Authenticity, Concordance, Dissonance, Fun: retrospectives on designing the *Extraterrestrial Virtual Field Experience*

Introduction

The *Extraterrestrial Virtual Field Experience* (EVFE) is an interactive educational simulation intended for use in middle and high school earth science courses and introductory undergraduate geology courses. The simulation immerses students in the role of a NASA science team tasked with remotely operating a rover and interpreting its telemetry. Multiple scenarios are eventually planned; The currently-designed scenario presents the Mars Exploration Rover *Opportunity's* operations at Meridiani Planum. Future scenarios are planned. They include both additional Mars operations and the operations by other probes and landers. A tentative list includes: *Opportunity* at Victoria Crater, *Spirit* at Home Plate, *Curiosity* in the north of Gale Crater, the Pathfinder mission, the Viking mission, *Cassini-Huygens*, and the last three Apollo missions.

This article is a discussion of the EVFE, its approach to science education, and some challenges we encountered in its design. It focuses on some unique and underexplored opportunities that immersive multimedia provide for teaching students about how science, especially science with remote sensing instruments, gets done. It further explores the implications some of our design challenges might have for interactive educational software in this general space, meaning historically accurate simulations, especially of complex scientific processes. In this discussion, we arrive at some possible limits of the category "game" in relationship to learning outcomes, historical fidelity, and modes of interactivity. We finally present several potential methodologies for further work in this space that might skirt or shift these limits.

We will begin with a brief overview of the current state and pedagogical aims of the EVFE. Further descriptions of the scientific background and intended supporting classroom activities of the Meridiani Planum scenario can be found in our Teacher's Guide[1], and a live version of the EVFE is available through the Cornell Spacecraft Planetary Image Facility[2]

Overview of the EVFE

The EVFE's base display[3] presents students with images from *Opportunity's* navigational camera and hazard cameras (NavCam and HazCam) depicting the rover's current surroundings. These views -- like all imagery utilized in the program -- consist of actual data products from the mission. These data products are sometimes cropped or scaled down in resolution, and often have highlights to indicate areas of interest, but are otherwise not altered.

Highlighted areas on this main view denote potential targets for more detailed observations. Clicking on one of these targets produces a secondary dialog that explains why the science team has identified it as interesting, what sort of data it might reasonably be expected to return, and potential instruments from the rover that can be used to make these detailed observations.

These observations return further data, along with a summary report that helps students interpret and develop meaningful cognitive models for the data. This is also the primary user experience loop of the simulation: completing these observations reveals new targets for investigation and new areas to explore. Following this sequence of observation, exploration, interpretation, and investigation allows students to virtually experience the narrative of a significant space science mission.

The EVFE's intended classroom role can be well-described, as we do in our Teacher's Guide, as a computer-based laboratory module. However, from a design perspective, it does not fit easily into existing software, pedagogical, or literary genres, so we have developed a working conceptual model for its organization and function. One may think of the EVFE as an *immersive narrative workbook*, or, equivalently, as an *interactive nonfiction.*

---

[1] *Teacher's Guide to the Extraterrestrial Virtual Field Experience: Mars Exploration Rover Opportunity at Meridiani Planum.* Chase Million and Don Duggan-Haas, 2017. https://cornellspif.com/wp-content/themes/spif2017/img/TeachersGuide141031.pdf
[2] https://cornellspif.com/educators/#evfe
[3] Images are not included in this version of this paper. Please refer to the *Teacher's Guide* or the EVFE software itself.

The first phrase describes its relationship to classroom practice; the second term describes its relationship to videogames.
The EVFE is a workbook in the sense that it is a text that enables its readers to perform exercises intended to produce particular learning outcomes. The simple fact of interactivity, though often lauded as a salient feature of digital technology, is a longstanding feature of such pedagogical devices. All workbooks, including traditional paper ones, are interactive and participatory. The 'work' is crucial; they are not merely passive repositories of information. However, the digital affordances of the EVFE add an immersive layer to the participatory qualities enabled by best-practice traditional workbook exercises.

Brenda Laurel's typology of components of interactivity is a useful point of reference here [4]. Laurel notes that the degree of interactivity of a system can be described as some function of the *range*, *frequency*, and *significance* of the inputs it accepts from its users, combined with *subjectively experienced immersion or participation*. The range of inputs the EVFE accepts is fairly small: selection of preplanned observations, toggling between camera views, and moving between sites. The frequency is moderate and paced, with rapid feedback, similar to a well-supervised classroom exercise (or turn-based strategy game). The significance of the input is, again, similar to a traditional classroom exercise; users may not proceed to later portions of the EVFE without correctly completing earlier portions, but variances in user input do not lead to entirely separate scenario outcomes. (There are no major narrative branches or hidden endings..)

In the dimension of experienced immersion, however, there *is* a sharp difference between the EVFE and a conventional workbook--or, for that matter, conventional interactive educational software. The interface is intended to feel familiar to our users, following best-practice contemporary user experience design standards in a generally flat style likely familiar to users from many mobile apps. The *content* of the game is not, however, intended to feel familiar. Rather than abstract 2D graphics, cartoonishly stylized 3D, or Unreal Engine-style faux-photorealistic 3D, it uses high-resolution photographic data taken directly from *Opportunity's* cameras. The effect is stark, breathtaking, and deeply immersive in a way that is vastly unfamiliar to most users.

---

[4] See the second edition of *Computers as Theatre* (2013), pp. 29-34, along with the entirety of her classic essay "Interface Mimesis" (in *User centered system design: New perspectives on human-computer interaction* (1986), eds. Don Norman and Stephen Draper).

More specifically, it immerses users in a narrative of scientific exploration. Users are, as one tester put it, “squarely in the driver’s seat of the mission.” They work through the *Opportunity* team’s meticulous search for hematite deposits at Meridiani Planum. They see what the team saw, engage with the team’s interpretations of collected data and rationales for further investigation, and chase the clues the team followed from site to site on the crater. It is an immersive narrative workbook, and can be deployed as such in a formal classroom, or in informal group or individual learning settings.

Although we currently do not think of the EVFE as a game (a point we will discuss in detail a little further on), we originally designed it as a game, and it bears structural similarities to some videogames. In particular, it has a close relationship to the genre of videogames known as *interactive fiction*. Broadly speaking, interactive fiction games are stories whose narrative form or content are responsive in some way to user input. It is the playful fraction of the larger domain that cybertext theorist Espen Aarseth called *ergodic literature:* texts that take work (above and beyond the usual labor of reading) to navigate.[5]

Interactive fiction is one of the oldest genres of videogames, and the forms that interactive fictions take are extremely diverse, from parser-based games that include some form of natural language processing (often along with rich physical simulation of environments and complex puzzles) to choice-based games that offer explicit, discrete options at decision points. They include games of pure reading and typing, games that imbed fictions in graphical environments, and games that mix entirely different game genres (action, strategy, etc.) with interactive fiction elements. The browser and mobile market spaces, along with the development of lightweight interactive fiction scripting environments (notably Twine), have led to a renaissance of popular interactive fiction in the past decade. Although its development history predates their release dates, the EVFE is particularly close to some recent exploration-focused interactive fictions with rich graphical environments and discrete decision trees, including *80 Days*, *Sunless Sea*, and the rover game *Extrasolar*. The EVFE is an educational expression of trends in game design that have pushed the entertainment gaming market back towards these forms.

However, to be clear: the EVFE is not interactive fiction precisely because it is not fiction. It meticulously follows a series of historical events. While it is possible for users to collect a few pieces of data out of strict historical order, these are relatively minor. There is no variation in ultimate outcome, including the possibility of failure that even the most linear, puzzle-heavy, classic text adventure (Infocom’s *Zork*, for instance) tends to implement. As such, we conceptualize the EVFE as an interactive nonfiction, combining some formal

[5] *Cybertext—Perspectives on Ergodic Literature* (1997).

discursive techniques from interactive fiction games with a pedagogical ethos more closely related to popular scientific nonfiction.

Background of the EVFE

Our original inspiration for the EVFE came from the realization that remote space science missions are excellent educational case studies. They involve dynamic demonstrations of a variety of scientific principles, especially related to geology and earth systems. Better yet, because of the extremely challenging and radically interdisciplinary qualities of remote planetary operations, they provide object lessons in many themes that tie the STEM disciplines together. These interconnecting themes are increasingly considered core parts of science education: they appear, for instance, as the "crosscutting concepts" of the Next Generation Science Standards[6]. Rover operations potentially demonstrate almost all of the NGSS crosscutting concepts. They are especially well-suited to express the often-neglected but crucial relationships between scientific research, engineering, and technology.

The Meridiani Planum scenario seemed like a very strong first choice for immersing students in a remote space science narrative. There are several central reasons for this. One is the generally high popularity of Mars rovers, including the broad circulation of images from rover-mounted cameras in both mainstream press and popular science communities. Rovers are both extremely powerful and highly interesting to laypeople. The Meridiani Planum scenario particularly highlights the power of rovers as investigative tools. Many of *Opportunity's* findings there were foundational for contemporary studies of Martian hydrology and relied on proximal observations that could not have been performed by orbital instruments. These observations demonstrate geoscience concepts deeply relevant to earth and environmental science, including erosive processes and the oxidation and leaching effects of water.

We were extremely familiar with *Opportunity's* operations at Meridiani Planum. Our team included members of the mission who had been instrumental in making day-to-day decisions about interpretation and mission direction. Therefore, we had deep access not only to the mission data as such, but more importantly, the mission data considered as part of a narrative of scientific investigation. We could immerse students in scientific thought as an unfolding process of gathering data and drawing conclusions that shaped further data-gathering choices.

[6] https://www.nextgenscience.org/

Furthermore, rover operations are themselves a conceptual model for for an existing, effective set of digital science pedagogy tools: the Paleontological Research Institution's (PRI's) Virtual Field Experiences (VFEs)[7]. The VFEs are collections of photographs, videos, diagrams, Google Earth files, worksheets, and lesson plans intended to provide students with guided, immersive access to and simulated research experiences in far-flung terrestrial environments. Several of our personnel were instrumental in the development of PRI's VFEs and had been extremely impressed by the effectiveness of these technically modest techniques for teaching geoscience concepts. As its name suggests, the EVFE is to some degree an extension of the VFE concept, taking it to Mars and integrating its disparate media into a coherent, interactive simulated environment.

There is also a sense in which the EVFE returns the VFE concept to its intellectual roots. As Ross and Duggan-Hass (2012) outline, the Mars rover missions provided PRI's VFE designers with their original model for the VFE. Rover missions are literally virtual fieldwork experiences. All fieldwork on Mars, simulated or otherwise, is virtual fieldwork. The terrestrial VFE is an application of planetary science remote sensing and operations methodologies to earth science pedagogy. This means that stakes of realism for a VFE simulation set on Mars are significantly higher.

Immersion on Mars[8]

Mars is not Earth, and it was crucial for us to keep this in mind as we attempted to create the correct *kind* of immersion. To expand on this point: The terrestrial VFEs work to create a kind of extension of users' bodies. They create unified viewpoints that (phenomenologically speaking) extend or project the bodies of users into (for instance) the Valles Caldera in New Mexico, or at least a simulation of it. The production of this extension relies on the prior documented and documentary presence of expert viewers at the Valles Caldera; simultaneously, it hides the presence of these others, working as a kind of promise of the possible future presence and the fictive present presence of the users' bodies at the Valles Caldera. Terrestrial VFEs work to make users feel like they're *really there*. Meridiani Planum is not like the Valles Caldera. No one has ever been to Meridiani Planum, and it is possible that no one ever will.

[7] https://www.priweb.org/science-education-programs-and-resources/virtual-field-experiences
[8] This is a pun about water.

There are two obvious problems with that last sentence. First, rovers have been to Mars, and certainly give a kind of experiential access to Mars' surface. Otherwise, we could not be writing this at all. Second, it is certainly possible to create *fictions* that project users into a simulation of the surface of Mars, but we are interested in user experiences that authentically convey nonfiction narratives of scientific investigation. So do rovers generally serve as points of experienced projection for scientific investigators? To put this another way, the proper "driver's seat of the mission" an experience of real-time, seamless operation of the rover, of users feeling coterminous with the rover on Mars?

Janet Vertesi takes this question up in *Seeing Like a Rover*, her comprehensive 2014 ethnography of the Mars Exploration Rover teams. Though Vertesi is concerned with a wide variety of social phenomena, mission members' embodied experiences of rover operations provide an organizing theme for her study. In the course of theorizing this category of experience, she articulates a careful distinction between two different modes of telemetric experience. The first relies on integration of remote instruments and their telemetry into operators' preexisting physical models of themselves, resulting in an experience of seamless unity in which the instruments recede into operators' gestalt experience of the physical world. The second relies on introjecting the qualities of remote instruments into operators' kinesthetic experiences, highlighting the instrument's subject position--and its radical gulf from its operators' usual subjectivities--in an attempt to more adequately understand how that instrument "sees" or otherwise experiences the physical world.

Vertesi articulates the first mode with reference to Rachel Prentice's work on the experiences of laparoscopic surgeons with their tools[9]. Using Prentice's gloss on the perceptual phenomenology of Maurice Merleau-Ponty, Vertesi calls remote instruments experienced in this way "proxies" for their users.

Delving slightly deeper into Merleau-Ponty, these are the sorts of instruments he describes as spaces of knowledge within a user's sensorium rather than discretely-experienced objects: an expert driver's car, an experienced blind navigator's stick, a well-known typewriter, the feather on the top of a long-familiar hat. "In the exploration of things," he writes, "the length of the stick does not enter expressly as a middle term: the blind man is rather aware of it through the position of objects than of the position of objects through it[10]." What Vertesi calls proxies are precisely these objects which, by dint of practice and habituation, users integrate into their gestalt sensorium. Proxies are objects people have made part of their subjectivities.

---

[9] Prentice's most complete discussion of this topic is in *Bodies in Formation* (2013), particularly chapter 5.
[10] *Phenomenology of Perception* (tr. Colin Smith). Taylor & Francis e-Library edition (2013).

Vertesi argues that in this sense, the rover generally does *not* serve as a proxy for mission members. Scientists and engineers certainly perceive a relationship between the rover and their own bodies, but the terms of this mode of experience are almost precisely opposite to proxyhood. Rather than projecting themselves transparently through the rover to the Mars surface, they introject the rover and its telemetry into their own bodies. They study PanCam images to make their eyes into adequate representations of the PanCam, they dig with their hands to understand how the rover might dig--but these are attempts to sharpen their awareness of the rover's objecthood, not to swallow it in order to gain unmediated access to the Martian surface.

This rejection of intuitive but incorrect assumptions about proxying has resonances with practical problems in user experience design. It is very close to classic distinctions between seamless and seamful design highlighted by ubiquitous computing pioneer Mark Weiser in the mid-1990s[11] and to the ongoing conversation about how seamless integration in learning discovery tools risks suppressing critical thought about research methodologies.

Further, though Vertesi does not explicitly make this critique of phenomenological accounts of tool use, they commonly valorize processes in which users map tools into their preexisting kinesthetic and proprioceptive experience and to treat said processes as both desirable and normative. Such values underlie and have distorted much popular discourse on the experience of technology, and it is significant that Vertesi most often observed mission members describing the rover as proxy only to the media (176).

The distinction may seem overly fine: both of these modes of telemetric experience are deeply involved, sensorily immersive experiences of human-machine interaction. The introjective mode, however, does not describe the kind of seamless experience of incorporation that expert drivers feel with their cars or expert prosthetic users feel with their prosthetics. It also does not describe an experienced or imagined projection of human presence onto the surface of Mars. The bodies of mission members provide them with a way of knowing the rover, but a way of knowing that continuously insists on the vast gulf and essential difference between their subject positions and the subjectivity of the rover.

This corresponds to a crucial difference between the EVFE and most existing rover simulations. The most common form of rover simulation is basically a real-time driving

[11] "Creating the Invisible Interface." *UIST '94: Proceedings of the 7th annual ACM symposium on User interface software and technology*.

game. Examples of this form include both EPO efforts like JPL's *Explore Mars!* and commercial offerings like Bohemia Interactive's *Take on Mars*. Unlike these games, the EVFE does not aim to recreate a direct experience of the Martian landscape, but rather the Martian landscape as seen and navigated by science missions. On the other end of immersion are programming and design games like NASA's *ROVER*, or, more fantastically, the UK Science Museum's *Rugged Rovers*, which do not aim to offer an experience of science missions at all.

The rover-control components of *Extrasolar* are the closest existing analog, in terms of desired mode of immersion, to the EVFE's design. They involve giving instructions to the rover and receiving telemetry in response to those instructions after a time delay. The overall gameplay of *Extrasolar*, however, is extremely different: it involves investigations of fictional planets, simple visual tagging of interesting objects rather than detailed investigation of their granular properties, and to a large extent focuses on an Earth-based thriller narrative revealed through emails and other communications.

Design Choices & Challenges

To summarize: we wanted to develop a piece of software that immersed students in the experience of the *Opportunity* team and thereby facilitated learning outcomes related to geoscience and crosscutting themes in science, engineering, and technology. This set of priorities, of course, does not entirely determine the design requirements of a simulation. The EVFE has had two major design iterations (disregarding primarily distribution/backend-oriented iterations like its recent port to Android). Its current iteration is, as we have noted, an immersive workbook or interactive nonfiction. The first iteration was a simulation game.

We should note that we're not using "simulation" in the very broad sense in which some educational software theorists, notably Clark Aldrich, use it, as a category of media which includes all games, digital microcosms, virtual practice environments, practiceware, and so on[12]. We're using "simulation game" here in the narrower sense it tends to be used in commercial game markets: a game that centrally works to present a model, possibly highly abstracted or simplified, of some game-extrinsic process or phenomenon that claims to grant insight (predictive, interpretive, or historical) into that process or phenomenon, or sometimes procedural skills related to that process or phenomenon, while also being fun.

---

[12] *The Complete Guide to Simulations and Serious Games* (2009).

The attempt to make the EVFE a simulation game was, in our opinion, basically unsuccessful. The subsequent sections of this article discuss the practical and theoretical grounding and implications of why this approach didn't--and perhaps couldn't--work.

Initially, we conceived the EVFE as a rover-command simulation game with a firm attachment to historical mission data. More specifically, we planned to create an exploration-investigation game with strong resource-management elements.
Many elements of the current EVFE, naturally, existed in our initial prototypes. The basic interaction sequence was even quite similar. We presented players with Pancam and Navcam views taken from original mission data, with visual highlights indicating potential targets for further investigation. Players could interact with these highlighted targets to access further information about why these targets had been identified by mission scientists as especially interesting. After reviewing their options at a given site, players could then enqueue observations for the sol, upload these instructions to the rover, receive data resulting from these observations, and subsequently use these data to plan further observations and possible site-to-site travel.

We retained the observation-queuing interaction loop because it has useful immersive qualities. Initially, however, we planned the observation queue as a primary site of time management and optimization. The primary difference between this model and our current workbook model is the presence of investigatory challenge produced by resource management challenges. We hoped that these elements of challenge would enhance a sense of immersion in the experience of control of an enormously complicated instrument separated by vast distances, an experience characterized by difficulty and interpretive work more than simple procedural action. We also hoped that they would highlight some elements of scientific investigation often ignored in secondary science education settings and especially important in rover operations--particularly, the mutually interdependent relationships between science and engineering priorities.

The most basic of these planned challenges was a limit on the number of sols players would have in which to make observations and travel between sites. This simulated limitations on the amount of power stored in the rover's batteries. If players didn't made the crucial observations necessary to verify the role of water in the geologic history of hematite at Meridiani Planum by the time their sol/power limit came up, they would lose the game. Similarly, we also strongly considered implementing bandwidth constraints, simulating real-life limits on the total amount of uploaded data per sol. However, we did not find an acceptably lightweight way to make these constraints feel meaningfully different from time/power constraints on action, and never actually implemented them.

In this design, time management was a crucial element of play. The amount of time required to conduct a particular observation--for instance, considerably more to engage in a complex physical manipulation like the use of the Rock Abrasion Tool than to take an additional PanCam image--made certain types of observations considerably “costlier” than others. This would require players to carefully study received data to ensure that targets for costly observations were well-justified. And travel, of course, was the most costly process of all, and we wanted to make users think quite hard about the choice to order a drive to a new site, to ensure that useful investigations at the current site were exhausted and that the data clearly pointed towards the target site. Ill-considered drives could easily run down the rover’s batteries with little useful scientific return.

These time/power management pressures, we hoped, would serve both game design and instructional design functions. They would enable a kind of puzzle-exploration gameplay in which careful planning and investigation was necessary to ensure success. This planning would require students to carefully interpret returned data.

We expected interpretation under resource management pressures to serve a variety of instructional purposes. These purposes obviously included technical reading comprehension and subject area knowledge. However, they also approached a more salient piece of our experiential, immersive science instruction. We wanted students to experience science as a concrete, material, iterative process of investigation and exploration. In secondary education settings, even when science is not presented as a collection of established facts or relations, it is often presented as a sort of merely logical exercise in which hypotheses spring full-formed from the brows of scientists to be verified or dismissed by some separate and entirely pre-planned process of data collection. This is both dull and inaccurate. Most real scientific investigations, especially those performed under constraints as sharp as those experienced by space mission teams, have feedback between experimentation and interpretation, between gathered data and processes of data collection. When time, mass, power, and bandwidth (just to name a few limiting factors) are in short supply, every act of observation has to count, and every observation can work to guide future observational choices.

Unfortunately, we ran into an irreconcilable difficulty at this point. One of the most generally-recognized fundamental game design principles is that games must offer players meaningful choices[13]. Another way to put this is that, for a game to be good, there has to be a way to play it badly (without merely refusing to play, or playing it incorrectly). In other words, in order to make this type of puzzle-exploration gameplay good, multiple

---

[13] Perhaps most famously prosecuted by Katie Salen and Eric Zimmerman in their 2003 *Rules of Play*.

paths must be available to players. In this case, it required the existence of sites and targets for observation that didn't lead directly to conclusions about the hydrology of Meridiani Planum, and possibly even entire inefficient but valid chains of observation and inference.

Since *Opportunity* in fact only traveled to the sites it traveled to and performed the observations it performed, implementing such red-herring or suboptimal paths would have required developing *historically counterfactual* sites, targets, and/or observations. However, one of our fundamental instructional and aesthetic principles was to use historically accurate data, so we did not have and could not produce historically counterfactual data. We were left with a "game" with a single path to victory, rather than a superfluity of options that required players to carefully analyze and select likely routes towards scientifically useful next steps.

As such, pruning the expectation of many options and transforming the EVFE into a full-fledged interactive nonfiction/immersive workbook was a natural route. This draws on the strengths of our rich historical dataset and uses it as a meticulously worked example of remote scientific investigation, rather than as a challenge in which students themselves attempt to meet the resource constraints of remote scientific investigation.

While this solution to the anti-game qualities of our instructional and educational design goals has been eminently workable, it is worth considering some implications of this 'ungameability' to the broader space of experiential learning products. We feel it has important lessons for the category of educational simulation games in general.

Learning, Fun, and Simulation

For the purposes of generalizing this set of difficulties, it is worth engaging in a brief review of some topics related to fun and educational games. In particular: what makes learning games fun? And are there modes of learning, or even types of educational content, that, even if potentially interesting and enjoyable, cannot be made into fun *games*?

We might begin by noting that designers of educational games (not to mention other, non-game categories of fun educational experiences) often perceive a tradeoff between 'fun' and 'educational value'. This perception is longstanding, and, despite regular bursts of hopefulness and boosterism, quite stable. Obstacles to educational play appear as central themes in all major bodies of work on learning games[14]. Amy Bruckman's late-90s description of the edutainment sector as "chocolate-covered broccoli" remains perennially apt[15].

Certain aspects of this opposition are unique to the particular characteristics of formal education environments, especially at the primary and secondary level. Multiple characteristics of classroom environments tend to militate against the kind of exploratory, self-directed immersion that makes games fun: a high concern with tracking and documentation of learning outcomes, an insistence on timed activities, and an emphasis on broad surveys of subject matter rather than in-depth investigation of particular concepts[16].

However, despite the contribution of these particular frictions between formal schooling practices and play as such, close analogs of this opposition can also be found outside the strictly educational space. Most relevant to the current discussion, makers of simulation games intended for the entertainment market often perceive a tradeoff between 'fun' and 'realism' (or 'accuracy', 'granularity', etc.), a term that has very close connections to learning. For instance, consider the following statement from a designer of *East India Company*, a game released by Paradox, one of the most successful mainstream commercial simulation game studios:

> We did have to consider realism versus accessibility and general gaming experience. Too realistic is not necessarily that much fun in the long run. For example, just getting on cannon range of the enemy could take hours of real time, so naturally we have cut some corners there.[17]

These perceptions reveal a shared aporia at the aesthetic heart of "educational game" and "simulation game." Written out explicitly, this aporia would read something like:

> *Games are about free play and the unreal; education and simulation are about constrained tasks and the real. To meet the demands of one is to sacrifice the demands of the other; to be accurate and educational is to be unfun.*

In both cases the present-time orientation suggested by the category of 'game' (having fun in the now without need of game-external referents) appears to militate against the future-time orientation suggested by the category of 'education' or 'simulation'

---

[14] Bjorn Marklund, *Games in formal educational settings: Obstacles for the development and use of learning games.* University of Skövde, 2013.
[15] Amy Bruckman, "Can Educational be Fun?" *Game Developer's Conference (GDC)* 1999.
[16] Suzanne de Castell and Jennifer Jenson, "Paying Attention to Attention." *Educational Theory 54:4* (2004).
[17] Kim Soares (2009), "Interview with Nitro Games' Kim Soares About East India Company" http://diehardgamefan.com/2009/04/08/east-india-company-interview-with-kim-soares/

(developing competencies for later use, or creating models of events with predictive value for future iterations of said events).

This aporia is, of course, by no means an absolute contradiction. There are no shortage of approaches to producing fun educational games. Frameworks for producing enjoyable learning games have existed since the late 1970s[18]. Despite the existence of these approaches and frameworks, we must admit that many educational games are unfun.

This is not a trivial problem. It calls into question the purpose of making an instructional object into a game at all. Moreover, if many educational gaming theorists are to be believed, it renders educational games considerably less effective. Many of the loudest voices in favor of deploying games in educational settings--even setting aside more radical arguments like Eric Zimmerman's that games are themselves the most proper and characteristic form of contemporary literacy[19] (Zimmerman 2013)--have emphasized their fun or otherwise motivating qualities as intrinsic components of their salient educational utility. For instance, James Gee's argument that "videogames are learning machines[20]" is worth keeping in mind here: every game teaches players some set of skills and content knowledge, and does so precisely because fun is a powerful motivator. Similarly, we might consider Raph Koster's famous, though rather less specific injunction that "fun is just another word for learning[21]," derived from Chris Crawford's "fun is the emotional response to learning.[22]"

Gee's statement is in some sense incontrovertibly true--if games didn't enable learning, players wouldn't be able to improve at games they play repeatedly, and wouldn't be able to play complex games at all. The Koster/Crawford position is more aspirational, and certainly more contentious. However, it again becomes very hard to argue with given a sufficiently narrow definition of "true fun" that excludes the three types of fun Nicole Lazzarro defines in her Four Keys[23] typology as "Easy Fun," "Serious Fun," and "People Fun," leaving only "Hard Fun," the kind of fun derived from overcoming challenges. Intellectual challenge certainly has a relationship to learning, particularly in regards to the development of problem-solving skills and virtues like "grit".

---

[18] See for instance Thomas Malone's "Toward a theory of intrinsically motivating instruction.", *Cognitive science*, 5(4) (1981).
[19] "Gaming Literacy: Game Design as a Model for Literacy in the Twenty-First Century." In the *Video Game Theory Reader 2* (2008).
[20] "Learning by Design: Good Video Games as Learning Machines." *E-Learning and Digital Media* 2.1 (2005).
[21] *A Theory of Fun for Game Design* (2004),
[22] *Chris Crawford on Game Design* (2003).
[23] First articulated in "Why We Play Games" (2004); http://xeodesign.com/xeodesign_whyweplaygames.pdf

As a counterpoint, these are closely related to what play scholar Brian Sutton-Smith, in a skeptical phase, called the "rhetoric of progress" in the study of play and games[24]. This rhetoric names a persistent position, especially prevalent in the study of child and animal play, that attempts to explain the confusingly pervasive phenomenon of apparently-useless play in functional terms related to improvement, learning, and general cognitive development. This rhetoric analogizes improvement at play tasks to improvement towards other desired pedagogical and/or developmental outcomes. It goes almost without saying that some version of this rhetoric has granted force to most of the modern uses of play in formal education.

It is not entirely obvious that fun makes gameplay more effective at producing learning (as opposed to simply making students more likely to engage in self-motivated play). *Every* complex task enables learning and acts as a kind of educational process, including entirely unfun ones. And as Sutton-Smith warns, although play and games may well produce development, the development they produce may be primarily in gameplay skill as such. Fun tasks may well have a relationship to learning, but there may not be a strong reason to believe that this relationship is unique to fun tasks.

Moreover, crucially, even when a game does enable skill development and content mastery portable to areas outside the game, these skills and masteries may not be closely related to learning outcomes desired by educational institutions. This is an especially difficult problem for content-rich areas like science education. It is even more difficult for component skills and subdisciplines of content-rich areas that have only weakly standardized curricula and pedagogical cultures. (All of *Rover Ops'* central interests -- experimental design, scientific project management, and planetary science -- fall into this category, at least at the secondary education level.)

On the other hand, anyone who has deployed a significant number of educational games in classroom settings knows that learners sometimes experience genuine enjoyment in the play of educational games that can be readily directed towards a specific. What games should be taken as exemplary of this phenomenon? Unfortunately, hard data on fun from classroom settings is scarce. There has recently been a certain amount of work on performance assessments of particular educational games' capacity to produce fun. However, these assessments remain distinctly secondary or even tertiary to learning outcome-focused assessments, do not treat learning outcomes and fun as related variables, and rarely assess educational games' 'funness' in comparison to related games.

---

[24] See *The Ambiguity of Play* (1997), particularly ch. 2 and 3.

With this in mind, we will briefly consider a few exemplary educational games on the basis of high commercial success in the general-audience market. Although commercial success in the general-audience market by no means an exact proxy for fun, it is reasonably good, and certainly better than success in the educational market. Classrooms are coercive. Students are in some sense compelled to engage in classroom-based learning activities. They may engage with an unfun game in the same way they would engage with any other unfun learning activity. Conversely, an educational game's success in the broader commercial space gives a strong signal that players consider it enjoyable in its own right.

Zachary Barth's 2011 *SpaceChem* and Masamitsu Shiino's 2001 *Typing of the Dead* are two particularly well-known examples of such commercially successful educational games. *SpaceChem* is a chemistry-themed puzzle game that requires players to script assembly-language- like loops in a visual programming environment narrativized as an extraterrestrial nanotech molecular assembler. *Typing of the Dead* is a reimagining of a 1998 zombie rail shooter from the same studio (Sega WOW), *The House of the Dead 2*. It requires players to quickly and accurately type prompted words and phrases in order to kill zombies.

Why were these incontrovertibly educational games successfully fun enough to become major commercial successes? They do have excellent graphic and sound design, interesting themes, and generally polished user experiences. However, they also align closely with the main line of theoretical approaches to the production of games that are both fun and educational. One very common theme in these approaches is: to make sure that the most basic player activities -- what game designers would call the fundamental mechanics and core loop of the game, the things that players do over and over and are rewarded for -- have some close relation to the desired learning outcomes. One strong way to do this is by making game skill match up with the skills that are important in the real-world knowledge domain. (In certain very strong cases, there might not be a bright-line difference between these skills at all: this is the promise of, for instance, pilot or driver training simulators.) This theme dates back to at least the early-80s work of Thomas Malone at Xerox PARC; in addition to supporting rapid meaningful feedback and satisfying curiosity about a knowledge domain, it is centrally related to the criterion he refers to as "intrinsic fantasy", in which "not only does the fantasy depend on the skill, but the skill also depends on the fantasy." Malone here uses "fantasy" to denote what, in contemporary writing on games, is more often called "narrative," "story," "theme," or "fiction:" some quality that evokes "mental images of things not present.[25]"

---

[25] "What makes things fun to learn? heuristics for designing instructional computer games." *SIGSMALL '80: Proceedings of the 3rd ACM SIGSMALL symposium and the first SIGPC symposium on Small systems.*

The contrary to this principle -- what good educational game designers are not supposed to do -- is to merely add a scoring mechanism to a repetitive learning exercise. This is what Amy Bruckman refers to as the "drill and practice" or "treating children like rats" model. And this is indeed the structure of many of the worst (considered as games) educational games.

These games are precisely in line with this principle. *Typing of the Dead* is a relatively simple example. A typing drill is a repetitive, time-sensitive task involving fine hand-eye coordination. This matches the theme of the game closely: shooting zombies on a linear course is also a repetitive, time-sensitive task involving fine hand-eye coordination. Typing drills themselves are, of course, exemplars of a boring "drill and practice" model, but *Typing's* exploding zombies are not mere positive reinforcement for good typing performance. By linking individual typing acts with high granularity to individual acts of zombie slaughter, *Typing* does not merely gamify or score typing, but meaningfully contextualizes it within a game environment. The game environment, moreover, is not simply a series of rewards, but an exemplary game from a perennially popular game genre that had already proven itself in arcade and console markets. This linkage probably worked so successfully precisely because rapid typing and target acquisition are both *somatically* difficult. Using this methodology for, say, multiplication drills probably could not work as well. (And, indeed, recent efforts in this direction like TapToLearn's 2015 *Math vs. Zombies* have not enjoyed any special success.)

SpaceChem is a slightly more complex instance. It is not really a chemistry-teaching game; chemistry is part of its fictional premise more than its educational outcomes. It teaches assembly programming--much like Barth's more explicit 2015 *TIS-100*, which abandons any pretense of a thematic frame and merely presents players with a series of mysteries that can only be unraveled by retrieving data on a virtual device coded in a purpose-built assembly language. The close compatibility between in-game skill (visual programming with representations of memory registers and so on) and intended learning outcomes (familiarity with low-level programming paradigms) is fairly obvious. Moreover, while, working with simplified programming environments is not as common a core game mechanic as twitch-reflex shooting, it is nevertheless a reasonably well-established fun core gameplay activity, with roots reaching back at least to Jones and Dewdney's 1984 *Core War*. Malone even identified computer programming as an intrinsically puzzle-gamelike activity, offering rapid feedback and subversion of expectations in a way that effectively reveals the underlying structure of the medium.[26]

---

[26] "Toward a theory of intrinsically motivating instruction.", *Cognitive science*, 5(4) (1981).

These examples point towards one possible solution to the "education vs. fun" dilemma: fun is possible in educational gaming when the superstructures of the relevant learning outcome and the game-mechanical requirements have compatible interiors. In other words, play can happen when the constraints on content and action created by learning outcome and game models don't pull player behavior in totally different directions or squish player behavior into a boring drill model. This kind of compatibility exists between typing and zombie shooting, and also between assembly programming and automation puzzles in spatially cramped environments.

We should add that a version of these principles applies to simulation games -- and perhaps even more strongly. This point hardly bears prolonged discussion. Indeed, when the theme or narrative of a game is realistic, to say that it possesses intrinsic rather than extrinsic fantasy elements is nearly equivalent to saying that it works in some way as a model or simulation of its subject matter. Much academic attention to problems of model representation in simulation gaming has focused on the ways in which simulations might be excessively convincing, persuading users that they correspond to reality much more fully or effectively than they actually do[27]. As we noted at the beginning of this section, however, practical problems in simulation game design much more often center on how much granularity of control is practically feasible or aesthetically desirable. This is similar to the dilemma in the learning games space: do the modeled parameters of control over the simulated environment facilitate the desired category of gameplay experience or not? To return to the initial example, the designers of a mercantile wargame like *East India Company* are primarily attempting to focus on tactical positioning, attrition, and resource management -- not, for instance, seamanship or celestial navigation.

Limits of Content-Mechanics Concordance

This sort of high valuation of content-mechanics concordance is more or less the design strategy we followed in developing *Rover Ops*. What, we asked, is the in-game skillset necessary to map the knowledge domains involved in doing planetary science via mobile remote sensing equipment? It didn't seem to us that models like driving games or visual programming exercises were appropriate. Even slow, free spatial exploration didn't seem correct. The important gameable skills to us seemed to be a mixture of strategic resource management and measured investigation.

[27] See for instance McKenzie Wark's *Gamer Theory* and Sherry Turkle's *Simulation and its Discontents*.

However, this strategy ran into a limit of content-mechanics concordance. Consider, briefly: what kind of educational or simulation constraints might be too tight to permit the design of meaningfully integrated play experiences? (Such constraints may also be too *loose*, creating virtual worlds or sandboxes rather than tightly-focused games -- but we are not concerned here with gestures in the direction of *Minecraft* or *Second Life*.)

Duration and *complexity* are two factors that present obvious limits. If a game, or its meaningful play increments, simply takes too long to play -- whether because of accurately modeling elapsed time in simulated events or because of high handling time related to encounters with data and procedures required by simulation constraints or desired learning outcomes -- it is not generally possible to design game mechanics that make this duration acceptable.

Here "too long" may apply to either allotted time for play, as in a classroom setting, or to player tolerance for time expenditure and tedium. If a game's material is so complex as to be impenetrable to its players -- whether because of excessively detailed simulation of events or because of inappropriately-targeted educational requirements -- fun and simulation-congruent game mechanics will not make this complexity tolerable. (Of course, brilliant game mechanics, like brilliant work in other explanatory genres, can make complex topics more digestible, but extremely good explanations only go so far.) What constitutes "too long" or "too complex" is, of course, population- and context-dependent. A game that is entirely approachable for 12th graders may well not be for 6th graders; a game that is perfect for hardcore gamers looking to grind away time is not appropriate for a classroom setting. In *Rover Ops*, we ran into some problems with duration and complexity, but found them generally manageable, principally through extensive hinting and careful management of the amount and complexity of presented text.

One subtler limiting factor is *rigidity*, or, equivalently, *fidelity*. How much flexibility a game requires depends, of course, on a game's core mechanics. Videogames don't necessarily require extreme variation in outcome. A platform game or rail shooter may have only binary variation (win or loss) or something like a single axis of quantitative variation (score). Videogames focused on atmospheric expression or experiential ambience don't necessarily require any variation in outcome at all. However, they nevertheless require a fair amount of flexibility in movement and perception, the parameters on which their sense of play rests.

The limiting question for simulation games in particular is whether the parameter's of the game's model are so tight as to restrict the mechanical operation of the game. For instance, a historical strategy game whose historical fidelity is so high that it didn't permit any ahistorical strategic choices wouldn't be especially entertaining. This example is especially apropos here: it is essentially the situation that *Rover Ops* found itself in. Player choices were essentially constrained to the historical choices of the *Opportunity* team.

The usual way to solve this dilemma in simulation gaming is through the production of plausible counterfactuals. In fact, one might convincingly argue that the production of variably plausible counterfactuals, depending on the genre and style of the simulation, is the central object of most simulation gaming. (Or at least semi-serious simulation gaming properly called: tongue-in-cheek efforts like *Goat Simulator* more correctly work to produce *implausible* counterfactuals.)

"Plausible counterfactuality" here can be readily understood through philosopher David Lewis's criterion of possible-world distance, which he informally describes as:

> "If kangaroos had no tails, they would topple over" seems to me to mean something like this: in any possible state of affairs in which kangaroos have no tails, and which resembles our actual state of affairs as much as kangaroos having no tails permits it to, the kangaroos topple over[28].

For the purposes of *Rover Ops*, we were interested in counterfactual questions like: "If the *Opportunity* team had chosen to investigate this geological feature rather than that geological feature, what telemetry would they have received? How might this different telemetry have altered their subsequent processes of investigation and interpretation?"

Unfortunately, plausible counterfactuals are extremely hard to generate in this setting. Very granular simulation and modeling often focuses on physical systems, and in videogame contexts, their graphical appearances in general. Higher-level simulation -- both in serious and entertainment contexts -- has historically focused on military operations and business management. As such, there are well-established methodologies for answering counterfactual questions like: "what would a particular terrain feature look like from an angle no camera has actually viewed it from?" And even: "what if the British had not retreated into Portugal in 1809?" However, there is no methodology we are aware of for generating a holistically meaningful counterfactual data trail and explicating its significance to learners.

Games that thematize scientific investigation and allow relatively free exploration tend to be satisfied with an extremely high level of abstraction about data and instrumentation, much higher than would be acceptable for our pedagogical objectives. You point the

---

[28] *Counterfactuals* (1973).

instrument, generally a camera or something indistinguishable from a camera, at the object of interest and conclusions, points, or labels of some kind pop out (perhaps after a delay, as in *Extrasolar*).

In other words, we simply did not have enough material to enlarge the space of the simulation enough to allow play. It would have required, at the minimum, a superfluity of "unused" data or some way to convincingly generate artificial data of comparable detail. And to be coherent and useful, we would have needed a narrative methodology for stitching this additional data together into plausible counterfactual scenarios.

One final problem with approaching *Rover Ops* as an educational game was that player failure was not an option, largely due to the conceptual complexity and detail of the material as such. Failure made the gameplay too slow for its classroom setting and gated access to the scientific narrative in ways instructors found undesirable. If failure is not an option, it is quite likely you should not be making a resource-management simulation game. Again, not all videogames require variation in outcome of this kind. But successful games may well require some possibility for failure[29]. Even games like highly forgiving point-and-click adventures (the *Monkey Island* series, for instance), which are formally quite similar to the interactive nonfiction model we have developed, have extensive possibilities for player failure and confusion in discrete parts of the game. Normal (and perhaps normative) play of these games involves extensive instances of failing to see the connection between two disparate inventory items or failing to notice interactive objects in the environment, requiring repetition, exploration of the visual space, and, in general, *playing around with stuff*. We worked very hard to eliminate the possibility of this category of failure: all interactive objects are highlighted and all options are flagged in the UI and explained at length.

Future directions for the EVFE concept

The success of the EVFE to date is a promising suggestion about the possibilities of this kind of interactive nonfiction simulation of remote sensing in planetary science. Future entries in the EVFE series could vastly expand its pedagogical range, potentially covering a very wide range of material about planetary science, scientific instrumentation, and the practical processes of scientific investigation. *Curiosity's* investigations at Gale Crater, for instance, could teach students about sedimentary processes and hydrology. The Cassini-Huygens mission could depict command planning in great detail, clarifying

[29] See Jesper Juul's extended discussion in *The Art of Failure* (2013).

constraints of investigations with very long lag times between commands and reception of subsequent telemetry -- not to mention topics ranging from the geology of the outer planets to general relativity.

We ultimately envision the EVFEs as a freely-accessible series of programs that can be readily used as enrichment exercises in a wide variety of courses. Although it would require additional curricular planning, the EVFEs could perhaps even form the basis for an interactive planetary science textbook aimed at nontraditional learners or general education university courses.

Revisiting the discarded game elements from *Rover Ops* in some separate project is also a possibility. There is, of course, something lost in the shift from exploration-puzzle game to interactive workbook. In particular, rather than experiential immersion in the feedback loop of data analysis, target selection, and data acquisition, students are instead presented with a carefully worked example of this feedback loop. A game version of this requires counterfactual possibilities, procedurally or hand-generated datasets that would enable more free exploration. Simply creating content-accurate maps is the simple part of this problem, using well-established techniques like generating 3D models from heightmap data. It is even possible that we could use some kind of procedurally generated texturing technique to generate instrument data, although this is much more speculative.

Contextualization, guidance, and assessment are the really difficult problems here. The simple lock-and-key or egg-hunt puzzles that most adventure and exploration games are built on are insufficient for the purposes of legitimately demonstrating processes of scientific decision-making without without extremely rich description of their individual nodes. While false-3D views can be easily generated, there is no methodology we are aware of for procedurally generating a holistically meaningful counterfactual data trail and explicating its significance to learners.

One way to shrink this problem space might be to use simple choose-your-own-adventure style narrative forms. The composition of a few plausible counterfactual possibilities presents a smaller difficulty than procedurally generating data and deduction paths. Although choice-based paradigms of this kind don't seem entirely satisfying as vehicles for exploration, it's possible that sufficiently clever writing could produce a convincing and satisfying branching narrative, potentially with 'better' and 'worse' outcomes. The problem of presenting visual materials as immersive as the real mission's data products still exists, though, as a major hurdle.

It's also quite possible that games that attempt to teach these skills should limit themselves to *fictional* environments, as fundamentally simpler testbeds. Games like *Dwarf Fortress* generate worlds with complex ecologies, geologies, and social structures. With additional emphasis on modeling scientific instrumentation, there is no particular reason that scientific investigation could not take place in such a simulation. One problem with this technique is that sophisticated, sandboxy simulation games have notoriously high learning curves. Even with excellent usability design, this methodology would be unlikely to generate games that would fit well within conventional curricular models. Such games would require considerably longer and more focused periods of engagement to produce meaningful learning outcomes.

The intense excitement, focus, and profound creativity produced by successful games makes them an appealing form for new pedagogies. This is especially true in recent years, as an explosion of emerging technologies -- the ubiquity of shockingly cheap laptops and mobile devices, software frameworks that vastly reduce the development costs of graphically rich environments and cross-platform distribution, and so on -- have colonized practically every niche of everyday life with play. But the immersive, participatory multimedia aspects of these technologies are as crucial as their playful aspects. Creative expression, free movement, and painless repetition have limitations, because they are twisty and never quite to the point. One of these limitations is that they are slow. Another is that they rest uncomfortably with historical truth. The EVFE is an excellent science education tool. Its development history is also an object lesson of a general principle: unless creativity or frequent repetition are accommodated in a learning environment, you probably shouldn't put a game in it.