Virtual reality is an immersive, interactive medium that manipulates the senses in order provide users with simulated experiences in computer-generated worlds. The visual design of virtual reality is an important issue, but literature has tended to stress the medium's instructional potential rather than setting forth a protocol for designing virtual environments. Furthermore, virtual reality is often considered a solution in search of a problem instead of a viable instructional tool. To counteract this notion, a study of technology designers, media experts, and instructional designers was conducted to share ideas and generate data regarding what constitutes an Instructional Virtual Environment (IVE), when one should be used, how best to design one, who should be involved in the design process, and how to evaluate one. Three rounds of Delphi surveys were administered by electronic mail, resulting in expert consensus on several points. Virtual reality (VR) is defined by the computer-mediated experiences it provides, and while it must be interactive, it does not require full immersion. VR should be used in educational situations requiring experience in particular settings, especially to present spatial or abstract information because of the advantages of sensorial feedback. The design of IVEs should consider constructivist learning principles, require a high level of user interaction, and support mastery learning. The design process should involve multi-disciplinary teams, including instructional designers, subject matter experts, end users, computer experts, ergonomic specialists, and artists. Finally, evaluators should study IVEs very thoroughly, examining all aspects of the environment and using many different methods. Qualitative data, focus groups, and comparative analysis are recommended directions for future study. (Contains 29 references.) (BEW)
Imagine floating among the planets, touching each of Saturn's rings as you count them, or pushing two molecules together in order to see their chemical reaction. Perhaps you could visit an archaeological dig in Egypt for the afternoon or drop in on a French village to practice your language skills. The scenes described above exemplify educational virtual reality (VR) experiences that could be created through the collaboration of instructional designers, computer scientists and media experts.

Virtual reality is an immersive, interactive medium that relies on computer-mediated manipulations of the visual, aural and tactile senses to provide users with simulated experiences in computer generated worlds. These virtual worlds may reflect either potentially real or fantasy settings.

The visual design of virtual reality, an image-oriented medium, is an important issue. Virtual reality literature reflects both the instructional potential of the medium and the need for guidance in designing the visual interface. There is no current protocol for designing instructional virtual environments (IVE's). Nilan (1992) states that the current orientation of VR research and development toward the technology's ability to simulate accurately a physical environment suggests that VR is a solution looking for a problem. Steuer (1992) believes that existing VR research has neglected to provide an understanding of many aspects of the technology, including effects, design considerations, and aesthetics. To use VR successfully for achieving instructional goals there are still many aspects of the medium that we must explore.

Virtual reality's status as a new medium means that few people can claim to be experts in its specialized applications. Helsel (1992) sums up the current situation: "To date, the agenda for VR has been set by the computer science community. . . . Yet, education has a tremendous wealth of information and experience to bring to the VR curriculum" (p. 42). Thus we need the collaboration of those experts who are knowledgeable in specialized aspects of the medium to set the research and development agenda of instructional virtual environments.

A study of technology designers, media experts and instructional designers was conducted to share ideas and generate data regarding how the visual interface of IVE's should be designed to maximize the success of this educational medium. Participants were asked to comment on: (a) what constitutes an IVE; (b) when IVE's should be used; (c) design principles for constructing IVE's; (d) who should be involved in the IVE design process; and (e) how IVE's should be evaluated.

Methodology

Problem Statement

Virtual reality as an educational medium is a means to an end. It can be seen through an
input-process-output paradigm in which the initial input is data from the human user, the process determines and presents appropriate instructional content mediated by a computer engine, and the output is increased human knowledge. The current technological focus of VR has resulted in the notion that it is a solution in search of a problem rather than a tool in the instructional technologist's toolbox. Roblyer (1993) finds this approach problematic: "Without a clear vision, we have no hope of getting to where we want to go, and the lack of vision, as the prophets warn us, could prove fatal" (p. 34). Thus, a blueprint or conceptual framework consistent with the theories and practice of instructional design is needed to guide the future of virtual reality.

Research Questions

The general question for this study was, "What principles should be considered when designing instructional virtual environments (IVE's)?" The specific research questions for this study were:
1. How do you define virtual reality?
2. How should virtual environments be designed to facilitate learning?
3. Who should be involved in the design of virtual reality, and what form should their involvement take?
4. How should educational virtual reality experiences be evaluated?

These questions were asked of participants in a modified Delphi study. The Delphi technique is a research method involving multiple surveys of a population. It is based on the adage "Two heads are better than one" (Dalkey, 1972, p. 15), with the objective of reaching group consensus on survey items.

The questions became more defined and new issues were raised during the Delphi process, which requested that respondents make statements related to the research questions in the first round and then reflect those responses in subsequent survey instruments. The questions given here, then, served as an initial guide for the Delphi participants.

Participants

Participants in this study represented three fields of research, instructional design, computer science and media. Instructional designers were chosen for their experience in designing instructional content. Computer scientists and media experts were chosen for their intimate knowledge how technology works and media design, respectively. Specific selection criteria required that participants be knowledgeable in one of these three fields as evidenced by their job positions, research, and educational backgrounds. Additional qualifications for all participants were access to and ability to use electronic mail through the Internet.

Data Collection Procedures

This study involved three rounds of Delphi surveys. The first round generated predictions and statements pertaining to the four broad research questions. The second round generated participant ratings of their agreement with these statements. The third round attempted to reach group consensus on the ratings of these statements.

Description of Conditions

All participant communication was conducted solely through electronic mail. Contact was made by the researcher only, and was limited to the initial request for participation and messages pertaining to and including three rounds of surveys.

Instruments

The Round One survey consisted of the four broad research questions developed from information generated by a literature review. Its intent was to elicit responses for use as survey items in succeeding data collection instruments.
The Round Two and Round Three surveys consisted of 97 items divided into five major sections:

I. Defining Virtual Reality (10 items)
II. Using Virtual Reality In Education (14 items)
III. How to Design Instructional Virtual Environments (19 items)
IV. Who Should Design Instructional Virtual Environments (32 items)
V. Evaluating Instructional Virtual Environments (22 items)

All of the Round Two and Three items were derived from responses to the Round One survey. Samples of Round Two and Three items from each of these survey parts are in Table 1.

<table>
<thead>
<tr>
<th>Survey Part</th>
<th>Sample Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>An element of VR is the interaction of human senses and computer peripherals</td>
</tr>
<tr>
<td></td>
<td>An element of VR is a computer simulations in which the user controls point of view</td>
</tr>
<tr>
<td>II</td>
<td>VR should be used in education to teach spatial information</td>
</tr>
<tr>
<td></td>
<td>VR should be used in education only when cost effective</td>
</tr>
<tr>
<td>III</td>
<td>IVE's should be systematically designed</td>
</tr>
<tr>
<td></td>
<td>IVE's should be fully immersive</td>
</tr>
<tr>
<td>IV</td>
<td>IVE's should be designed by a team Instructional designers should be involved in the design of IVE's</td>
</tr>
<tr>
<td>V</td>
<td>IVE's should be evaluated to determine evidence of learning</td>
</tr>
<tr>
<td></td>
<td>IVE's should be evaluated using user observation</td>
</tr>
</tbody>
</table>

Table 1: Sample survey items

Participants were asked to rate each item on a given Likert scale. Part I. Defining Virtual Reality: contained definitions or parts of definitions and requested that participants indicate how these statements compare to their personal concepts of virtual reality, ranging from "1" (Much broader) to "5" (Much narrower). Parts II - V asked participants to classify their feelings about statements according to the following scale, originally used by Holden and Wedman (1994):

1 - Strongly disagree
2 - Disagree
3 - Disagree with reservation
4 - Agree with reservation
5 - Agree
6 - Strongly agree

Respondents also were encouraged to add any comments they felt necessary to their responses.

Data Analysis

Data generated by the final survey round were numerical ratings from Likert scales. The mean, median and quartile deviation again were calculated to determine definition of "virtual reality" (Part I only), importance of concept (Parts II - V only), and level of consensus for each survey item.

The Part I responses rated elements of possible definitions of virtual reality. The median responses were used to determine whether or not the study group and field subgroups felt a particular phrase broadly (median $\geq 2.5$), closely (median $< 2.5$ and $> 3.5$) or narrowly (median $\geq 3.5$) applies to virtual reality.

Parts II - V asked participants to provide their level of agreements with certain phrases. Those items with which the study population either agreed or disagreed strongly (median $> 2$ or $< 5$) were considered extremely important to the design of instructional virtual environments. Items rated with solid agreement or disagreement (medians $\leq 2$ and $> 3$ or $< 4$ and $\geq 5$) were considered important. When reservations were expressed on either agreement or disagreement with a statement (median $\leq 3$ and $\geq 4$), that item was considered to be of negligible or uncertain importance to instructional virtual environments.

Level of consensus was determined in accordance with the Delphi studies of Faherty (1979) and Holden and Wedman (1994). Using quartile deviation as the indicator, items were divided into those achieving high, moderate and
A statement was said to have achieved high consensus when the quartile deviation was $\leq .50$ for Part I and $\leq .60$ for Parts II - V or, in other words, when the deviation of responses in the interquartile range was less than one point on the Likert scale. Moderate consensus was indicated by quartile deviations $> .50$ and $\leq 1.00$ for Part I and $> .60$ and $\leq 1.00$ for Parts II - V. When quartile deviation exceeded 1.00 in any survey part, the item was classified as low consensus.

Results and Conclusions

The 97 items surveyed in five categories yielded 65% items attaining a high level of consensus and 31% with a moderate level of consensus. From these items, a total of 14 conclusions were made.

Virtual Reality Definitions (Part I)

It was important to have survey participants consider their own definitions of virtual reality for two reasons. First, it made each participant carefully consider properties and elements of the medium they were about to comment on in the rest of the study. Second, it provided a basis for understanding respondent ratings and comments on other survey items.

There were ten items in this section, of which five achieved high consensus, three achieved moderate consensus and two did not reach consensus. For most items the median ratings for all sub-groups fell within the range of one point. The study population did not find any of the given definitions too narrow, although definitions specifying only imaginary worlds and synthetic experiences were considered too broad.

The participants’ definitions of virtual reality centered around computer simulations and human-computer interfaces. The following elements were considered additional parts of an appropriate definition: (a) involves the user in a realistic 3-D world, (b) exists on a continuum from marginal (computer desktop) to a full-immersion alternate reality system, (c) user controls point of view, (d) user feels presence, and (e) user interaction is required in the form of direct object manipulation.

Two conclusions were drawn from the definition section of the survey:

1. Virtual reality is defined by the computer-mediated experiences that it provides for users.
2. Virtual reality, while interactive, does not require full immersion.

This first conclusion is significant because the term “virtual reality” often is used in reference to a certain hardware-software configuration (Heim, 1993). In this study, not one of the participants used specific equipment to describe the medium. Instead, their definitions were in agreement with Ferrington and Loge (1992) and Pimentel and Teixeira (1993), who refer to VR as a computer-mediated experience. While respondents included different elements in their definitions, such as 3-D worlds and direct object manipulation, all agreed upon the basic necessity of an experience provided by a human computer interface. The study population’s strong agreement on this issue suggests that as the technology develops the definition of the medium is becoming more conceptual and less equipment-dependent.

The second conclusion addresses a key variable of virtual environments. Immersion exists on a continuum (Pimentel & Teixeira, 1993) and it is unclear from the literature exactly when the level of immersion is deep enough for the computer-mediated experience to cross over from simulation to virtual reality. While participants did not overtly address the location of this crossover point, they did agree that (a) immersion exists on a continuum from marginal to full, (b) mere computer simulation or computer-mediated communication is not immersive enough to be considered virtual reality, and (c) full immersion and full sensorial stimulation are not essential to instructional virtual environments.

As a variable, immersion should be determined by need; not all IVE’s will require the highest level to reach maximum efficacy. For example, an IVE designed to teach molecular
docking would rely heavily on visual stimuli, but little on aural factors. Cost of production and use increases alongside level of immersion, and participants suggested that virtual reality should be used when cost effective. Inclusion of extensive aural stimuli in this example scenario might cost a significant amount without greatly increasing the efficacy of the environment. For these reasons IVE’s should be evaluated carefully during the design phase to determine the immersion level that is appropriate to the content and context of the instruction.

When to Use VR for Instruction
(Part II)

This part of the survey contained 14 items with the consensus being high for eight items, moderate for five items and low for one item. For most items the median ratings for all sub-groups fell within the range of one point.

Overall, participants felt that VR is not a learning tool to be used in all situations; more specifically, they indicated that IVE’s should be used for more than just gaining user attention and should be cost effective. This data resulted in the following conclusion:

3. Virtual reality is an appropriate medium for some educational situations. This statement supports the writings and research of instructional designers and computer and media technologists, such as Ferrington and Loge (1992), Fritz (1991), Helsel (1992), and McCluskey (1991). They propose VR as a useful instructional medium because of its immersive qualities and interactive nature; participants highlighted the same features in their ratings and comments, noting also that there are some situations in which VR might not be the best instructional medium. Lanier (1992) similarly finds in an interview with Frank Biocca about VR’s future that there are some things students should be learning in the physical world rather than in instructional virtual environments.

Based on the conclusion that virtual reality is an appropriate instructional medium for some educational situations, other conclusions can be made. The data collected support the following statements about when VR should be used in education:

4. Virtual reality should be used in educational situations requiring experience in particular settings. This conclusion was based upon strong participant ratings for usage in the following situations: (a) when dangerous real-life scenarios are involved, (b) when a setting that would otherwise be inaccessible is involved, and (c) when detailed demonstrations or illustrations are needed. This finding supports Bricken (1990), who states that experience is the basis of the virtual reality learning process. Ferrington and Loge (1990) suggest VR as an excellent medium for learning about hostile environments and dangerous materials; Fritz (1991) similarly recommends it for medical and flight instruction. Helsel (1992) mentions using VR to visit a historical era. These examples all indicate that the technology’s potential for experience-based learning is high.

5. Virtual reality should be used for presenting spatial or abstract information. Participants rated both of these often-related areas highly. Virtual environments can remove the abstract nature of a topic and increase spatial understanding through sensorial stimuli. Abstract concepts are more easily learned when presented in a visual form (Krueger, 1991). Helsel (1992) and Regian, Shebilske and Monk (1992) support the use of imagery for learning because of the brain’s superior visual capabilities. Fritz (1991) discusses spatial opportunities, such as a dance training program in which the dancer is rewarded with pleasant music for correct moves and discordant sounds for incorrect ones. In all of these cases, the belief in sensorial feedback rather than just verbal presentation is strong.

Design Philosophy and Features
(Part III)

This section contained 19 items, 17 of which attained high consensus levels; the remaining two attained moderate consensus. Most items were rated important and the median responses for different sub-groups fell within a
one-point range except where noted in this section.

**Design Philosophy**

Participants felt that a systematic, iterative design process, with objectives established from the beginning, was important. Responses further indicated that the IVE should provide learning in the context in which it is to be recalled. Isolating the concepts to be learned was considered appropriate when important contextual elements are retained. Respondents did not feel that IVE's should be designed necessarily according to the same principles as other media. Comments on this item noted that it may be too soon to tell whether or not such principles are relevant to IVE's and that we should not view VR as merely "an extension of the old technology."

Support for constructivist learning was rated important overall, but the sub-groups of the study, all of which attained high consensus, placed significantly varying levels of importance on this item. Instructional designers felt most strongly, with a median response of six; computer technologists felt most uncertain, with a median response of four.

Participant ratings on how the design process should be approached resulted in two conclusions:

6. *Instructional virtual environments should be created according to the principles of instructional design.* Systematic design, iteration and educational objectives were all considered important parts of the design process. Gagné, Briggs and Wager (1992) state that effective learning must be planned, with goals determined at the beginning. The iterative nature of instructional design is exemplified by Dick and Carey's (1990) model, which provides a continuous option to evaluate and then return to a previous design stage for revision. These principles are supported in the literature of virtual reality by Ferrington and Loge (1992) and Henderson (1991), who feel that motivated, well-constructed virtual environments will provide the most useful learning tools.

7. *The design of instructional virtual environments should consider constructivist learning principles.* Participants were in support of environments that facilitate individual construction of meaning from presented stimuli. This item, along with respondents' suggested experience-based nature of virtual learning, shows the need to explore constructivism when designing instructional virtual worlds. Virtual reality researchers Bricken (1990) and Winn (1993) heavily support this claim. Lebow (1994) states that constructivist values can be used to guide the instructional design process, suggesting that the synergy of instructional design principles (Conclusion 6) and constructivism (Conclusion 7) will provide effective virtual learning experiences.

**Design Features**

A virtual environment that allows users to see and correct their errors was considered very important. Other elements rated highly included (a) user-friendliness, (b) a high level of required user interaction, (c) built-in tools for interaction and exploration, (d) a way for teachers to enter and interact within the IVE, and (e) allowance for user experimentation and expression.

Providing users with the ability to change environment and object parameters fell at the bottom end of important design features. Comments indicated that this low rating was related not to the item's value as a design factor but rather conditions of its use. Participants felt that environment and object parameter changes should be permitted carefully, judiciously, and only when relevant to the learning experience.

Full immersion and full-sensorial stimulation were not considered essential to instructional virtual realities. In both cases, computer technologists were the sub-group most against mandating these elements. Participant remarks stated that they are variables which need to be addressed on a case-by-case basis, depending heavily on the instructional goal of the learning situation.

From these results, the following conclusions about design features have been made:

8. *Instructional virtual environments should require a high level of user interaction.*
This conclusion was based on participant indications that: (a) much interaction is desirable, (b) tools for interaction and exploration should be included in the virtual environment, (c) teachers should be able to communicate with students in the virtual environments, and (d) user experimentation should be encouraged within the virtual environment. Additional comments indicate that an IVE with a low level of interaction would not be taking full advantage of the medium. Interaction is cited by many researchers (Aukstakalnis & Blatner, 1992; Ferrington & Loge, 1992; Steuer, 1992) as an important part of virtual environments, with particular issues including navigation and user control (Pimentel & Teixeira, 1993). As a medium recommended for experience-based learning (Conclusion 4), virtual reality should not only permit but require users to maintain an active position.

9. Instructional virtual environments should support mastery learning. A potential feature of IVE’s that was rated highly is a user feedback system that shows error and allows the user to keep trying until the correct response or performance is attained. Also considered important was the ability of a virtual environment to adjust to user needs through user- and task-centered analysis. These items support mastery learning, as defined by Gagné et al. (1992), by allowing users the opportunity to learn from their mistakes until they accomplish the performance objective within an environment constructed to provide extra time, determine missing prerequisite skills and alter the example or approach to the material as needed.

Virtual Environment Designers (Part IV)

This survey section, which asked what types of people should be involved in the design of IVE’s and what their roles should entail, contained 32 items. Seventeen items attained high consensus levels, 14 were moderate, and one remained at a low consensus level. The responses of each sub-group were closely clustered around the overall median, except as noted within this section.

Team-based design, with members representing a broad spectrum of disciplines and having primarily educational aims, was considered important. Participants agreed that those who choose to get involved in IVE design should not let themselves be limited by the technology. Instructional designers, subject matter experts and end-users were considered the most important members of a design team, followed by computer experts, ergonomic specialists and artists. The potential contributions of psychologists, investors and computer hackers were viewed as less necessary to the design process.

Participant indications that many skills and different types of input are required when designing IVE’s led to the following conclusions:

10. Instructional virtual environments should be designed by multi-disciplinary teams.

11. Core members of an IVE design team should be instructional designers, subject matter experts and end-users.

12. Computer experts, ergonomic specialists and artists may have useful contributions to make to IVE design teams.

Helsel (1992) states that both educators and computer technologists have much to bring to VR design. It is common instructional design practice as well to include subject matter experts (SME’s) and users at some point in the process (Dick & Carey, 1990). Pett and Grabinger (1991) suggest the addition of an instructional producer, or someone who has the production skills to execute the design. Shrock (1991) cites a designer-SME-producer triad as the core of a design team.

All of these people have skills necessary to the design of instructional virtual environments. A core of designer-SME-users can examine closely the instructional content and context and evaluate its effectiveness. While Pett and Grabinger (1991) and Shrock (1991) both recommend having a team member with production skills, participants did not find such a person, in this case a computer expert, to be a necessary a member of the design team. Their comments, however, indicate that it may be a partially semantic issue. Many respondents noted that a computer expert would be necessary for
execution of the project but should not assume primary responsibility for the initial conception or decisions on how learning should occur.

Evaluating IVE's (Part V)

This section of the study addressed two main issues: (a) what aspects of IVE's should be evaluated, and (b) with what methods they should be evaluated. Sixteen of the twenty-two survey items attained high consensus, with the remaining six reaching moderate levels. Overall results indicated that IVE's should be evaluated for many purposes and through many means. Median responses for all sub-groups were close to the study population medians, with instructional designers consistently rating items slightly higher than the median.

Twelve aspects of IVE's were recommended for evaluation in the Round One survey. All were considered important evaluation foci by the participants, with the highest ratings being placed on studying an IVE's ability to facilitate learning and evidence of learning when an IVE is used.

The Round One survey resulted in ten suggestions for how IVE's should be evaluated. The only items not rated important were traditional evaluation methods and written exams. Participants stated that traditional methods may not be appropriate since they were designed for other media and therefore may not accurately measure the learning that takes place in an instructional virtual environment. Written exams were considered necessary for now, due to their prevalence in our society, but some participants felt there were better evaluation techniques that could be used.

The remaining eight evaluation methods were rated important, but many participants indicated that their response in favor of any one system did not mean that it should be the only one used. Those who commented on how IVE's should be evaluated expressed a belief that many methods should be used as each would produce different information.

Conclusions were made about two aspects of IVE evaluation:

13. It is important to evaluate all facets of instructional virtual environments. Participants recommended looking at all aspects of IVE's to determine their success. These aspects include user performance, user perception of IVE, and cost effectiveness. This finding is in agreement with many virtual reality scholars. For example, Auld and Pantelidis (1994) suggest evaluating cost, feasibility in a school setting and educational value. Scriven (1991) embraces the notion that evaluations can be conducted in a multitude of ways for a multitude of purposes. Particularly in light of this technology's youth, it follows that extensive evaluation of every aspect and application is in order.

14. Instructional virtual environments should be studied using many evaluation methods. As a natural progression from the many potential evaluation purposes (Conclusion 13), participants indicated the importance of numerous evaluation methods. While responses tended to eschew traditional methods such as written tests in favor of observation and interaction-based evaluations, comments indicated that any method is better than none at all and that each method will have its own merits due to the different types of information it will generate. This finding has support in evaluation scholarship (Scriven, 1991) as well as in instructional design (Gagné et al., 1992). Ferrington and Loge (1992) suggest evaluation of student learning in a virtual environment through observations, interviews, personal reflection or user data generated by the environment. Virtual environments to date have been evaluated through such methods as user performance (Regian et al., 1992), observation (Pimentel & Blau, 1994), and survey (Osberg, 1992).

Recommendations for Future Research

As is often the case with research, for each conclusion that was made many new research paths were uncovered. There are several recommendations that can be made for the directions of future studies.
This study began with very broad research questions and, through the Delphi process, narrowed large issues into somewhat smaller ones. However, there is still a great deal of room for each survey item to be explored in depth. Much of what needs to be done is collaborative in nature; ideas need to be shared and built upon, until a well-functioning systems evolves.

The following are suggestions for types of future studies:

1. A collection of qualitative data, through interviews or long answer survey instruments, from field experts on small, specific topics within IVE design. Topics that could be explored in this way include the roles and tasks of specific design team members, how interfaces for teacher-student interactions within IVE's should be designed, and efficacy of different evaluation methods.

2. A focus group of various field experts, which would permit the free and immediate exchange of ideas, upon which new research avenues could be discovered. This type of study could generate a great deal of qualitative data as well as further the research of the focus group members through ideas gained in the collaborative process.

3. A comparison study from observations of users within different IVE's could help determine which designed elements are useful in practice.

These are just a few examples of future studies that could be conducted. Each, in turn, would generate more new studies, bringing us closer to a conceptual framework for designing instructional virtual realities. For now, the most important thing is to keep studying this area, sharing ideas, and building on each other’s work.

References


promise (pp. 19-26). Westport, CT: Meckler Publishing.


