonidentical with structure of energy in the simulated systems. Our contention is that physical law places severe and irremediable constraints on the conditions in which a simulator can give rise to patterns of energy that are identical to those created by the simulated system.

The Optic and Acoustic Arrays

There are a wide variety of differences between optic and acoustic arrays generated by simulated events and those generated by simulations of those events (for a detailed list of examples in optics, see Hochberg, 1986; for a related discussion about audition, see Gilkey & Anderson, 1997). Many of these will be remediated through technological development, but some may prove to be fundamental, such that they cannot be eliminated (for examples and discussion, see Gibson, 1971; Edgar & Bex, 1995; Hochberg, 1986; Wann, Rushton, & Mon-Williams, 1995).[5]

Mechanical and Inertial Arrays

The structure of haptic and inertial arrays in any fixed-based simulator will be identical to haptic and inertial structure in a simulated vehicle when the simulated motion is in a straight line at constant velocity. For all other motions, haptic and inertial structure in the simulator will differ from that in the simulated vehicle (cf. Stoffregen & Riccio, 1991). Rather than being subtle (such that they might be below detection threshold), these differences will often be of great magnitude.

Inertial displacement of the body always structures stimulation of the vestibular and somatosensory systems (and usually of the visual and auditory systems). Thus, in order to simulate the totality of (multimodal) inputs that result from inertial motion, it is necessary to control the structure of a variety of different forms of

---

[4]Our position contrasts with that of Michaels and Beek (1995), who argued that there can be patterns of stimulation that are not specific to the physical events that cause them. As examples of their position, Michaels and Beek (p. 274) included "any information created by artifice (e.g., a hologram or computer simulation)." Michaels and Beek appear to be asserting that simulation is not specified as such.

[5]It should be possible to control for the possibility that simulation, as such, is specified by artifacts of optical or acoustic display technologies. This could be done by creating a vehicle in which the operator could not see or hear the outside world directly, but had access to it only through electronic media (e.g., video and radio) In such a vehicle, many distortions of audiovisual imaging should be the same as those found in vehicle simulators. Subjects would operate this experimental vehicle and would also operate a state-of-the-art full-motion simulator, with identical optical and acoustic display technology (in the vehicle the content of the displays would be "real," whereas in the simulator the content of the displays would be computer generated). The dependent variable would be whether users could differentiate the simulator from the vehicle. If so, this would suggest that specification of the simulator, as such, was not limited to artifacts arising from audiovisual image generation and display technology.

energy, not only the optic array, but also the acoustic array, the haptic array, the gravito-inertial array, and so on. This is true both for motion-base (inertial) and for fixed-base (noninertial) simulations. We will discuss these separately.

Consider an inertially stationary observer seated at the center of a rotating drum (e.g., Brandt, Dichgins, & Koenig, 1973). During constant-velocity angular rotation, optical and inertial stimulation might be identical to that created by inertial rotation of the observer within a stationary drum. Similarly, during constant velocity linear motion of a "moving room" (Lishman & Lee, 1973; Stoffregen, 1985), an inertially stationary observer might experience optical and inertial stimulation that was identical to what would occur if the observer moved at constant linear velocity within a stationary room. In these situations, observers have no basis for perceptual differentiation of the simulator (optical motion with inertial stasis) from the simulated (inertial motion with optical stasis). But the stimulus fidelity that exists with these "simulations" ceases to exist if the simulated event includes any inertial motion, such as changes in velocity of the observer, the environment, or both. The optical display may be faithful to the optical consequences of the simulated event, but stimulation of the vestibular and somatosensory systems will not be.

These examples show that while it is possible to present the optical consequences of inertial motion, such optical depictions will not produce the changes in stimulation of haptic and vestibular systems that accompany inertial motion (Riccio, 1995). In the scenario presented at the beginning of this chapter, changes in direction or velocity of motion will produce nonidentities (in stimulation) in the simulator relative to the real aircraft. Because the scenario is likely to include substantial motion, these nonidentities will be large. For example, in flight, a climb gives rise to large and sustained increases in the magnitude and direction of the gravito-inertial force vector. In a fixed-base flight simulator, these changes will be wholly absent. In a motion-base simulator, a simulated climb can produce changes in the inertial array, but these will tend to be of lower magnitude, in a different direction, and of much shorter duration than those found in-flight.

In general, vehicles have a much greater range of motion than do simulator motion bases. Many aircraft can execute 360-deg rolls and turns, pulling more than 8 g acceleration, with very high rates of linear and angular motion. By contrast, simulator motion bases are greatly restricted. Consider the Link 6-DOF "six-post" motion platform, one of the more sophisticated systems currently in use. This system has 170-cm linear and 50° angular excursion, with peak velocity of 60 cm/sec (20 deg/sec), and peak acceleration of 0.8 g (linear) and less than 60 deg/sec (angular). Unlike an aircraft, these capabilities are independent and cannot be achieved simultaneously. The capabilities of the aircraft exceed those of the motion base in a variety of dimensions. Similarly, there are gross deviations from stimulus fidelity in any device that simulates walking (e.g., a head-mounted optical display, in which there is no inertial translation of the body) or in virtual reality "scooters" (Smets, Stappers, Overbeeke, & Mast, 1995).

Designers of flight simulators sometimes attempt to minimize conscious awareness of the limitations of motion bases by employing washout algorithms (Ellis, 1991; Nahon & Reid, 1990). The general strategy is to reach the limits of the motion base in such a manner that the termination of inertial motion is below detection threshold. Although they may enjoy success as psychophysical strategies, washout algorithms cannot reduce or overcome the fundamental physical limits of motion bases. Accordingly, we would predict that users should be able to differentiate motion-base simulations from the simulated vehicles even when washout algorithms are in use. We are not aware of any direct tests of this hypothesis.

We have seen that inertial stimulation in a simulator generally is different from inertial stimulation in the simulated system, and that this is true for both motion-base and fixed-base simulations. A major consequence of this is that simulations must be inaccurate if they simulate motion relative to the inertial environment, either of the body as a whole or of its parts separately (Stoffregen & Bardy, 2001; Stoffregen & Riccio, 1991). Stimulus fidelity can occur only in the absence of such differential motions. As we noted earlier, this means that for a moving observer, stimulus fidelity in the haptic and inertial arrays can be achieved only for motion in a straight line at constant velocity, for example, straight and level flight. Thus, for any simulation that includes inertial motion, there will be nonidentities between the structure of energy arrays by the simulator and the simulated. Based on the hypothesis of specificity between physical reality and the structure of ambient energy arrays, we conclude that for any simulation that depicts inertial motion of the user, the simulation is specified as such.

Within inferential theories of perception, nonidentities between the simulator and the simulated are not problematic. In these views, it is assumed that specificity does not exist and that (as a result) perception is primarily an inferential process operating on impoverished sensory stimulation that bears only a probabalistic relation to reality. In such theories, the sensory stimulation in a simulator might be probabalistically consistent with sensory stimulation in the simulated system without being identical to it. For this reason, such theories would tend to predict that simulations can be and are truly mistaken for the simulated systems and events. By contrast, theories based on the concept of specificity would tend to predict that nonidentities between the simulator and the simulated would result in each being perceived as such. This raises the question of whether simulation, being specified as such, actually is perceived as such: Is perception in simulators veridical?

WHAT IS PERCEIVED?

It is widely believed that many types of nonidentity can be "tolerated" because of the psychophysical limitations of sensory systems and perceptual processes (e.g., Hochberg, 1986). Arguments of this kind may be plausible when applied to optical and acoustic stimulation, in which nonidentities can be relatively small in magnitude and may shrink as technology continues to develop. However, we believe they are less plausible in the context of nonidentities in the structuring of haptic and inertial arrays by simulators and the simulated, which are often of great magnitude. Given the existence of these large nonidentities, will there be a perception of reality (an illusion), and if so, under what circumstances?

Contemporary flight simulation typically uses very sophisticated display technology (e.g., fractal terrain models, multiaxis motion bases, and closed-loop control). Despite this technological sophistication, it is widely acknowledged that users easily differentiate the simulator from the simulated, detecting the simulation as such (Edgar & Bex, 1995; Hochberg, 1986; Stoffregen, 1997). Even multimillion-dollar flight simulators are still easily discriminated from "the real thing." For example, the Synthesized Immersion Research Environment (SIRE) at Wright-Patterson Air Force Base features a Silicon Graphics Onyx computer image-generation system specially modified to drive six high-resolution channels, which are projected onto a specially constructed hemispherical dome. The graphics have a resolution of 2-min arc/pixel and an update rate of 30 Hz. The system produces exceptionally robust vection. However, pilots using the facility have no difficulty discriminating it from physical flight, even for "things that we actually try hard to model as highly convincing representations," (L. Hettinger, personal communication, June 1995).

The pickup and use of information in nonidentities need not imply conscious awareness. This is suggested by the occurrence of motion sickness in simulators during maneuvers that do not produce sickness in the simulated systems (cf. Biocca, 1992). Also, adaptation to sensorimotor "rearrangements" in simulators (e.g., DiZio & Lackner, 1992) can be interpreted as evidence that the information specifying the simulation as such has been detected and used.

The examples discussed in this section are suggestive, but they are not definitive. Very few existing data relate directly to the question of whether users perceive simulations as such. In our view, this is because there has been very little research involving direct comparisons of perception in the simulator and in the simulated (most comparisons that have been done have been concerned primarily with transfer of training and not with whether there is differentiation of simulators and simulated; e.g., Moroney, Hampton, Beirs, & Kirton, 1994). In our Conclusion, we propose a program of research of this kind.

The Role of Exploration

Perceivers do not pick up all of the information that is available to them at any given time, or in any given situation. This means that perception is selective. Selective perception requires controlled motion of perceptual systems. This can include not only motion of receptor organs, but of the entire body, as when we walk toward an object to get a better view of it. We refer to the activity of perceptual selection as *exploration* (Gibson & Rader, 1979; Stoffregen, 1997). In this section, we discuss ways in which perception can be influenced through either the promotion or restriction of exploration.

Our discussion has shown that stimulus fidelity will occur only in highly limited circumstances. Given this fact, it might seem impossible for a simulator to be effective. However, simulation may be effective if users can be prevented from picking up the information that specifies the simulation as such. Conversely, simulation can be perceived as such if users are able to pick up patterns in ambient energy that specify it.

Exploration produces changes in information (in the relation between action and sensory stimulation) that can disambiguate situations. For example, circular vection that is created by a rotating visual surround ceases if the observer is permitted to execute head movements (Dichgans & Brandt, 1978; cf. Prothero, Parker, Furness, & Wells, 1995, p. 362). The activity of moving the head produces changes in the stimulation of the visual, haptic, and vestibular systems. The changes that occur with the rotating visual surround differ dramatically from the changes that occur when head movements are executed during inertial body rotation. For example, head movements with inertial rotation will lead to coriolis forces, which are not present when rotation is solely optical. This difference specifies the actual motion; that is, it specifies whether the motion is optical or inertial.[6]

We have argued that stimulus fidelity exists only when the simulated motions are of constant velocity. Simulations can be perceived as such if the user has and exploits opportunities to explore beyond this limited situation, that is, if the user can alter the simulated velocity (cf. Mark, Balliet, Craver, Douglas, & Fox, 1990). We regard this as being logically identical to the disambiguating effect of head movements in circular vection. The use of bite bars and chin rests in experiments on circular vection can be interpreted as means to restrict perceptual exploration (cf. Mark et al., 1990). This is consistent with the fact that many simulations produce subjective realism only if exploration is limited through physical restraint (e.g., Hochberg, 1986). If it is important that the user perceive the simulator to be the simulated system, designers must develop ways to limit users' exploratory activities. That vection can be so easily reduced or destroyed by simple head movements illustrates the power of exploration in perception, and is consistent with our hypothesis that stimulus fidelity exists only when a simulation does not attempt to depict inertial motion. Further research is needed to determine the kinds of exploratory activity that permit differentiation of simulators from the simulated.

Restrictions on exploration (such as bite bars and chin rests) may themselves alter sensory stimulation in ways that specify the fact that the user is not in the simulated situation. The use of a bite bar, for example, causes changes in the structure of ambient energy that specify that the user is not in an aircraft, automobile, or other

---

[6]This should be true also of head movements in a moving room, and so raises the question of why, if the moving room is specified as such, people sway in it. It may be that they perceive it quite accurately, but sway so as to maintain stable looking (Stoffregen, Smart, Bardy, & Pagulayan, 1999).

vehicle. A related example is restrictions on the acceleration and excursion of a simulation that are mandated by the limited range of its motion base. Such restrictions impose limits on the motions that the user can execute (and correspondingly limit the training utility of the simulation). Because these limits differ from those found in the simulated system, they provide information about the simulation.

Our earlier analysis of specification leads us to conclude that in almost all cases, simulation is specified in patterns of ambient energy that are available to perceivers. This provides a logical basis for perceiving the simulation as such. Such a percept would be accurate and not an illusion. There is little experimental evidence that relates directly to the issue of whether users actually differentiate simulations from simulated systems. We have reviewed anecdotal reports and some indirect evidence that suggest that such differentiation does take place in at least some situations. In the next section, we consider implications of this possibility for the evaluation of simulator fidelity.

IMPLICATIONS FOR EVALUATING SIMULATOR FIDELITY

What are the appropriate criteria for evaluating the effectiveness of simulations? This is an area of considerable uncertainty: "A major concern relating to the use of simulation for training stems from the difficulty of determining, in many cases, the simulation's adequacy for training purposes" (Nickerson, 1992, p. 155). In terms of evaluation, we equate effectiveness with fidelity (that is, with the faithfulness of the user's behavior in the simulator to their behavior in the simulated). Metrics for fidelity fall into two classes: fidelity of subjective experience and fidelity of performance. Following Riccio (1995) we refer to these as *experiential fidelity* and *action fidelity*, respectively. Riccio (1995, p. 136), noted that "there is a lack of general agreement about the criteria for fidelity of flight simulators. This makes it difficult to resolve controversies about the sufficiency of particular displays. Progress in flight simulation has been limited by a poor understanding of experiential fidelity and action fidelity." We will argue that metrics of action fidelity are more useful as constraints on the design and evaluation of vehicular simulations. We will further argue that a profound limitation of any simulation is the range of circumstances (e.g., the type of simulated events) over which it can produce action fidelity.

Experiential Fidelity: Presence

Experiential fidelity is the extent to which a simulation gives rise to a subjective experience of "being there" (e.g., Held & Durlach, 1991; Smets, 1995). Prothero and colleagues (1995, p. 359) have argued that experiential fidelity should be the sole criterion for the design of virtual environments. Recently, there have been attempts to formalize experiential fidelity in the concept of *presence*. This effort is derived largely from studies of situation awareness and vection (e.g., Prothero et al., 1995), and may be related to immersion. Presence has been operationalized in terms of conscious reports such as questionaires or numerical intensity ratings. For example, Prothero and colleagues (1995) assessed presence by asking subjects to give numerical ratings of the degree to which a simulation seemed "real." Presence is widely believed to be common in simulators (e.g., Carr, 1995; Prothero & colleagues 1995; Stappers & Smets, 1995).

Presence is widely understood to be an illusory percept. Prothero and colleagues (1995, p. 359; see also Slater, Usoh, & Steed, 1994) defined presence as "an illusion of position and orientation" (this complements *vection*, which is defined as an illusion of self-motion; e.g., Dichgans & Brandt, 1978). Carr (1995, p. 1) described virtual reality as "fooling people into accepting as real what is only perceived." Similarly, Ellis (1991, p. 323) referred to creating "the illusion of an enveloping environment," whereas Christou and Parker (1995, p. 53) asserted that with virtual environments "any sense of reality is . . . illusory" and that "it is possible to fool the perceiver by making it difficult for them to discern that the world they are experiencing is artificial." These definitions reflect widespread agreement that presence can exist if and only if people are fooled by the simulation.

Reality or Realism?

Despite its intuitive appeal, the concept of presence is not entirely straightforward. We regard as critically important a complication that has been noted by Carr: "It is important to distinguish between the perception of realism and the perception of reality . . . A 'sense of reality' does not necessarily imply belief in reality" (Carr, 1995, p. 6; see also Stoffregen, 1997). We equate "perception of realism" with perception of the simulation as such, and "perception of reality" with perception of that which is simulated.[7] This is a critical distinction that has not been addressed in the literature on presence. Researchers often appear to confuse presence with realism (e.g., Christou & Parker, 1995, p. 53–54).

The distinction between perceiving realism and perceiving reality is important because these percepts may differ qualitatively. A perception of realism would be accurate, reflecting an actual resemblance of the simulation to the simulated. For example, a painting of a building may be said to resemble the building, and an

[7]Is there a difference between the "immersion" of simulation and the "suspension of disbelief" (Goffman, 1974) that is required for appreciation of fictional entertainment? Suspension of disbelief embodies the conceptual distinction between realism and reality in theatrical performance. The suspension of disbelief implies a perception of the simulation as such. In the case of the theater, it implies a perception of the fact of theatrical performance (for a detailed discussion, see Goffman, 1974). That is, suspension of disbelief implies that theater and movie patrons experience the situation as realistic, but not as reality (Stoffregen, 1997). This underscores the fact that the distinction is not peculiar to simulation technology, but is an extension of issues that predate the development of computer-based simulation (cf. Steuer, 1992).

impersonation may be said to resemble the person who is impersonated. Thus, a perception of realism would not be an illusion. In a simulator, by contrast, a perception of reality would necessarily be erroneous; the person is not in the "real" system, and to perceive otherwise would be an error. Because it is defined as an illusion, presence cannot refer to the (accurate) perception of realism but only to an (inaccurate) perception of reality. Thus, a more precise definition of *presence* might be "an illusory (false) perception that the simulator is the simulated."

The logical distinction between perception of realism and perception of reality has implications for methods that are used to assess presence. Questionnaires that are intended to measure the illusory perception of reality (presence) may, instead, measure accurate perception of realism. For instance, Prothero and colleagues (1995) exposed experimental participants to a virtual environment. After exiting the virtual environment, participants were asked: "How real did the virtual world seem to you?" However, Slater and colleagues (1994) asked to what extent the "computer-generated world" was "more real" than the "real world." In their phrasing, these questions inform the subject that the situation they experienced was "virtual," or "computer generated." A subject who had truly been fooled by the simulation would be disabused of their error by this information. In addition, all possible answers to these questions are in terms of accurate, nonillusory perceptions of realism.

These questions are poorly formulated if the goal is to assess an illusory perception that the simulator is the simulated. Rather than asking, "How real does this seem?" a better question might be: "Is this a simulation, or is it reality?" (participants might also be asked to give numerical certainty ratings). With existing simulation systems, such questions have little credibility; the status of the simulation as such is obvious (this suggests that, rather than being common, presence may be very rare). A rigorous empirical evaluation of this would require the development of paired situations, one member of which was real whereas the other was simulated. In most contemporary research on presence, participants are exposed only to simulations (e.g., Prothero et al., 1995).

As a comparison, consider identical twins, Jane and Mary, both of whom are known to you. If one day you encounter Jane, you may correctly identify her as Jane. If you are asked, "Does this person resemble Mary?" the question would seem reasonable and you would reply in the affirmative. Another possibility is that in encountering Jane you erroneously perceive her to be Mary. In this case the question "Does this person resemble Mary?" would be nonsensical (it makes little sense to say that Mary resembles herself). In the same way, to ask whether a simulation "seems real" is to assume that the participant already knows that it is *not* real, that is, that there is no illusory perception of reality (presence) but only an accurate perception of realism.

One reason that users do not experience the simulation as being real is that they have prior knowledge that it is not. They know this because they put on a headset rather than getting into an aircraft, because they are being paid to use a simulator rather than an aircraft, and so on.[8] This is a powerful deterrent to the occurrence of any "belief in reality" (Carr, 1995, p. 6). As we noted earlier, it is extremely difficult to control or eliminate situational information. Users almost always know, before using a system, whether it is a simulator or the real thing. With adult humans, such prior knowledge could be eliminated only by draconian interventions of the kind found in the scenario that opens this chapter.

A related example occurs in the cinema. Despite recent advances in film technology (e.g., Omnimax cinema), movie patrons do not run away from cinematic dinosaurs, murderers, and so on. Presumably, this is because they experience realism but do not have any perception of reality, and this, in turn, is due to stimulus differences between film and reality (Stoffregen, 1997). When film was first developed, there were incidents in which patrons failed to differentiate events in a film from physical reality (Shapiro & McDonald, 1992). Early in this century, a Montana patron fell asleep during a film. When he awoke, the film included a bear. He mistook this for an actual bear and fired a gun. Similarly, it is reported that at the Lumiere brothers' first public exhibition of the new motion picture technology, audience members confused a scene of a train with a real train and fled from the theater in terror. The rarity of reactions of this kind is testimony to the fact that patrons perceive realism rather than reality.

Action Fidelity

We have defined stimulus fidelity in terms of prepsychological relations between reality and ambient energy arrays. Action fidelity (Riccio, 1995; cf. Caro, 1979) is defined in terms of relations between performance in the simulator and performance in the simulated system (this is similar to the concept of *functional fidelity*; Moroney & Moroney, 1998). Action fidelity exists when performance in the simulator transfers to behavior in the simulated system. An appropriate measure of action fidelity is transfer of learning, or transfer of training (e.g., Flach, Riccio, Mcmillan, & Warren, 1986; Kozak, Hancock, Arthur, & Chrysler, 1993; Moroney, Hampton, Beirs, & Kirton, 1994).

We have seen that experiential fidelity is measured via subjective reports. By contrast, action fidelity is measured in terms of task performance. Common metrics that could be used to compare performance in a simulator and in the simulated system are time to completion of a task, variance in performance across trials, and trials to criterion (Kozak et al., 1993; Moroney et al., 1994). Performance metrics appropriate for a flight control task might include time on course and magnitude

[8]Knowledge of this kind is often thought to originate in cognition. However, scheduling, waiting rooms, donning of simulation hardware, and other preparations all produce characteristic patterns of sensory stimulation. This is specification; the sense is global, but real nevertheless. Thus, such knowledge may have a perceptual basis (cf. Stoffregen, Gorday, Sheng, & Flynn, 1999).

of heading and altitude deviation (e.g., Moroney et al., 1994). Appropriate metrics for a manual tracking task might include position errors, time, and phase (Knight, 1987), whereas appropriate metrics for a telemanipulation task might include the accuracy with which users can position objects in a closed-loop 3-D video image (Smets, 1995, p. 193).

It can be very difficult to obtain data on transfer of skills from a simulation to the simulated system. An example is high-performance jet aircraft, for which it is both difficult and very expensive to collect data on actual flight control. Another example might be virtual environments for which there is not a corresponding physical system, such as a "walk-through" of the human body. For these systems, it may be impractical to proceed directly to the development of action fidelity metrics. An alternative might be to begin by concentrating on systems for which transfer of training can be measured more directly. For example, rather than concentrating on high-performance aircraft, measures of action fidelity might be first developed in automobiles.

Reliance on experiential fidelity motivates designers to maximize stimulus fidelity, in an effort to maximize the subjective experience. By contrast, action fidelity does not mandate a concentration on stimulus fidelity. Transfer of skills from the simulator to the simulated system may occur despite departures from stimulus fidelity. A display that does not look or feel realistic may nevertheless facilitate performance. For example, Moroney and colleagues (1994) studied the acquisition of instrument flight skills. Some participants were trained in an FAA-standard simulator, whereas others were trained using a PC-based desktop retail flight simulation program (whose cost was 96% less). Both groups were later tested for instrument flight skills in actual flight. Results showed that there was not a significant difference in actual flight skills between the groups trained on the FAA-approved simulator and those trained with the desktop system, despite the fact that the desktop system was less "realistic."

In some cases, departures from stimulus fidelity may produce improvements in action fidelity. Consider a situation in which a simulator is used to teach nap-of-the-earth flight (minimum controllable altitude, following terrain contours). Simulator training at realistically low altitudes would be inefficient because novices could not control the aircraft (this is why training is needed). There would be many crashes, and learning would be reduced or delayed. Training at high altitudes would be inefficient, because at high altitudes the optical consequences of changes in altitude are reduced. A solution might be deliberately to reduce the stimulus fidelity of the simulation; changes in optical splay resulting from controlled changes in altitude could be made to be greater than in real flight at the same altitude (i.e., an increase is the gain of the closed-loop optical splay function). This deliberate departure from subjective realism might lead to improvements in training (Warren & Riccio, 1986; cf. Flach et al., 1986; Stappers et al., 2002) despite a reduction in subjective realism.

Students of presence have not offered a clear reason why the user's subjective experience should be of interest to or important for the evaluation of simulator fidelity. The importance of presence appears to be assumed: Presence "is thought to correlate with improved task performance in virtual environments" (Prothero et al., 1995, p. 361; cf. Held & Durlach, 1992). However, the relevance of presence is an assumption that can, and should, be evaluated empirically. A consequence of our analysis is that the utility of experiences of realism and reality will be application specific: If the purpose of a simulation is to influence subjective awareness (e.g., entertainment), then experiential measures may be most appropriate. However, if the simulation is intended to influence behavior (e.g., vehicular control and training), then experiential measures may be insufficient.

CONCLUSION

In this chapter, we have developed several arguments: (1) that simulation is specified as such, in patterns of ambient energy that can serve as stimuli for perception; (2) that because of this, the simulator can be, and almost always is, differentiated from the simulated system; and (3) this reflects an accurate perception of realism (the reality of simulation), and not illusory perception of reality (the illusion of the simulated). These conclusions have implications for the scenario that we presented at the beginning of this chapter. If you awoke in this situation, we would expect that you would quickly perceive the true nature of the situation (whether you were in a simulator or real aircraft), and that your behavior would vary greatly depending on which situation you were in, with these variations being, for the most part, adaptive. Your percepts would be accurate, that is, you would perceive the reality of the situation. There would be no illusion. If you were in the simulator, you might enjoy noticing that it was highly realistic. This, too, would be an accurate (nonillusory) percept, and it would not interfere with your simultaneous perception of the simulator as such.

The arguments developed in this chapter may have important consequences for the design of experiments in VE and for the assessement of simulator fidelity. Experiments are needed that evaluate the perception of reality in addition to (or rather than) evaluating the perception of realism. These experiments are difficult to design because they require that subjects have no a priori knowledge about whether they will be in a simulator or in the simulated system. However, such experiments appear to be essential to a satisfying assessment of the perception of reality in simulations. A second need is for experiments on action fidelity, using paradigms such as transfer of learning. In particular, research is needed on the possibility that task-specific departures from subjective realism could be used to improve users' sensitivity to the relevant dynamics of the simulated system. Finally, it may be important to determine the actual correlations between performance and subjective experiences (perception of realism and, separately, reality). Work of this kind may help to determine the respective and complementary contribution of action fidelity and experiential fidelity to the design and evaluation of simulations.

ACKNOWLEDGMENTS

Preparation of this chapter was supported by grants from the National Science Foundation (SBR-9601351 and INT-9603315) to Thomas A. Stoffregen and by a grant from the National Center for Scientific Research (CNRS/NSF-3899) to Benoît G. Bardy, with additional support from the French Ministry National Education, Research and Technology.

REFERENCES

Biocca, F. (1992). Will simulation sickness slow down the diffusion of virtual environment technology? *Presence, 1*, 334–343.

Brandt, T., Dichgans, J., & Koenig, E. (1973). Differential effects of central versus peripheral vision on egocentric and exocentric motion perception. *Experimental Brain Research, 16*, 476–491.

Caro, P. W. (1979). Relationship between flight simulator motion and training requirements. *Human Factors, 21*, 493–501.

Carr, K. (1995). Introduction. In K. Carr & R. England (Eds.), *Simulated and virtual realities* (pp. 1–10). Bristol, PA: Taylor & Francis.

Christou, C., & Parker, A. (1995). Visual realism and virtual reality: A psychological perspective. In K. Carr & R. England (Eds.), *Simulated and virtual realities* (pp. 53–84). Bristol, PA: Taylor & Francis.

Dichgans, J., & Brandt, T. (1978). Visual–vestibular interaction: Effects on self-motion perception and postural control. In R. Held, H. Leibowitz, & H. Teuber (Eds.), *Handbook of sensory physiology* (Vol. 8, pp. 755–804). New York: Springer-Verlag.

DiZio, P., & Lackner, J. R. (1992). Spatial orientation, adaptation, and motion sickness in real and virtual environments. *Presence, 1*, 319–328.

Edgar, G. K., & Bex, P. J. (1995). Vision and displays. In K. Carr & R. England (Eds.), *Simulated and virtual realities* (pp. 85–102). Bristol, PA: Taylor & Francis.

Ellis, S. R. (1991). Nature and origins of virtual environments: A biblographical essay. *Computing Systems in Engineering, 2*, 321–347.

England, R. (1995). Sensory–motor systems in virtual manipulation. In K. Carr & R. England (Eds.), *Simulated and virtual realities* (pp. 131–178). Bristol, PA: Taylor & Francis.

Flach, J. M., Riccio, G. E., McMillan, G., & Warren, R. (1986). Psychophysical methods for equating performance between alternative motion simulators. *Ergonomics, 29*, 1423–1438.

Gibson, E. J., & Rader, N. (1979). Attention: The perceiver as performer. In G. A. Hale & M. Lewis (Eds.), *Attention and cognitive development* (pp. 1–21). New York: Plenum.

Gibson, J. J. (1971). The information available in pictures. *Leonardo,4*, 27–35.

Gibson, J. J. (1986). *The ecological approach to visual perception.* Mahwah, NJ: Lawrence Erlbaum Associates. (Original work published 1979)

Gilkey, R. H., & Anderson, T. R. (1997). *Binaural and spatial hearing in real and virtual environments.* Mahwah, NJ: Lawrence Erlbaum Associates.

Goffman, E. (1974). *Frame analysis.* New York: Harper & Row.

Held, R., & Durlach, N. (1992). Telepresence. *Presence, 1*, 109–112.

Hochberg, J. (1986). Representation of motion and space in video and cinematic displays. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1, Chapter 22). New York: Wiley.

Kozak, J. J., Hancock, P. A., Arthur, E. J., & Chrysler, S. T. (1993). Transfer of training from virtual reality. *Ergonomics, 36*, 777–784.

Knight, J. R., Jr. (1987). Manual control and tracking. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 182–218). New York: Wiley.

Lishman, J. R., Lee, D. N. The autonomy of visual kineasethesis. *Perception, 2*, 287–294.

Mark, L. S., Balliet, J. A., Craver, K. D., Douglas, S. D., & Fox, T. (1990). What an actor must do in order to perceive the affordance for sitting. *Ecological Psychology, 2*, 325–366.

Moroney, W. F., Hampton, S., Beirs, D. W., & Kirton, T. (1994). The use of personal computer-based training devices in teaching instrument flying: A comparative study. In *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting* (pp. 95–99). Santa Monica, CA: Human Factors and Ergonomics Society.

Moroney, W. F., & Moroney, B. W. (1998). Simulation. In D. J. Garland, J. A. Wise, & V. D. Hopkin (Eds.), *Human factors in aviation systems* (pp. 358–388). Mahwah, NJ: Lawrence Erlbaum Associates.

Michaels, C., & Beek, P. (1995). The state of ecological psychology. Ecological Psychology, 7, 259–278.

Nahon, M. A., & Reid, L. D. (1990). Simulator motion-drive algorithms: A designer's perspective. *Journal of Guidance, Control, and Dynamics, 13*, 356–362.

Nickerson, R. S. (1992). *Looking ahead: Human factors challenges in a changing world.* Mahwah, NJ: Lawrence Erlbaum Associates.

Pausch, R., Crea, T., & Conway, M. (1992). A litearture survey for virtual environments: Military flight simulator visual systems and simulator sickness. *Presence, 1*, 344–363.

Prothero, J. D., Parker, D. E., Furness, T. A., & Wells, M. J. (1995). Towards a robust, quantitative measure for presence. In *Proceedings, Experimental analysis and measurement of situation awareness* (pp. 359–366). Dayton Beach, FL: Embry-Riddle Aeronautical Univerisity Press.

Riccio, G. E., & Stoffregen, T. A. (1991). An ecological theory of motion sickness and postural instability. *Ecological Psychology, 3*, 195–240.

Riccio, G. E. (1995). Coordination of postural control and vehicular control: Implications for multimodal perception and simulation of self-motion. In P. Hancock, J. Flach, J. Caird, & K. Vicente (Eds.), *Local applications of the ecologial approach to human–machine systems* (pp. 122–181). Hillsdale, NJ: Lawrence Erlbaum Associates.

Runeson, S., & Vedeler, D. (1993). The indispensability of precollision kinematics in the visual perception of relative mass. *Perception & Psychophysics, 53*, 617–632.

Shapiro, M. A., & McDonald, D. G. (1992). I'm not a real doctor, but I play one in virtual reality: Implications of virtual reality for judgments about reality. *Journal of Communication, 42*, 94–114.

Slater, M., Usoh, M., & Steed, A. (1994). Depth of presence in virtual environments. *Presence, 3*, 130–144.

Smets, G. J. F. (1995). Designing for telepresence: The Delft Virtual Window System. In P. Hancock, J. Flach, J. Caird, & K. Vicente (Eds.), *Local applications of the ecologcial approach to human–machine systems* (pp. 182–207). Mahwah, NJ: Lawrence Erlbaum Associates.

Smets, G. J. F., Stappers, P. J., Overbeeke, K. J., & Mast, C. (1995). Designing in virtual reality: Perception–action coupling and affordances. In K. Carr & R. England (Eds.), *Simulated and virtual realities* (pp. 189–208). Bristol, PA: Taylor & Francis.

Stappers, P. J., Overbeeke, C. J., & Gaver, W. W. (in press). Beyond the limits of real-time realism: Uses, necessity, and theoretical foundations of non-realistic virtual reality. In M. W. Haas & L. J. Hettinger (Eds.), *Psychological issues in the design and use of virtual and adaptive environments.* Mahwah, NJ: Lawrence Erlbaums Associates.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication, 42*, 73–93.

Stoffregen, T. A. (1985). Flow structure versus retinal location in the optical control of stance. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 554–565.

Stoffregen, T. A. (1997). Filming the world: An essay review of Anderson's *The reality of illusion. Ecological Psychology, 9*, 161–177.

Stoffregen, T. A., & Bardy, B. G. (2001). On specification and the senses. *Behavioral and Brain Sciences, 24*, 195–261.

Stoffregen, T. A., Gorday, K. M., Sheng, Y-Y., & Flynn, S. B. (1999). Perceiving affordances for another person's actions. *Journal of Experimental Psychology: Human Perception & Performance, 25*, 120–136.

Stoffregen, T. A., & Riccio, G. E. (1988). An ecological theory of orientation and the vestibular system. *Psychological Review, 95*, 3–12.

Stoffregen, T. A., & Riccio, G. E. (1990). Responses to optical looming in the retinal center and periphery. *Ecological Psychology, 2*, 251–274.

Stoffregen, T. A., Smart, L. J., Bardy, B. G., & Pagulayan, R. J. (1999). Postural stabilization of looking. *Journal of Experimental Psychology: Human Perception & Performance, 25*, 1641–1658.

Wann, J. P., Rushton, S. K., & Mon-Williams, M. (1995). Natural problems for stereoscopic depth perception in virtual environments. *Vision Research, 19*, 2713–2736.

Warren, R., & Riccio, G. E. (1986). Visual cue dominance hierarchies: Implications for simulator design. *Transactions of the Society for Automotive Engineering, 6*, 937–951.