Intuitive Navigation in the Targeting of Radiation Therapy Treatment Beams

TR94-025

1994

James Che-Ming Chung

Department of Computer Science
The University of North Carolina
Chapel Hill, NC 27599-3175

Intuitive Navigation in the Targeting of Radiation Therapy Treatment Beams

James C. Chung
(Under the direction of Frederick P. Brooks, Jr.)

ABSTRACT

This research focused on the possible benefits to be gained from using the intuitive navigation made possible by head-mounted displays (HMDs) in the targeting of radiation therapy treatment beams. A tumor surrounded by various types of healthy tissue can present a very complex 3-D situation which must be thoroughly understood for effective treatment to be possible. Conventional 2-D treatment planning suffers from reliance on 2-D diagnostic imaging and dose computations to represent an inherently 3-D problem. Current state-of-the-art treatment planning systems use 3-D models of the patient and compute 3-D dose distributions, but complete exploration of the solution space of possible beam orientations can be hindered by not-so-intuitive navigation controls. The thesis of this dissertation was that the head-mounted display will permit freer exploration of the patient anatomy and range of possible beam orientations, thereby resulting in more efficient treatment planning.

To that end, I developed a new, intuitive navigation mode, which used the orientation of a HMD to determine the direction from which its wearer viewed a 3-D model of the patient's anatomy. When the user looked down, he viewed the anatomy from above. Looking up yielded a view from below, and turning his head horizontally in a circle completely scanned around the model. Although it did not have a real-world metaphor, this mode (dubbed Orbital mode) proved to be surprisingly easy to use and well-suited to the task of targeting treatment beams. I compared Orbital mode with more conventional joystick-based rotation in a user study in which radiation oncology professionals designed treatment beam configurations for lung tumor cases. Slightly faster performance was found with Orbital mode, although there appeared to be no significant difference in the quality of the beam configurations. Movement in Orbital mode was more omnidirectional than with the joystick, most likely due to the mechanical construction of the joystick, which preferentially encouraged left-right and forward-back deflection. The overall conclusion from this study was that HMDs in their present state are not appropriate for clinical use, but with future developments in HMD technology, the benefits of intuitive navigation may appear in mainstream radiation treatment planning.

Acknowledgements

I am immeasurably indebted to a great many people for helping me along in my pursuit of this Ph.D. First and foremost is my advisor, Professor Frederick P. Brooks, Jr., whose wisdom and advice extended beyond academic affairs. Dr. Brooks's model of professional conduct and responsibility will stay with me and guide my own actions in the future. Dr. Julian G. Rosenman served as a valuable resource for information and advice concerning radiation oncology, and provided many useful comments on my dissertation. Professor David V. Beard engaged me in interesting discussions of experimental design and also served as a dissertation reader. Professors Henry Fuchs and Stephen M. Pizer always kept me on my toes with sharp questions and interesting insights.

I of course owe much to my subjects, whose data and comments gave me much to think about in the course of this research.

The staff of the Department of Radiation Oncology at UNC Hospitals was especially helpful. Mitch Soltys was my software guru, providing pointers on how to alter existing Radiation Oncology programs to suit my needs. Aziz Boxwala assisted me in locating data for case studies. System administrator Jeff Lewis made sure I had what I needed. Medical Physicist Kathy Deschesne went out of her way on numerous occasions to accommodate my observation of the clinical practice of radiation oncology.

Professor Keith Muller of UNC's Department of Biostatistics kept my head above water during the data analysis. He cleared up many statistical fuzzy points for me and provided reassurance that what I was doing was the right thing to do.

My fellow members of the Head-mounted Display Project were invaluable sources of support and technical assistance. I hope to continue these relationships as colleagues and friends as we make our way into the real world. Nothing would ever get done in the UNC Graphics Lab without the nuts-and-bolts assistance of GLab keepers John Hughes and David Harrison.

Last, but not least, I owe much of my successful completion of the Ph.D. program to my wife Nancy Striniste. Her level-headedness keep me on target at all the crucial points, and she sacrificed much to indulge me this endeavor. My daughter Abbey joined us about halfway through this journey and her presence has made it all worthwhile. She will go down in history as the creator (at age 2) of the "blankie" paradigm for interacting with virtual worlds.

I wish to gratefully acknowledge the financial support I received through the years from the following agencies: National Science Foundation, Defense Advanced Research Projects Agency, Office for Naval Research, Digital Equipment Corporation, the National Institutes of Health, and the UNC Department of Computer Science.

Table of Contents

List of Tables

List of Figures

Abbreviations and Symbols

°	degrees
3-D	three-dimensional
AP-PA	antero-posterior—postero-anterior
BEV	beam's-eye view
cm.	centimeters
CT	computed tomography
DRR	digitally-reconstructed radiograph
HLS	hue-lightness-saturation
hr.	hours.
lb.	pounds
M.D.	Doctor of Medicine
MeV	megavolts
mg.	milligrams
ms.	milliseconds
RTP	radiation treatment planning
SAD	source-axis distance
sec.	seconds
UNC	University of North Carolina
VIMS	visually-induced motion sickness

Chapter 1

Introduction

At the 1965 Congress of the International Federation for Information processing Ivan Sutherland first proposed what he termed the "Ultimate Display." (Sutherland 1965) While such a display, which would essentially "be a room within which the computer can control the existence of matter," can only be considered fantasy in today's world, Sutherland also discussed an intermediate step, a *kinesthetic display*. A kinesthetic display would respond to positions and movements of the human body, varying its information content according to such parameters as where the viewer is looking.

Several years later Sutherland did build a head-mounted kinesthetic display (Sutherland 1968), and although Clark found the system somewhat useful for designing three-dimensional surfaces (Clark 1976), it has been only recently that advances in computational and image generation technology have made possible head-mounted display systems that are inexpensive enough to be accessible by a large market and powerful enough to raise the notion that such systems may be useful for real-world tasks.

Within the past few years, the promise of this new technology has captured the imaginations of people the world over. *Virtual Reality*, as this area of endeavor has come to be known, has become a hot buzzword and the focus of much attention in the print and broadcast media, as well as popular cinema. Visionaries have proclaimed the advent of Virtual Reality to be a major revolution in human communication that will have tremendous consequences for human culture and society. People of all walks, from artists to Wall Street financial wizards, ask themselves, "What can Virtual Reality do for me?" The nature of this question illustrates the oft-heard criticism of Virtual Reality as a solution looking for a problem. The technology has arrived, but what is it good for? The promises are great, but can they be fulfilled?

Frederick P. Brooks, Jr., Kenan Professor of Computer Science at the University of North Carolina at Chapel Hill, expresses these reservations regarding Virtual Reality as, "Can we do it? If so, so what?" (personal communication) Today's commercially available head-mounted display technology, readily available to researchers and developers with modest budgets, allows us to answer the first question with a qualified yes. Yes, we are able to immerse a user in a computer-generated environment, but unlike the builders of high-end flight simulators, we are not yet able to fully entice that user into forgetting about the real world and accepting the virtual world as real. Such suspension of disbelief will become easier with the inevitable advances in technology, and each passing year will bring us closer to attaining a sense of presence in the virtual environment. It is merely a matter of time.

Brooks's second question is somewhat more difficult to answer. Granted, we can do it—kind of—but so what? Assuming that it is possible to seamlessly immerse a person in a computer-generated environment, what possible benefits might be derived from such interaction between human and computer? The research reported here addressed this question by exploring the use of a head-mounted display in the targeting of radiation beams for cancer treatment. I sought to determine if a head-mounted display could be used to advantage by radiation therapists in designing treatment plans for their patients. A head-mounted display enables its user to make use of kinesthetic and proprioceptive information that is not available to a user seated in front of a conventional workstation, and thus makes possible methods for exploring and navigating three-dimensional spaces that are more intuitive and natural. For the three-dimensional situation presented by a tumor growing in the midst of normal, healthy tissue, I anticipated that the radiotherapist using a head-mounted display would be able to more readily and more fully explore the complete range of possible beam configurations. The end result would be better treatment plans based upon beam configurations that do a better job of irradiating the tumor while avoiding radiosensitive healthy tissue.

1.1. Radiation therapy planning

Radiation therapy treatment planning is a complex procedure, the goal of which is to determine the best use of radiation that will effectively control of the tumor while minimizing the complications resulting from the treatment. (Perez and Purdy 1992) In this research, I was concerned only with *external beam therapy*, in which the radiation source is outside the patient's body; I ignored *brachytherapy*, in which the radiation source is embedded in the patient's body near the tumor.

The treatment planning goal of having external beams irradiate the tumor and a conservative safety margin around the tumor, while avoiding the healthy surrounding tissue as much as possible, presents a major challenge to the radiotherapist. Because tumors may often wrap themselves around organs and have micro-extensions that snake away into the surrounding tissue, satisfying the constraints to the best possible advantage for the patient requires a thorough understanding of the complex three-dimensional situation in the neighborhood of a tumor.

Conventional planning of radiation treatment is based upon the use of a radiation treatment machine simulator to provide therapists with information regarding what tissue will receive radiation from a proposed treatment beam. This information comes in the form of two-dimensional projection radiographs of the anatomical structures in the proposed treatment field, taken from the point of view of the beam source. (Mosher et al. 1988) Since the radiographs are perspective projection views of the anatomy, so-called *beam's-eye views* (Goitein, Abrams et al. 1983), the boundary of the beam can be represented by a simple closed line on the image, and all anatomical structures found within the beam outline will receive radiation[1] Unfortunately, radiographs taken for beam directions that are not the standard anterior-posterior or lateral views are somewhat more difficult to interpret—three-dimensional structure is not so easily reconstructed from the two-dimensional image for views that are unfamiliar to the physician. As a result, there is a bias among radiation therapists to rely on cardinal angle beam approaches (i.e. anterior-posterior and lateral approaches), and this is almost sure to lead to suboptimal radiation treatment plans. (Rosenman 1991)

Within the past decade, three-dimensional treatment planning has been devel-

---

1. Actually, scattering of the radiation will irradiate some amount of tissue outside the beam boundary. Additionally, scattering, differential absorption and the divergence of the beam will produce a non-uniform dose distribution.

oped in response to the drawbacks of two-dimensional planning. As imaging and display technology improved, along with computational power, it became feasible to build interactive systems that helped the radiotherapist understand the wealth of three-dimensional information available. Chin et al. explored the advantages of non-traditional beam orientations, and suggested that any three-dimensional planning system should comprise three modules: an interactive patient data evaluation module to analyze and integrate various sources of diagnostic data, a module to compute doses using three-dimensional patient and beam geometry data, and an interactive dose-display module. (Chin et al. 1985) Goitein and colleagues developed the basic tools found in most three-dimensional treatment planning systems today: anatomy contouring, beam's-eye view, back projection from beam aperture onto CT slice, and projection through the CT sections from any arbitrary point. (Goitein and Abrams 1983; Goitein, Abrams et al. 1983)

According to Purdy and Emami the advantages provided by three-dimensional treatment planning include better targeting of the tumor and better conformation between dose and the target volume, better quantification of normal tissue tolerances arising from improved delineation of normal tissues, and the ability to explore novel treatment strategies that permit higher doses within the tumor while minimizing detrimental effects to normal tissue. (Purdy and Emami 1992) Despite these advantages, three-dimensional treatment planning has yet to find its way into the mainstream of clinical practice. Purdy and Emami cite such problems as incomplete quality assurance of three-dimensional planning systems, inefficient delineation of critical anatomical structures, inadequate dose calculation algorithms, inadequate criteria for evaluation and comparison of treatment plans, and inadequate patient positioning and immobilization techniques, as significant obstacles to the routine use of three-dimensional treatment planning.

1.2. Radiation treatment planning and virtual environments

Current state-of-the-art three-dimensional treatment planning systems all use high-resolution graphics workstations as the visual interface between the therapist and the three-dimensional computer model of the patient's anatomy. The therapist views the anatomy on the screen and explores the anatomy by manipulating it with a variety of translational and rotational devices. Stereo images may be used to enhance the three-dimensional percept.

With these systems the therapist must rely on visual cues only to maintain bearing[2] as the patient's anatomy can be manipulated into any arbitrary position and orientation relative to the treatment planner. Personal observation of radiation oncologists at UNC Hospitals using their treatment planning program *xvsim* revealed that despite being familiar with general human anatomy, even a trained physician could have some difficulty in re-establishing his bearing relative to the patient's anatomy when presented with a new view of the anatomical model. At the time *xvsim's* update rate was so low that user's saw rotations as discontinuous jumps from one view to another. Smooth motion would no doubt have helped the physicians. In addition, navigating through the patient's anatomy often proved to be problematical, as desired changes in position and orientation had to be decomposed into a series of knob rotations and mouse movements. Such additional cognitive strain[3] interferes with the regular problem-solving process of exploring prospective beam configurations.

Use of a head-mounted display promises to aid the task of beam-targeting, because it provides a natural means of navigation that allows the radiotherapist to concentrate on the main task of exploring and evaluating prospective beam configurations. The head-mounted display adds kinesthetic information to the visual information already available to the treatment planner. If used smartly by the system designer, this kinesthetic information can greatly lessen the cognitive load, for human beings are trained from birth to respond to kinesthetic feedback while navigating their environment.

An ideal problem-solving strategy is one that both completes the task as efficiently as possible and minimizes cognitive strain. (Bourne et al. 1986) Natural, intuitive navigation helps in both respects. Less time is spent figuring out where one is and how to move to somewhere else, and more time is spent on the evaluation of different possible solutions. In addition, it was the thesis of this research that intuitive navigation would lead to fuller exploration of the solution space, which in turn would lead to the discovery of better solutions, on average, than would have been possible without the intuitive

---

2. In this document the term *orientation* will refer to the rotational relationship between two objects or frames of reference, and *position* will be used to indicate the translational relationship. *Bearing* will refer to a person's understanding of his situation relative to an object or to his surroundings, which is somewhat akin to the psychological use of the term *orientation* as one's awareness of one's surroundings in relation to one's self.

3. Mental effort or stress on one's information processing capacity. (Bourne et al. 1986)

navigation.

In the realm of radiation treatment planning, this means the beam targeting stage will produce better beam configurations, from which better treatment plans can be designed. Better treatment plans will result in higher success rates of tumor control and fewer side effects from radiation treatment.

1.3. This research

The research reported in this dissertation was aimed at evaluating the effect of natural navigation provided by head-tracked head-mounted displays in the context of targeting radiation treatment beams. This evaluation involved a user study that investigated the effect of navigation mode on beam configuration, in which radiation oncology professionals were asked to target treatment beams in real cancer cases.

It should be stressed here that this research was restricted to studying only the beam configurations that came out of the beam-targeting step in the full treatment planning procedure, and did not evaluate complete treatment plans. The question was basically, "Which navigation mode produced better beam configurations, where a better beam configuration could be thought of as one that presented a more promising starting point for the design of the ultimate treatment plan?" All the steps that follow beam targeting in the design of a full treatment plan either did not provide opportunity for improvement through the use of natural navigation (e.g. dose computation and plan evaluation), or were considered to be beyond the scope of this research (e.g. evaluation of the three-dimensional dose field with respect to the anatomy).

Although I anticipated that intuitive navigation would be shown to produce better beam configurations for the reasons stated above, neither I nor my doctoral committee believed that the results were guaranteed. We feared that technological problems with current head-mounted displays would introduce confounding factors that would mask the effect being studied. As will be discussed in the following chapters, statistically significant effects of intuitive navigation were difficult to find. One important result of the quantitative analyses or performance used in this study was that a new intuitive navigation mode, which to my knowledge has not been used before as a means of exploring three-dimensional virtual objects, produced significantly faster performances than the conventional joystick rotation. This is interesting in light of the less significant suggestion arising from this study that the intuitive navigation mode produced worse solutions than the joystick rotation. The task of targeting treatment beams is a design problem that involves numerous constraints and requires many trade-offs to be made. I found that individual differences in approach introduced much unanticipated variability into the data, and this, in addition to the technological deficiencies, made it almost impossible to find any effect. Much interesting qualitative information, however, came from observational data. Future work will be aimed at more thoroughly understanding these observations.

Chapter 2

Radiation Treatment Planning

To understand the effect of the human computer interface on the task of designing a radiation treatment plan, it is necessary to understand the basic principles of radiation therapy, for these principles guide the planning process and factor into the decisions and trade-offs required of every treatment planner. This chapter presents a basic primer of the field of radiation oncology, and is included in this dissertation to enable the reader to better appreciate the complexity of the treatment planning process and to better understand the issues raised by this research.

2.1. Radiation therapy

2.1.1. Introduction

Cancer treatment involves the use of *ionizing radiation*, the absorption of which results in localized release of relatively large amounts of energy great enough to break chemical bonds and initiate the chain of events that ultimately leads to a biologic effect. The basic goal of radiation treatment of a cancerous tumor is to irradiate the tumor with a dosage sufficient to kill the malignant cells without further harming the patient. Under ideal radiation treatment, only tumor cells would receive radiation and healthy tissue would receive no radiation and suffer no damage. In practice, however, this is seldom the case. The use of treatment beams to deliver the radiation makes it impossible to avoid exposing the healthy tissue surrounding the tumor. Even brachytherapy, which involves the implantation of the radiation source within the patient's body, cannot guarantee perfect conformation between the lethal dose and the target volume. Some careful planning is required, then, to ensure the treatment applies uniform lethal dosage to the tumor while delivering tolerable dosage to healthy tissue. The term *radiation treatment planning* refers to the complex process by which a radiation oncologist designs the best treatment for the patient.

2.1.2. Brief history

In 1895 German physicist Wilhelm Conrad Roentgen discovered X-rays, noting that "this new kind of ray" could blacken photographic film sealed in a container and stored in a drawer, and could also pass through materials opaque to light, such as cardboard and wood. Roentgen's distinction as the father of diagnostic radiology and radiation physics probably originated when, during a demonstration of x-ray production, he asked a colleague to put his hand in front of a sheet of photographic film, thereby producing the first radiograph displaying the bony structure of the hand. There is some controversy regarding the first therapeutic use of x-rays, but in 1897 Professor Freund demonstrated to the Vienna Medical Society the disappearance of a hairy mole due to x-ray exposure. (Hall and Cox 1989) The first successful application in America came in 1899 with the treatment of a malignant nose tumor, and that patient was still disease-free in 1920. (Raven 1990)

Parallel to the discovery of x-rays, radioactivity was discovered in 1898 by Becquerel, who inadvertently left a container with 200 mg. radium in his vest pocket for six hours. Becquerel noted that his skin became inflamed within two weeks, and the subsequent ulceration that developed required several weeks to heal. The basic tenets of radiobiology were established by simple radiobiologic experiments conducted in the early 1900's, which showed that radiosensitivity is greatest in rapidly dividing cells and lowest in differentiated tissues, and that oxygenated cells were more responsive to radiation than hypoxic cells.

In the 1920's and 1930's experiments in radiation sterilization of rams demonstrated that while a large single dose of radiation resulted in severe damage to skin and other healthy tissues, dividing the exposure into lower dosage fractions applied over several days achieved sterilization with minimal skin damage. This was the basis of the "Paris" approach to radium treatment of uterine cancer, which held that the use of low-activity sources for longer periods of time improved the therapeutic ratio. This concept was extended by others who also sought the optimal arrangements of the radium sources to achieve a consistent dose rate and a homogeneous dose distribution through the tumor volume while sparing the surrounding tissue. (Hall and Cox 1989)

Still, early radiation oncologists operated much in the dark. Little was known about how much dose was required to destroy a tumor and how such a dose should be administered, and radiobiologic studies of dose-rate effects were decades away. Tumor localization was also a problem, as early oncologists had to rely on physical examination to determine the location and extent of the tumor volume. (Rosenman 1991)

While some treatment centers collected large quantities of radium to use in external beam therapy, the 30's and 40's saw the development of much higher energy x-ray generators, and in the 1950's ^{60}Co units and the first medical linear accelerators became available. The use of atomic weapons in World War II spurred great activity in radiobiology research in an effort to understand radiotherapy, as well as radiation-related mutagenesis and carcinogenesis.

2.2. Principles

2.2.1. External beam therapy

2.2.1.1. Targeting

The treatment beam has to hit the tumor, but it is important to spare as much of the normal surrounding tissue as possible.

2.2.1.2. Shaping

The collimators on most treatment units only provide rectangular field shapes. In most cases treatment planning includes the design and fabrication of custom blocks (typically cast in cerrobend, an alloy of bismuth, lead, tin, cadmium) that carefully shape the beam's cross-section to miss normal tissue that might otherwise have been hit by a plain rectangular beam.

2.2.1.3. Multiple beams

The use of multiple treatment beams that intersect at the tumor is another important strategy in striving for the optimal dose distribution. Whereas each individual beam's radiation is low enough to avoid serious damage to tissue, the intersection of the beams provides a localized volume of high dosage strong enough to affect the tumor within.

2.2.1.4. Beam modifiers

Beam modifiers placed in the beam path and are commonly used to make small adjustments to the dose distribution. *Wedges* are usually made of lead, brass, or tungsten, and when placed near the beam source, they attenuate the beam differentially across the beam's cross section in proportion to the thickness of the wedge. Wedges are often used to smooth out uneven beam distributions produced by non-opposed beam pairs, and can also be used to compensate for skin contours that are not perpendicular to the beam path to prevent hot spots on the skin. *Tissue compensators* can be thought of as generalized wedges, as they are designed to cause non-uniform beam attenuation to account for irregular skin topography, such that the resulting dose is uniform in the underlying tissue. Made of tissue-equivalent material, *bolus* is used when it is necessary to treat the skin surface with a high dose.

2.2.1.5. Beam weighting

Beam weighting refers to the adjustment of the relative strengths of multiple beams to improve the resulting dose distribution. Weighting and wedging are the most common forms of beam modification.

2.2.2. Types of radiation

2.2.2.1. Electromagnetic: X-rays and gamma rays

Sometimes collectively referred to in the literature as *photon therapy*, X-rays and gamma rays are equivalent in nature and behavior, and differ only in source and production. X-rays are generated by accelerating electrons to high energy and stopping them abruptly in a target, usually of tungsten or gold. The kinetic energy of each electron is converted into photons of x-rays. Gamma rays are emitted by decaying radioactive isotopes (e.g. ^{60}Co, ^{226}Ra, ^{137}Cs), often along with alpha rays (positively charged helium nuclei) and beta rays (negatively and positively charged electrons). They are much better suited for brachytherapy because their sources are implantable. Another clinically significant difference between the two types of radiation is that the linear accelerator (x-rays) produces a beam with a much sharper edge than a cobalt unit's beam (gamma rays), which may be important when irradiating close to critical anatomical structures.

The high-energy electromagnetic radiation used therapeutically today is indirectly ionizing, i.e. it does not itself produce chemical and biologic damage, but is absorbed by the medium to then produce fast-moving electrons. These secondary electrons scatter mostly forward in the direction of the original incident photon, yielding a cumulative energy deposition that reaches a maximum at some distance below the skin surface. This is the cause of the skin-sparing effect that allows the effective treatment of deep tumors, and which was not possible with the low-energy x-rays used in early radiation treatment.

2.2.2.2. Particulate

A variety of particle beams have come into use for radiation therapy. Because of their densely ionizing nature, in which ionizing events occur close together in contrast with the widely separated events of the sparsely ionizing x-rays, particulate radiation enables the delivery of sharply defined dose distributions that have very small penumbras. This is very important when treating tumors that are close to sensitive normal structures. A variety of particulate radiation is described below.

Electrons are accelerated to high energy and close to light speed by a betatron or a linear accelerator. They have a limited range in tissues because they are rapidly absorbed and the beam density falls off steeply with depth of penetration. For this reason the depth at which a certain dose is required must be accurately determined so that the appropriate electron beam energy can be selected. *Protons* have 2000 times the mass of electrons and require more complex and expensive equipment to accelerate them to useful energies. They have a specific depth at which the maximum dose can be delivered, requiring an extremely accurate definition of the tumor. For example, 160 MeV. protons have a depth range of about 12 cm. in tissue. *Alpha particles* are positively charged helium nuclei consisting of two protons and two neutrons. They can be accelerated in the same manner as protons, and are also emitted in the decay of some radioactive isotopes. *Neutrons* are produced by colliding charged particles such as deuterons or protons with a suitable target. They are also emitted by the fission of heavy radioactive atoms. Neutron radiation is the only indirectly ionizing particulate radiation. Its biologic effects are produced by the recoil protons, alpha particles and heavier nuclear fragments resulting from its interaction with the medium. *Negative pi-mesons* are negatively charged particles with a mass 273 times larger than the electron. They are produced only in large linear accelerators or synchrocyclotrons capable of accelerating protons to energies of 400 to 800 MeV. They behave like overweight electrons at relativistic speeds, but at slower speeds they are absorbed by atomic nuclei which then explode to produce neutrons, alpha particles and larger nuclear fragments. *Heavy ion* radiation contains the nuclei of elements such as nitrogen, carbon, neon, argon, or silicon stripped of some or all of their electrons. They must be accelerated to energies of thousands of megavolts to be useful, which limits their use to only a few laboratories in the world.

2.2.3. Dose

Because successful radiotherapy depends upon the ability to repeatedly deliver specified amounts of radiation with precision and accuracy and is largely based upon the collected body of accurate clinical data, the quantification of dose very important. Biologic effect is strongly correlated with absorbed dose, which is expressed in terms of amount of energy absorbed per unit mass of tissue. Previously the rad, equal to 100 ergs/gram, was commonly used as the standard unit of dose, but this has been replaced in modern radiotherapy by the gray (Gy), equal to 1 joule/kilogram or 100 rads.

Specifications for a treatment plan usually set forth the *minimum tumor dose*, which is the dose which must be met or exceeded by each cell in the treatment volume. Similarly, irradiation constraints for specific normal tissue is specified in terms of the *maximum dose*, which is the dose that no cell in the particular normal structure should exceed. Minimum tumor dose must be high enough to ensure tumor destruction, and maximum dose for normal tissues should be low enough to minimize adverse side effects of the treatment. In all cases, total dose specifications are practically meaningless without qualifying them with the the fractionation regimen used (see Section 2.2.5 below). (Hall and Cox 1989)

2.2.4. Treatment effects

The use of external treatment beams guarantees the exposure of some normal tissue to radiation, resulting in damage to and impaired function of the exposed anatomical structure. Some such complications are reversible, while others are serious and possibly fatal. Rottenberg points out that neurological complications are often misdiagnosed as tumor recurrence, resulting in the patient being subjected to unnecessary, costly, and painful diagnostic procedures. (Rottenberg 1991) The best approach in dealing with treatment complications is to avoid them in the first place as much as possible. To do this, radiation portals must be carefully designed, radiation doses must be carefully calculated, and potential side effects must be recognized and considered in treatment planning.

Even with the best precautions, complications are impossible to avoid for a number of reasons. Safe radiation thresholds have not been precisely established because the variability among patients of the sensitivity of normal tissues and malignant tumors to ionizing radiation is large. In addition, there exists a lack of clinical data that can be used to relate the tolerance of normal tissues to the fractional volume of the organ irradiated, the nature and function of the irradiated organs, and the stage of cancer treated. Injury thresholds may be lowered by prior surgery, concomitant chemotherapy, or systemic illness. (Rottenberg 1991)

Treatment effects are classified by the length of time from treatment to their onset. *Acute* or *early* effects appear within hours to days after treatment. Tissues most susceptible to acute effects include bone marrow, ovary, and testis. *Late* effects generally do not appear until after treatment has stopped and are most likely to appear in lung, kidney, liver, and heart tissues.

Symptoms of radiation damage vary depending on the location of the damage. Irradiation of the abdomen or pelvis can result in appetite changes, diarrhea, nausea, and sexual dysfunction. Irradiation of the arms and legs can produce skin reactions, decreased function, and fluid retention. Chest- and breast-related effects include breathing difficulty and skin reactions. Head and neck treatment can result in swallowing and chewing difficulty, taste and smell changes, and hair loss. (Dodd 1987)

2.2.5. Fractionation

Hall and Cox refer to fractionation as the most important conceptual development in the history of clinical radiation oncology. (Hall and Cox 1989) Fractionation is the practice of dividing the total radiation dose into fractional parts, and allowing some amount of time to pass between delivery of successive fractions. This allows the repair and repopulation of normal cells, while tumor cells, which do not repair themselves to the same extent as normal tissue does, become reoxygenated, thereby increasing their radiosensitivity. The result of fractionation is tumor control with much less severe acute reactions, but late effects are not affected by the practice. The size and timing of the individual fractions is crucial to successfully minimizing acute effects without permitting surviving malignant cells to proliferate.

2.3. Procedure

2.3.1. Personnel

Radiation therapy is a complex field requiring the collaboration of different people with complementary areas of expertise. A basic radiotherapy team typically consists of a radiation oncologist, a radiation physicist, a dosimetrist and a variety of technologists. Table 2-1 presents the staff members involved in the various steps of radiation therapy. The treatment planning stage, in which treatment beams are targeted and evaluated, primarily involves the physicist, the radiation oncologist and the dosimetrist.

Perez and Purdy describe the characteristics desirable in a *radiation oncologist*, the only M.D. on the team and the person ultimately responsible for the radiation therapy: 1)

Task	Key Staff	Supportive Role
Clinical evaluation	Radiation oncologist	
Therapeutic decision	Radiation oncologist	
Target volume localization		
Tumor volume	Radiation oncologist	Simulat'n technol. / Dosimetrist
Sensitive critical organs	Radiation oncologist	Simulat'n technol. / Dosimetrist
Patient contour	Dosimetrist	Simulat'n technol. / Dosimetrist
Treatment planning		
Beam data-computerization	Physicist	
Computation of beams	Physicist	Dosimetrist
Shield'g blocks, treatm't aids, etc.	Dosimetrist / Mold room technologist	Radiat'n oncol. / Physicist
Analysis of alternate plans	Radiation oncologist / Physicist	Dosimetrist
Selection of treatment plan	Radiation oncologist / Physicist	
Dose calculation	Dosimetrist	Physicist
Simulat'n/verif'n of treatm't plan	Radiation oncol. / Simulat'n technol.	Dosimetrist / Physicist
Treatment		
First day setup	Radiat'n oncol. / Physicist / Therapy technol.	Dosimetrist / Physicist
Localization films	Radiat'n oncol. / Therapy technol.	
Dosimetry chks / Init. chart review	Physicist / Radiation oncologist	Dosimetrist / Chief technologist
Repositioning / Retreatment	Therapy technologist	Dosimetrist / Chief technologist
Periodic evaluation (during treatment)		
Tumor response / tolerance	Radiation oncologist	Nurses / Radiat'n therapy technol.
Follow-up evaluation	Radiation oncologist	Nurses

Table 2-1. Process and staff in clinical radiation therapy. (Perez and Purdy 1992)

sufficient training to interpret treatment planning information and to guide the physicist and/or dosimetrist in achieving the best dose distribution; 2) sufficient knowledge to select best possible combination of dose and fractionation for a given site and volume; 3) competence in the judgment of dose distribution quality, and the technical feasibility and accuracy of a proposed plan; and 4) understanding of the capabilities and limitations of the staff and radiation treatment planning process. (Perez and Purdy 1992)

The *medical physicist* must be familiar with clinical needs, and have extensive knowledge of the capabilities of the equipment. The medical physicist oversees the diagnostic imaging of the patient, and assists the radiation oncologist in treatment planning.

The *dosimetrist* is responsible for performing dose calculations for proposed treatment plans, and works closely with the radiation oncologist and the medical physicist in evaluating and revising proposed treatment plans. The dosimetrist also supervises the fabrication of shielding blocks and any other treatment aids required.

Radiation therapy technologists, simulation technologists and *CT technologists* perform most of the physical work required by radiation treatment and have the most contact with the patient on a day-to-day basis. Technologists will usually work in teams of two to increase efficiency and reduce the chance of error.

2.3.2. Evaluation

The first step in radiation therapy is a complete evaluation of the patient's disease. This requires a thorough knowledge of the natural history and the pathologic characteristics of the tumor. Staging procedures are followed to determine the full extent of tumor. From this information, including the stage and type of the tumor and the routes of spread, a treatment strategy is defined, which addresses whether the treatment will be curative or palliative (relief of suffering and prolonging of life when cure is not deemed possible) and which method or combination of methods will be used.

2.3.3. Planning

One of the oldest and most important principles of medicine can be stated as, "Above all else, do no harm." This is especially important in radiation therapy, where poorly-planned and poorly-delivered treatments can be more detrimental to the patient than no treatment at all (Bentel et al. 1989). In light of this it is of utmost importance in planning radiation treatment to always consider the effects of the proposed treatment on the normal tissues and organs exposed to the radiation. Dosimetry is a vital component in treatment planning, as it provides the basis for evaluating and comparing prospective treatment plans. The product of the treatment planning process is a complete description of the patient's treatment, including a definition of the treatment volume, the intended total tumor dose, the fractionation regimen (number of treatments, dose per treatment, frequency of treatment), and in the case of external beam therapy, specification of treatment machine settings and necessary treatment aids and beam modifiers. The process of treatment planning will be discussed in more detail below in Section 2.4.

2.3.4. Treatment

Figure 2-1 shows a computer model of a typical radiation therapy treatment unit for external beam therapy. A basic treatment machine has six degrees-of-freedom by which the patient and beam source can be placed in the proper relationship as specified by the treatment plan. The table can translate along three orthogonal axes. The gantry, which contains the radiation source in its head, can rotate about a horizontal axis. The collimator, located on the head of the gantry and through which the beam emanates, rotates about an axis perpendicular to the gantry rotation axis. Lastly, the table itself can

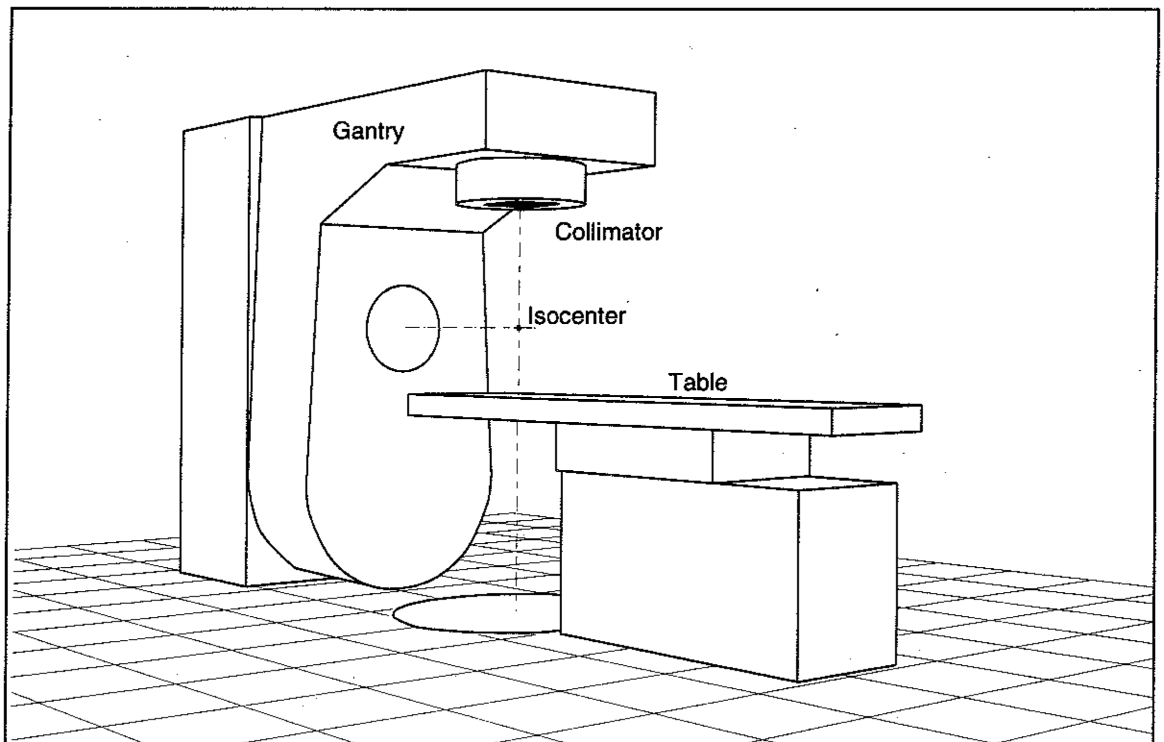


Figure 2-1. Illustration of typical radiation treatment machine. Machine isocenter is defined as the intersection of the rotational axes of the gantry, the collimator and the table.

rotate horizontally about a vertical axis. The three rotation axes intersect at a fixed point that does not change, regardless of the machine settings. In treatment, table translation is used to position the center of the tumor at this point, called the unit's *isocenter*. After that, delivery of the different beams in the treatment plan requires only gantry and table rotation, and perhaps collimator rotation to move from one beam to the next. This method of treatment, called *isocentric treatment*, greatly reduces setup time and chance for error by eliminating table translation between beams.

Fractionation regimens will typically require the patient to receive daily treatments over the course of several weeks. This frequent repetition of the same treatment plan requires accurate and reproducible repositioning and immobilization techniques. Quality control is very important for successful treatment. Regularly throughout the treatment, at the first visit and roughly once a week thereafter, *port films*[1] are taken on the treatment unit for comparison with simulation films and verification that the treatment is being administered correctly.

2.3.5. Follow-up

Patients undergo periodic evaluation during and after therapy to assess the effects of the treatment on the tumor and and to monitor side effects. Not only is this important for the health of the patient, but the clinical data collected contributes to the general body of radiation therapy knowledge that is used to plan therapies for future patients.

2.3.6. Limitations

Failure to eradicate a tumor can result from a number of factors. Clinical factors include the inadequate appraisal of the full extent of the tumor in the surrounding tissues, and clinically unrecognized distant and regional lymph node metastases that go untreated. Physical and technical factors include inaccurate definition of the target volume, substandard treatment planning resulting in sublethal tumor exposure, unreliable patient repositioning and immobilization techniques, and inadequate plan and treatment verification. Biologic factors include the initial cell burden, as large tumors are more difficult to kill than small ones; hypoxic cell subpopulations, which require greater doses of irradiation (this problem is partly resolved by the reoxygenation that occurs between

---

1. Two-dimensional radiographs taken on the treatment machine by briefly exposing photgraphic film on the opposite side of the patient from the beam source to treatment radiation. Port films typically suffer from lower contrast than normal diagnotic X-ray films, because of the higher energy radiation used.

fractions); damage repair between fractions; variation in radiosensitivity at different stages of the cell life cycle; insufficient knowledge of human cell kinetics and biologic equivalents for various dose rate-fractionation regimens; and limited tolerance of normal tissues to irradiation. Other less well-defined factors include the general condition, nutritional status, metabolism, and immune response of the patient. (Perez and Purdy 1992)

2.4. Radiation treatment planning

2.4.1. Objective

The goal of radiation treatment planning is to design the optimal treatment which achieves the greatest possible dose optimization in the irradiated volume. For external beam irradiation Perez and Purdy characterize an optimal dose distribution as having small entrance and exit doses, small side-scattering dose and a narrow penumbra, small differential tissue absorption, homogeneous dose distribution across the treatment volume, and a minimal dose contribution to normal tissue weighted by the tissue's radiosensitivity. The essence of a good dose distribution, however, is that it delivers at least the minimum tumor dose to the cancerous cells and does not exceed the appropriate maximum dose for each type of normal tissue. Treatment planning is most important for patients undergoing high-dosage curative treatment and for selected patients receiving special palliative therapy who require complex treatment techniques or high irradiation doses. (Perez and Purdy 1992)

2.4.2. Simulation and two-dimensional treatment planning

Conventional planning of radiation treatment is based upon the use of a radiation treatment machine simulator. The simulator has the same geometry, dynamics, and controls as the treatment unit, but a standard x-ray imaging tube replaces the high energy treatment source. A combination fluoroscope/film-tray assembly is mounted opposite to the x-ray source. The fluoroscope provides dynamic exploration of the patient's anatomy, and the film-tray is used to produce hardcopy radiographs showing patient position and prospective beam's-eye views.

2.4.2.1. Procedure

As described by Gerbi (1992), two-dimensional treatment planning begins with a simulator session, during which the patient lies on the simulator table in the proposed treatment position. The stability and reproducibility of the patient's position are important factors in the success of the treatment. As such, patient comfort is a major consideration, for the patient must be able to remain still during the 15 minutes or so of the simulation and during each treatment. Polyurethane molds and other devices can be used to improve immobilization.

Once the patient is situated properly, orthogonal anterior and lateral films are taken to record the three-dimensional positioning of the region of interest. Radio-opaque fiducial marks drawn on the skin appear on the radiographs and provide landmarks in addition to bony structures that will help in future patient positioning during treatment.

The vicinity of the tumor is explored with the fluoroscope and prospective treatment beam orientations are investigated. For good prospects, films are taken while the patient and machine are in the proposed treatment position. Each film then is a projection radiograph of the anatomical structures in the proposed treatment field, taken from the point of view of the x-ray source. This *beams-eye view* shows which anatomical structures will be irradiated by the proposed beam. The simulator automatically places in these radiographs a calibrated crosshair showing the locations of the beam's central axis and the field edges. The relationship of these references to fixed bony landmarks is used for subsequent treatment verification. All pertinent beam parameters such as machine settings are also recorded. While the patient is still in the treatment position, a cross-sectional skin contour at the tumor position is recorded, to be used in dosimetric calculations and the specification of beam modifiers. The simulator session ends when all candidate beams have been filmed and recorded.

For each beam, the cross-sectional shape is determined by drawing an outline on the corresponding simulator film. The outline is carefully drawn to include the entire treatment volume and to exclude as much radiosensitive normal tissue as possible. This outline is then used to fabricate a field-shaping block that is mounted on treatment machine's collimator.

The beam directions, field sizes, and the outline of the patient's body are used to compute radiation dose distributions for the proposed treatment. The computed dose distribution serves as the basis for acceptance or rejection of the proposed treatment plan. Qualitative inspection of isodose contours quickly reveals whether the tumor region will receive a therapeutic dose, as well as whether normal tissue will not receive too much radiation. If the dose distribution is unsatisfactory, the plan can be adjusted with beam modifiers. It is possible that some unanticipated condition will prevent any acceptable dose distribution to be designed, in which case another simulator session is required to explore other possibilities.

2.4.2.2. LIMITATIONS

This conventional method of radiation treatment planning is called *two-dimensional* because the design and decision-making process relies upon two-dimensional information. First, the images of the patient's anatomy are two-dimensional radiographs in which one dimension of the three-dimensional arrangement of the anatomical structures is lost. For views that are frequently seen by physicians (anterior, posterior, lateral) it might not be difficult to reconstruct the third dimension from the radiograph. Oblique views that stray from these common views, however, quickly become very difficult to understand. It becomes much more difficult to determine with certainty where the landmarks are and where the tumor is relative to those landmarks. As a result, radiotherapists are reluctant to deviate from the well-understood cardinal angle approaches, even though oblique beams might yield some therapeutic advantage. (Ling et al. 1983) Such conservatism may result in ineffective treatment, for portions of the tumor might receive less than the prescribed dosage and portions of normal tissue might be unnecessarily exposed. (Rosenman 1991)

Therefore, a better three-dimensional understanding of the patient's anatomy should result in more effectively targeted treatment beams. But that is only half the story, for the term *two-dimensional* also refers to the computed dose distribution. Because of limitations in computational power, the dose distribution conventionally is computed only for a two-dimensional cross-sectional slice through the patient. (See Figure 2.2.) This restriction forces treatment plans to use only beams whose central rays lie in the plane for which dose is computed. For such *coplanar* treatment plans, the two-dimensional dose distribution is taken to be a reasonable approximation of the actual dose delivered by the treatment beams to the three-dimensional slab of the patient that contains the target volume.

Although they may be adequately represented by two-dimensional dose distributions, coplanar treatment plans may not be the best for the patient, for a beam that

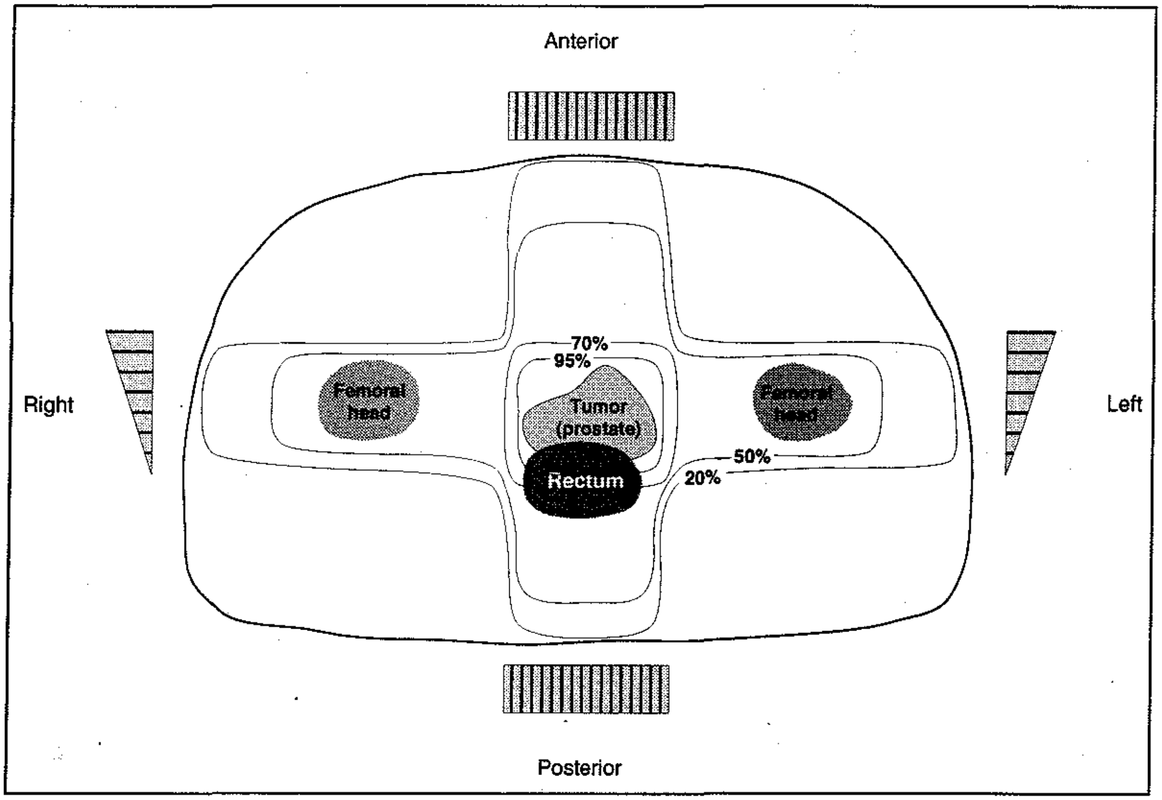


Figure 2-2. Example of a two-dimensional dose distribution computed for a four-field treatment plan. Triangles indicate left and right lateral beams modified by wedges to compensate for sloped skin surface. Anterior and posterior beams are not wedged. Dose contours are labeled as fraction of maximum dose.

strikes the tumor from an inferior or superior direction may do a better job of avoiding radiosensitive healthy tissue. Such *non-coplanar* treatment plans, however, cannot be evaluated with two-dimensional dose calculations, nor with inadequate visualization of a three-dimensional field. The latter is important, for some physicians believe that the lack of suitable methods to display three-dimensional dose distributions has contributed to the general practice of ignoring three-dimensional considerations in radiation treatment planning. (Rosenman et al. 1989)

In addition to the two-dimensional limitations, Goitein listed the following shortcomings of conventional treatment planning: a lack of software aids in defining the clinical problem, no estimate of error in the calculated dose, lack of tools for assessing alternative plans, inadequate geometric definition of anatomic structures, and inadequate tools for specifying and verifying the accuracy in treatment delivery. (Goitein 1982)

2.4.3. 3-D treatment planning

Optimal radiation treatment planning requires thorough comprehension of the three-dimensional arrangement of the patient's anatomy. The two-dimensional information used in conventional treatment planning is insufficient for designing and evaluating the more complicated, non-coplanar treatment plans that promise to more effectively control disease. It is not surprising, then, that the advances over the past 15 years in computed tomography (CT) and other three-dimensional imaging modalities have spurred the development of new treatment planning systems that enable radiation therapists to make design decisions on the basis of three-dimensional information.

By making use of accurate three-dimensional patient data, three-dimensional treatment planning allows: 1) improved targeting of the tumor and conformation of the dose to the target volume; 2) improved delineation of normal tissues, and thus better quantification of normal tissue exposure; and 3) the ability to develop new treatment approaches involving non-coplanar fields that allow the use of higher doses for better tumor control while minimizing normal tissue complications. (Purdy and Emami 1992; Sailer et al. 1992)

The contribution of CT to treatment planning can be significant, although the gain may be minimal for patients for whom the best treatment may be a single beam or a simple opposed beam pair (either antero-posterior—postero-anterior), i.e. plans that could be adequately designed and evaluated on a two-dimensional system. (Perez and Purdy 1992) Sailer et al. (1992) found that treatment planning with their three-dimensional planning system brought about changes when compared with standard treatment techniques in 143 of 144 patients. A review by Tremewan (1988) of 12 studies involving a total of 1293 patients showed that 42% of the treatment plans overall (ranging from 35% for pelvic tumors to 72% for abdominal malignancies) were changed due to CT examination. These changes usually involved enlargement of the treatment volume, although reduction was found to be appropriate in one case. Badcock (1984) used CT in 205 patients and of these, 118 (58%) were helped by CT, 36% had changes in field geometry, 39% had alterations in dose calculations. Estimating how these changes in therapy would affect tumor control, Badcock estimated the resulting improvement in cure rate due to CT to be 4.5%. Similarly, Goitein (1979a, 1979b, 1980) estimated an improvement in survival rate due to CT treatment planning between 3% and 4%. These seemingly small improvements are actually quite significant when one considers that each year approximately 500,000 new cancer diagnoses are treated with radiation. (Perez and Purdy 1992)

2.4.3.1. 3-D TREATMENT PLANNING SYSTEMS

Taking full advantage of recent advances in computational power and display technology, modern radiation treatment planning systems take on a broad range of functions, including: synthesis of all relevant diagnostic information in patient evaluation; appreciation and delineation of relevant anatomical structures (normal tissues, tumor, and target volume); simulation of therapy with generation of digitally reconstructed radiographs (DRR); design of treatment aids (e.g. compensators, blocks); calculation of three-dimensional dose distributions; dose optimization; and treatment plan evaluation. (Perez and Purdy 1992) Chin et al. explored the advantages of non-traditional non-coplanar beam orientations, and suggested that any three-dimensional planning system should comprise three modules: an interactive patient data evaluation module to analyze and integrate various sources of diagnostic data, a module to compute doses (relative and absolute) using three-dimensional patient and beam geometry data, and an interactive dose display module. (Chin et al. 1985) Goitein et al. developed the basic tools found in most three-dimensional treatment planning systems today: anatomy contouring, beam's-eye view, back projection from beam aperture onto CT slice, and projection through the CT sections from any arbitrary point to compute DRR's. (Goitein and Abrams 1983; Goitein, Abrams et al. 1983) Purdy et al. add that the ability to view multi-beam arrangements from any arbitrary position is a necessary complement to the beam's-eye view. (Purdy et al. 1987)

2.4.3.2. PROCEDURE

The procedure for three-dimensional treatment planning is basically the same as that for two-dimensional planning, but an extra step is required to obtain the three-dimensional patient data. At the University of North Carolina the entire treatment planning process, from the initial CT scan to the beginning of treatment, typically requires 6 hours spread over 3 to 4 days. (Sailer et al. 1992)

2.4.3.2.1. Pre-planning and localization

Immobilization devices are fabricated, if necessary, to position the patient in the proposed treatment position. Fiducial marks are painted on the patient's skin using radio-opaque paint. Localization radiographs are taken to document patient position for subsequent plan and treatment verification.

2.4.3.2.2. CT imaging study

The pertinent portion of the patient is imaged by CT, while the patient is in treatment position.

2.4.3.2.3. Delineate planning volumes

The CT data is loaded into the radiation treatment planning (RTP) system and contours are drawn around the tumor, the treatment volume, and the normal anatomical structures on each CT slice. These contours yield a three-dimensional, "stack-of-wire-l oops" representation of the patient anatomy. For well-defined structures, some automatic contouring may be possible to alleviate some of the tedium of this step. Target volume delineation is more difficult than organ or tumor delineation because the target volume has no well-defined shape and depends upon the grade, stage, histology, and routes of spread for the disease. The boundary of the target volume must be defined by a radiation oncologist on the basis of clinical experience.

2.4.3.2.4. Designing beams and field shaping

A three-dimensional RTP system should enable the therapist to take full advantage of the three-dimensional targeting capability provided by the treatment machine. It should also alert the planner to impossible setups. Beams's-eye view is invaluable for quick evaluation of shielding requirements and beam coverage of the tumor and normal anatomy. (Purdy and Emami 1992) has also found the "room view" helpful for appreciating non-coplanar multi-beam plans.

2.4.3.2.5. 3-D dose calculation

Evaluation of a treatment plan depends largely upon the merits of the resulting dose distribution. (Perez and Purdy 1992) Dose calculations are based on dosimetry data obtained with water-filled phantoms, with inhomogeneity corrections used to account for variations in dose due to density variations, such as lungs and bones. Current correction algorithms are crude, and the clinical physicist must be on the lookout for suspect calculations. A new generation of three-dimensional dose calculation methods is currently under development and evaluation. (Purdy et al. 1987)

2.4.3.2.6. Treatment plan evaluation and optimization

The evaluation and optimization of a treatment plan is an iterative process. The initial dose distribution is evaluated qualitatively. Changes are made in the plan to adjust the dose distribution. New dose calculations are performed, and the cycle repeats.

Evaluation of the dose distribution is usually based on a qualitative, visual examination that takes note of tumor coverage, normal anatomy exposure, and severity and location of hot and cold spots in the dose field. The most common way of displaying the dose information is to draw isodose curves on the CT slices, but integration of the individual slices into an appreciation of the three-dimensional shape of the dose field is difficult. Several methods for effectively displaying three-dimensional dose information in conjunction with three-dimensional anatomy are currently being studied. These make use of isodose surface representations, volume rendering, animation and interactivity, and texture. (Rosenman 1991)

In certain cases, qualitative examination is insufficient and quantitative aids are also used by the radiation oncologist. These include the *dose-volume histogram*, which concisely shows how much of a particular structure would receive more or less than a given dose dose level. The dose-volume histogram is useful, but does not provide spatial distribution information, and thus, can only complement, but not replace, the spatial distribution display. Also useful are the biologic models of *normal tissue complication probability* and *tumor control probability*. Goitein and Schultheiss (1985) represented tumor control probability by a clinically determined function of dose. Both Kutcher and Burman (1989) and Lyman (1985) have proposed computing normal tissue complication probability for a given anatomical structure from its dose-volume histogram.

Automated optimization approaches using computer algorithms to determine beam positions, beam weights, and other parameters have not gained widespread acceptance. The design space of radiation treatment planning involves too many parameters and involves optimization criteria that are too vague and too variable from physician to physician to permit automation. More feasible is the use of narrowly-scoped expert systems as consultants to the physician. It is unlikely that technological advances in radiation treatment planning systems will ever totally replace the skills of the radiation oncologist, medical physicist, and dosimetrist. The ultimate responsibility for the selection and execution of radiation therapy techniques, as well as, as for its consequences, will always rest with the radiation oncologist, for the clinical knowledge and intuition gained from years of experience is too difficult to quantitate. (Rosenman et al. 1991; Purdy and Emami 1992)

2.4.3.2.7. Verification of plan and treatment

Before treatment is begun, verification simulation is performed to confirm validity and accuracy of three-dimensional CT-based plan. With the patient again on the simulator in treatment position, proposed treatment fields are filmed and compared with the corresponding DRR's computed from the CT data. If everything compares favorably, the plan is approved for implementation.

Repeated delivery of unusual treatment plans places greater demands on patient immobilization and repositioning, possibly calling for treatment verification through treatment unit port films more frequently than the usual once per week. Oblique beams will be especially difficult to verify, because the poor contrast of port films will exacerbate the difficulty of interpreting an oblique view of the anatomy. Future treatment systems will use on-line imaging systems that can provide convenient, accurate, daily verification of each treatment

2.4.3.3. LIMITATIONS

Three-dimensional radiation treatment planning systems have been demonstrated to be superior to conventional two-dimensional systems in a small number of sites around the world. Commercial systems are even now available. Despite this success, three-dimensional treatment planning tools still contain some opportunities for improvement. Delineation of target volumes and critical structures could be made more efficient and less time-consuming. Dose computation algorithms need to be improved. Plan evaluation and optimization tools could be improved and better integrated into the system. Plan verification techniques need to be improved, as do patient immobilization and setup accuracy. (Perez and Purdy 1992; Purdy and Emami 1992)

Chapter 3

Rationale

3.1. Introduction

To plan optimal radiation treatment the radiation oncologist must fully comprehend the complex three-dimensional situation presented by the tumor within the patient's normal anatomy. The gains achieved by today's state-of-the-art three-dimensional treatment planning systems derive largely from more effective visualization of the patient's anatomy. This improved visualization is important in finding optimal beam directions in the planning stage, and also in correct evaluation of the resulting three-dimensional dose distribution.

The goal of this research was to take things one step further. Could the user interface to the computer-aided treatment planning device be further improved? Specifically, if the means of navigating and exploring the patient's anatomy were more intuitive, would better treatment plans be produced? I hypothesized that intuitive navigation would permit more complete and more efficient searching of the beam targeting solution space, thereby resulting in better treatment plans being produced in less time.

From the dictionary definition, an *intuitive* user interface is one that can be immediately understood without evident rational thought and inference. (Woolf 1975) Ware and Slipp (1991) used the term to describe the user's ability to develop a mental model of the effect of his or her actions in a given context. With an intuitive interface the user is able to anticipate the results of a particular action with minimal increase in cognitive load or short-term memory demands. "Intuitiveness" is not a binary attribute—one should not characterize a user interface as merely being "yes, intuitive" or "no, not intuitive." Instead, intuitiveness should be measured by a continuous scale that reflects how much effort is required for the user to build and maintain the mental model allows him to use the interface effectively. If little effort is required, then the interface can be thought of as more intuitive. If the user is able to use the interface only after prolonged practice, then the interface could be characterized as less intuitive.

Inherent in this definition is variability of the scale from user to user. Different users have different skills and aptitudes, and these differences will produce different assessments of the same user interface. Even so, some user interfaces can be considered more intuitive for an entire population of users, by virtue of their more effective use of basic skills common to all human beings, or common to a defined subset with shared prior experience and training. The head-mounted display is such an interface, using skills common to all human beings.

3.2. Head-mounted display

Unlike a conventional graphics display, whose image moves out of a user's field of view if he or she looks away, the head-mounted display's screens are fixed relative to the user's head such that its images are always in the wearer's field of view. In an *immersive head-mounted display*, the wearer's view of the physical environment is obscured so that all he can see are the computer generated images.

The other key aspect of the head-mounted display is the tracking of the position and orientation of the unit. With this information a new image appropriate to the headset's current situation can be computed and displayed. What is considered appropriate for the new position and orientation of the headset can vary depending upon the application, but what is conventionally done in immersive virtual environment applications is to fix relative to the the headset coordinates the view parameters used to compute the images. In other words, the viewpoint and view frame are always the same relative to the headset. Consequently, the changes in images seen by the user resulting from head movement are identical to the changes in view that result from the same head movement in real life. This familiar response to head movement visually immerses the user in a three-dimensional virtual world. Other possible responses to head movement have also been considered in this research and will also be discussed in later chapters.

Using the head-mounted display, navigating the virtual world can be intuitive and natural, for the user can take full advantage of proprioceptive cues (position information obtained from the muscles, tendons, and joints) and vestibular cues to develop an understanding of the spatial relationships present in the virtual world, and to maintain his bearings within the virtual world. Informal observations in our lab suggest that this effect is particularly true if the angular relationships between the physically real lab space and the virtual environment are held constant. When this relationship changes, as it does when using a treadmill outfitted with a steerable handlebar to explore a virtual building, users will often become lost. A similar situation exists with a stationary-monitor-based interface, which allows the virtual environment to translate and rotate relative to user's physical environment. In such a situation, the user is able to use only visual cues to maintain his or her bearings, which may be difficult in a visually unfamiliar virtual world. The head-mounted display allows the user to make use of skills developed from birth for navigating three-dimensional environments. Desired changes in view no longer need to be decomposed into a sequence of knob turns and mouse movements, but can be naturally effected by the user turning his head and taking a step.

3.3. Head-mounted displays and radiation treatment planning

Will the head-mounted display provide advantages in the targeting of treatment beams over conventional three-dimensional radiation treatment planning interfaces? We know that human beings have perceptual systems finely tuned to interpret visual information in terms of spatial properties, and that spatial visualization is useful in gaining conceptual insights. (Kaiser 1991) Treatment planning, however, requires more than conceptual insights. It requires accuracy and precision in the evaluation of the spatial relationships between anatomical structures, and we also know that human beings sometimes have trouble accurately judging such spatial properties as slant (Perrone and Wenderoth 1991) and direction (Ellis et al. 1991) from static images. Adding motion to the images may help remedy these problems, for human perceptual processes have developed to respond to motion, which is a rich source of information for perceiving a variety of environmental properties. (Proffitt and Kaiser 1991)

This still does not distinguish the head-mounted display from the stationary monitor display. Both can provide us with dynamic, three-dimensional images. The difference is that the head-mounted display enables proprioceptive and kinesthetic information to be used to enhance comprehension of the environment. The advantage of such information is illustrated by Gregory (1991) who recounts the story of a man who was born blind but given vision in mid-life with corneal grafts. This patient learned to judge horizontal distances to objects, such as chairs at the end of a hallway, quite well, but he believed if he hung by his fingertips from a fourth story window his feet would touch the ground. Gregory concludes that the experience of walking was necessary for seeing distance. Haptic information strongly enhances visual information in understanding the environment, and this enhancement is calibrated through the exploratory, trial-and-error activity of childhood.

Because the images it displays depend upon its position and orientation, the head-mounted display provides this haptic enhancement, whereas stationary monitors do not. To effectively use this enhancement in the virtual world, however, it is important to employ the same ground rules as are in effect in the real world. This means that the virtual environment must remain fixed, and any change in a person's relationship to the environment must be due only to the person's movements, which are evident through proprioception. Under these conditions, a person can intuitively explore and appreciate the virtual environment. For radiation treatment planning, this could translate into more accurate beam targeting arising from a better understanding of the three-dimensional anatomical structure. Additionally, searching the treatment beam solution space could become more efficient, because it will now be easier to maintain one's bearings and know what beam directions have already been considered and where the good prospects are, and what possibilities have yet to be considered.

3.4. Norman's design principles

To better understand how the intuitive interface provided by the head-mounted display might benefit radiation treatment planning, let us consider the principles proposed by Norman for the design of human interfaces for tools and machines. (Norman 1988)

Norman characterizes human action as having three components: goals, execution and evaluation. Execution involves doing something, and evaluation is the comparison of what happened in the world with what we wanted to happen. Execution and evaluation can be further decomposed, resulting in the seven stages of the Action Cycle: 1) First, the goal is formed. The goal is the state that is to be achieved. 2) Next, the intention, or the statement of what needs to be done to achieve the goal is formulated. 3) Then an action sequence of specific commands to execute is specified to satisfy the intention. 4) The action sequence is carried out. Steps 2-4 make up the execution aspect of human action. The following steps compose the evaluation aspect. 5) The state of the world is perceived. 6) The state of the world is interpreted with respect to the goals and expectations. 7) Finally, the state of the world is evaluated, and what actually happened is compared to what was wanted.

As an example, consider the targeting of radiation treatment beams. In the process of designing a treatment plan one might want to examine the efficacy of a prospective beam directly opposed to a particular beam. This represents a goal, for which an intention is then formulated to change the beam's-eye view provided by the system to represent the prospective beam. To satisfy this intention, an action sequence of specific commands is built and executed. The action sequence depends on the capabilities provided by the treatment planning system. One may have to type in new gantry, collimator, and table parameters, or one might be able to just click on a "View opposed beam" button. After the execution is completed, evaluation begins with perception and interpretation of the state of the world. What does the beam's-eye view look like now? What beam is represented by this beam's-eye view? Finally, the new beam and beam's-eye view is compared with the desired beam and view. If they do not match, then the action cycle must be repeated.

The success of the action cycle depends entirely upon the correspondence between mental intentions and interpretations and physical actions and states, and Norman describes two "gulfs" that separate mental states from physical ones. The *Gulf of Execution* is the distance between the user's intentions and the actions allowed by the system. Does the system provide actions that correspond to the user's intentions? In the example above, a system with a "View opposed beam" button would have a small Gulf of Execution for that particular goal, but a system that could only respond to gantry, collimator and table parameters would have a large Gulf that could be bridged only with great difficulty. The *Gulf of Evaluation* represents the amount of effort required to interpret the state of the system and to determine how well expectations and intentions have been met. A beam's-eye view does much to bridge the Gulf of Evaluation, since physicians can usually orient themselves quite readily from a perspective view of human anatomy. Some enhancements, though, can still be helpful in shrinking the Gulf. Shaded surface representations of anatomy may be more readily comprehended than wire-loop representations, especially when the viewing direction is parallel to the planes of the wire loops. And some representation of the original beam would be helpful in determining how well the action sequence has succeeded in changing the beam's-eye view to direct opposition to the original beam.

Norman's strategy for improving user interfaces relies on a handful of design principles, which are briefly described here.

Make things visible: bridge the gulfs of Execution and Evaluation. Users need to know what is possible and how actions should be performed. These possible actions provided by the system should match the user's intentions. Users also need to be able to readily understand the effects of their actions. Timely feedback must be provided to the user.

Use both knowledge in the world and knowledge in the head. Users learn more quickly and feel more comfortable when the knowledge required for a particular task is available externally, but they perform more efficiently after they are able to internalize the knowledge. The system should support both resources.

Simplify the structure of tasks. System designers must pay attention to the limits of the user's cognitive resources, such as short- and long-term memory and attention. Tasks should require as little planning and problem solving as possible. Interruptions should be minimized, and easy recovery from interruptions should be possible.

Get the mappings right—exploit natural mappings. Natural mappings help the user to understand the relationships between intentions and possible actions, between actions and their effects, between the actual system state and what is perceivable, and between the perceived system state and the intentions and expectations of the user.

Exploit the power of constraints, both natural and artificial. Constraints help direct the user toward valid actions and away from possible errors.

Design for error. The designer must assume that any error that can be made will be made, and must plan for it. The user should be able to recover from errors and reverse any unwanted outcome, and to learn from the error by understanding what was done and what resulted.

When something can't be designed without arbitrary mappings and difficulties, standardize. When related actions work in the same manner, their operation need only be learned once.

For exploration of a virtual world, intuitive navigation with a head-mounted display satisfies most of these principles, because it makes use of natural skills possessed by almost all human beings. Users know what actions are possible because they are the same familiar actions with which the user explores the real world. Their experience in the real world has taught them how to use these actions, so the required knowledge has been internalized and the mappings between intentions, actions and effects are readily understood. Because an intuitive interface ideally requires no thought or attention, the user is able to concentrate on the real task at hand. In radiation treatment planning, this means that the therapist can concentrate on evaluating prospective treatment beam directions without the disruption of having to figure out how to move from one prospective direction to another. The interface with the computer requires very little attention, and ideally becomes invisible to the user. Exploration of the solution space is simplified, and backtracking from design dead-ends becomes easier. Navigation errors are easily reversed or corrected, although this may become more difficult when system lag becomes excessively large or update rates drop below the threshold for perceiving smooth motion. These problems are discussed in more detail below.

Treatment beam targeting also presents a good opportunity for the intelligent application of constraints. Zeltzer (1992) upholds Norman's principle and recommends that input operations be properly organized so as to reduce functionally the number of degrees of freedom that must be directly controlled. Head-mounted displays, which report position and orientation, provide six degrees-of-freedom that can be used for navigation control. This is more than what is necessary for targeting treatment beams, because treatment beams themselves are constrained to hit the target and to originate a fixed distance away from the target. Under these conditions, beam targeting becomes essentially a two-dimensional problem whose solution space is the surface of a sphere centered at the target. Of the four intuitive navigation modes studied in this research, one used seven degrees of freedom (six head-movement and one model translation), one used 10 degrees of freedom (six head-movement, one model translation, and three model rotation), and two used only three degrees of freedom (three head rotation). As might be expected from the above design principles, the results reported in the next chapter demonstrate that "less is more" for this task.

3.5. Problems

The research reported here attempts to determine if an intuitive, head-mounted display-based user interface enables us to better perform the task of targeting radiation treatment beams. Such an interface shows great promise, for it employs our innate spatially-oriented skills and perceptual systems and enables us to devote more resources to solving the problem at hand. We must recognize, however, that there are severe problems and limitations with current head-mounted display systems that counteract these benefits and that may mask any performance improvement brought about by the interface. It is a very real possibility with this kind of study that inconclusive results will be obtained—that no improvement in performance will be seen with intuitive navigation. In that situation it will be impossible to determine without further research whether intuitive navigation truly had no effect or system problems masked the effect of intuitive navigation.

Below are discussed the known difficulties in using intuitive navigation with a head-mounted display. They are divided into two categories. Technological problems are problems with hardware and software that diminish the usability of the head-mounted display system. If the goal of the intuitive interface is to make the best use of natural skills developed through interaction with the real environment, then the system must present sensory stimuli that is consistent with the expectations of the perceptual system. Technological problems include system characteristics that interfere with meeting these expectations, but also include basic ergonomic concerns regarding the apparatus itself. Perceptual problems are complications that arise within the user's perceptual and cognitive processes, and can largely be thought of as the consequences of the presence of technological problems. Perceptual problems are not nearly as well understood as technological problems. The limited field-of-view of a head-mounted display may be easily measured, but its effect on the wearer's appreciation of a virtual architectural design has yet to be thoroughly studied. Some of the problems discussed below are not specific to the head-mounted display, i.e. they are also present in stationary monitor displays, but they are included here because they are sources of error in the targeting of treatment beams.

3.5.1. Technological problems

3.5.1.1. INADEQUATE PATIENT MODEL

Three-dimensional radiation treatment planning relies heavily on the three-dimensional model of the patient's anatomy built from a CT study. Anatomical structures are modeled from their apparent locations in the CT dataset. Possible errors in position and intensity values of individual voxels, as well as errors in registration of multiple slices of the dataset will produce inaccuracies in the patient model. Treatment plans designed on the inaccurate model may be optimal for the patient model, but may also be sub-optimal for the real patient.

3.5.1.2. INADEQUATE IMAGE

Even if a perfectly accurate model of the patient could be guaranteed, the images generated to display the model will introduce errors. (Holloway 1993) Practically all treatment planning systems use a raster display, which means that the image of the anatomical model must be rendered into a discrete frame buffer. The result is an image that suffers from either aliasing or blurring, depending upon whether anti-aliasing techniques have been used. Significant errors can also be produced from inaccurate view parameters in the generation of the image. Using incorrect values for field-of-view or inter-pupillary distance (for stereo images) will yield images that may not be correctly interpretable in terms of spatial relationships. Also, off-center projections may be required by the geometry of the display and the viewer.

3.5.1.3. POOR DISPLAY QUALITY

Current commercially-available, off-the-shelf head-mounted display units typically use liquid crystal displays, which are inherently deficient in resolution and contrast ratio. When magnified through wide-angle optics these displays, typically 120x240 pixels, yield visual acuity that is so poor as to render the user legally blind, i.e. worse than 20/200. The low contrast ratio further hinders appreciation of fine detail in the image. Poor display quality can be especially damaging in targeting precise beams that have thin safety margins. In such cases fine details of the anatomical structures must be visible to avoid unintentional irradiation of radiosensitive tissue.

3.5.1.4. Optics errors and distortion

Because the liquid crystal screens in head-mounted displays are too close to the wearer's eyes to focus on directly, optical systems are interposed to move the image to a comfortable focal length. These optics are also used to magnify the image so that it occupies a larger portion of the wearer's field of view. Such optical systems will produce pin-cushion distortion of the image, causing straight lines to appear as concave toward the periphery. With normal human development, the human perceptual system has not been calibrated to this distortion, and the result is altered spatial perception and possible motion sickness resulting from mis-coordinated visual and vestibular percepts. Robinett and Rolland (Robinett and Rolland 1992) describe mathematically the geometric errors introduced by LEEP optics (Howlett 1983), which are popular in commercial head-mounted displays, and prescribe a corrective predistortion that can be applied to the image to neutralize the optics distortion.

3.5.1.5. Construction and calibration

Assuming that the stereoscopic views of the patient's anatomy are computed with the correct parameters, the headset itself must be properly constructed and calibrated to the individual user to correctly implement the desired percept. (Holloway 1993) Misadjustments in the head-mounted display can result in disagreement between the physical values for field-of-view, interpupillary distance, and center of projection, and the values used in image computation. Such mismatches will affect spatial perception.

3.5.1.6. Tracker error

The errors in the information reported by the tracking device produce errors in the resulting images, because the images are computed from incorrect view information. (Holloway 1993)

3.5.1.7. Image lag

No matter how fast and powerful the system components are, there will always be some finite amount of time between the user's movement and the corresponding change in the image. Mine attributes this delay to four sources: *tracking system delay*—the time required for measurement and calculation of the user's movement and transmission of that information to the host by the tracker; *host-processing delay*—the time required for the host to collect all pertinent sensor data (trackers and input devices) and manipulate the display list accordingly; *rendering delay*—the time involved in traversing the display list and rendering the image into frame buffers for the two eyes; and *display delay*—the time required to scanout the contents of the frame buffers into the head-mounted display's liquid crystal screens. (Mine 1993)

Stationary monitor systems also suffer from image lag, which is composed of the same four components. In these systems tracking system delay can be defined more generally as the time required for the input device being used to manipulate the model to measure and report the user's action (e.g. joystick deflection, button press) to the host. In general, however, the image lag for stationary monitor systems is much less noticeable than that for head-mounted display systems. This difference may be due to a number of causes. First, the image lag for stationary monitor systems may very well be less than that for head-mounted display systems. The tracker delays for these input devices is usually very small compared to the delay for a head-mounted display's spatial tracker, because they typically are much simpler devices that report easily measured quantities to the host through an analog-to-digital interface. It is possible, of course, for the input device to incorporate a complex device, such as a spatial tracker, in which case the required computation will incur an appropriate delay penalty.

Even if the lags were quantitatively equivalent, however, they may still be more noticeable in head-mounted display systems than in stationary monitor systems. Human sensorimotor systems are finely-tuned to detect discrepancies between the expected consequences of voluntary movement and the actual sensory feedback. Image lag creates such a discrepancy, but may be more noticeable with head-mounted display systems if the user's head-neck control is more sensitive than the arm-hand-finger control typically employed by stationary monitor systems.

Also, input devices used with stationary monitor systems typically have a limited range of movement and are therefore most often used as velocity controls for rotation or translation. On the other hand, head-mounted display trackers are used for positional control of translation and orientation. If the human perceptual system is more sensitive to changes in position than to changes in velocity, then image lags associated with velocity control devices would be less discernible than those associated with position control devices.

For a simple scene consisting of only one triangle, Mine (1993) estimates the total image lag introduced by UNC's head-mounted display system (Polhemus Fastrak tracker, Sun-4 applications host, Pixel-Planes 5 image generator, and Virtual Research Flight Helmet head-mounted display) to be approximately 80 ms. For a typical application, which uses a manual input device and more complicated scene, lags of 250 ms. are not unusual. The effect is debilitating, for the perceived behavior of the virtual environment is so unrealistic as to thwart any illusion of immersion or sense of presence in the virtual world. At the start of a movement such as a head turn to the left the virtual world is perceived to rotate with the user's head. Then, when the images corresponding to the beginning of the head turn are finally displayed, the virtual world is seen to rotate to the right as it should. When the user stops the head, the virtual world still appears to rotate to the right for a short time before becoming still. The effect can be defined as the virtual world not being stationary, but instead "swimming" about the user. The response of the user is to usually to minimize the swimming by slowing down his movements. And not only is performance degraded by this slower, less precise movement, but image lag is also a major contributor to the onset of motion sickness.

3.5.1.8. Update Rate

In the real world, objects move smoothly in the visual field, and not in discrete steps. The dynamic images used in head-mounted displays can only be generated in discrete time steps, but decreasing the time step (increasing the update rate) can help to induce the perception of smooth motion. Asa rule of thumb, developers of head-mounted display applications aim for a minimum update rate of 15 updates/sec. (Holloway and Lastra 1993; Pausch, Chung et al. 1993) The cost of having to generate each scene update more quickly, however, is that scenes must usually be made simpler for the system to satisfy the time constraint, and this is often a difficult trade-off. Consequently, head-mounted systems often suffer from jerky motion, and there is evidence that such low update rates negatively impact spatial perception and task performance. (Tharp et al. 1993)

3.5.1.9. Field of View

The image seen by a single eye in a head-mounted display occupies a certain portion of that eye's total field of view. The angles subtended horizontally and vertically by the image are called the horizontal and vertical *physical* fields-of-view. It is important that the *computational* fields-of-view used in calculating the displayed perspective image match exactly the physical counterparts to ensure veridical presentation of the virtual environment to the user.

When using two eyes in a head-mounted display, the *binocular* and *overlapped* fields-of-view also become important. The overlapped field-of-view is the intersection of the individual fields-of-view of the two eyes, and represents the portion of the visual field in which stereoscopic vision is possible. The binocular field-of-view is the union of the two individual fields-of-view, and is a measure of the size of the total visual field allowed by the device. Using their computational model of the optics in a head-mounted display Robinett and Rolland calculated the horizontal fields-of-view for the VPL Eyephone formerly used at UNC to be 75.3° for a single eye, 60.6° for overlapped, and 90.0° for binocular. (Robinett and Rolland 1992)

Robinett and Rolland believe that a larger binocular field-of-view will strengthen one's feeling of actual presence in a virtual world. This makes sense but the effect has yet to be studied. Similarly, little is known about the effect of field-of-view—monocular, binocular, or overlapped—on task performance, or about the changes in spatial orientation caused by mismatched computational and physical fields-of-view. (Robinett and Rolland 1992)

3.5.1.10. Accommodation

All head-mounted displays in use today provide only a single focal distance for all virtual objects in the scene, regardless of their actual distance from the user. Therefore, even though the user's eyes may change their vergence angle when shifting attention from a far object to a near object, the eyes' accommodation does not change. Normally, changes in vergence are closely linked to changes in accommodation. Head-mounted displays force the user to counteract this reflex, and Rushton and Wann have found incorrect accommodation to be a major factor in producing negative side-effects from head-mounted display use. (Rushton and Wann 1993)

3.5.1.11. Ergonomics

Although the engineering of head-mounted displays is continually improving, current systems still are uncomfortable to the average user. They are heavy and their stabilization usually requires a headband tightened to the brink of painfulness. If the headband is not sufficient, then one or both hands must be used to stabilize the unit. Ventilation inside the head-mounted display is usually not very good, causing users to get hot and causing their eyeglasses to fog up. These effects are constant reminders to the user that he is standing in a lab with a funny device on his head, and not truly in the virtual world he sees. They run counter to the idea of a natural, intuitive interface, and they fatigue the user.

3.5.2. Perceptual problems

3.5.2.1. MOTION SICKNESS

Motion sickness and *simulator sickness* are two terms often used to describe the undesirable side effects of using head-mounted displays. The use of the terms is confusing, as they seem to mean different things to different people. Kennedy et al. (1992) distinguish the two by using *motion sickness* to refer to the symptoms experienced when subject to abrupt, periodic, or unnatural accelerations, and *simulator sickness* to refer to symptoms caused by incorrect aspects of simulation. The nausea brought on by the faithful simulation of a turbulent airplane flight is an example of motion sickness. On the other hand, visually-induced motion sickness (VIMS), in which the user becomes sick through vection (the illusion of self-motion) without any vestibular stimulation, is an example of simulator sickness. Although the collection of signs and symptoms for simulator sickness overlaps greatly with that for motion sickness (pallor, sweating, salivation, nausea, drowsiness, general discomfort, apathy, stomach awareness, disorientation, fatigue, incapacitation), the profile of symptoms for typical simulator sickness is different from that of motion sickness. For example, vomiting and retching are rare in simulator sickness, and headache, eyestrain and blurred vision are found mostly in simulator sickness. Simulator sickness also produces residual aftereffects: illusory sensations of climbing and turning, perceived visual field inversions, and disrupted motor control.

Using these definitions, the symptoms brought on by using a head-mounted display, in which vestibular stimulation is minimal, can be classified as simulator sickness. Yet, other authors call it motion sickness, invoking a broader definition of the term, that of symptoms produced by ambiguities among visual, vestibular, proprioceptive, and auditory senses to which the user has not yet adapted (McCauley and Sharkey 1992). DiZio and Lackner use the term *sensorimotor rearrangement*, which can be said to exist when changes occur to the user or to the environment such that sensorimotor information interacts in unfamiliar ways (e.g. VIMS, in which visual information conflicts with vestibular cues and posture control (Hettinger and Riccio 1992)), when sensorimotor systems are disrupted or forced to work outside normal parameters (e.g. wearing a 4 lb. head-mounted display, which significantly changes the head's effective weight and moment of inertia), or when there is a violation of the constraints to which adaptation has currently calibrated the spatial orientation systems (e.g. space vehicles that rotate to create "artificial" gravity). (DiZio and Lackner 1992) This definition of motion sickness, which we shall use for the following discussion, does not require direct inertial stimulation of the vestibular apparatus, but does require an intact, fully-functioning vestibular system. (Ebenholtz 1992)

The risk of motion sickness in head-mounted display applications is greatest in those that McCauley and Sharkey (1992) call *far* applications. Far applications involve distant objects—objects too far apart to permit walking from one to another within the bounds of the typical tracking device. The solution is for the user to "fly" from one object to another, to invoke with a gesture or a command self-motion from one location in the virtual world to another. In the absence of corroborating vestibular cues, the strong vection produced by the visual cues can result in VIMS. (Hettinger and Riccio 1992; McCauley and Sharkey 1992)

The targeting of radiation treatment beams, on the other hand, is more of a *near* application, characterized by proximate objects, a stationary self, and the absence of vection. (McCauley and Sharkey 1992) In near applications, vestibular function is limited primarily to head movements and whole-body rotations and linear accelerations are not encountered. Motion sickness is not expected to occur unless frequent head movements are required and some aspect of the virtual environment has stressed the vestibular-ocular reflex. (Ebenholtz 1992) warns that any condition yielding error in eye movement control, along with the ensuing feedback and error-correcting signal, is a potential source of motion sickness. This may be the root of the second general situation in which VIMS occurs according to (Hettinger and Riççio 1992) (the first being when strong vection is evoked)—when perceivable image lags exist between head movements and the corresponding image changes. The significant image lag in the current UNC Pixel-Planes 5/head-mounted display system can be expected to produce motion sickness.

The danger of motion sickness for this study was that, even though users may be able to adapt to the sensorimotor rearrangements produced by the head-mounted display, adaptation takes time and negatively affects performance. When there is risk of motion sickness, the common wisdom dictates that initial exposures be limited to short durations to facilitate adaptation with only minimal motion sickness symptoms, and to allow adequate time for readaptation to the real world before engaging in potentially dangerous activities. (Hettinger and Riccio 1992; Ebenholtz 1992; McCauley and Sharkey 1992) Our subjects were closely monitored for signs and symptoms of motion sickness.

3.5.2.2. SYSTEMATIC VIEWER BIASES

Perspective displays, even when viewed from correct viewpoint, are subject to systematic viewer biases. Ellis found sinusoidal variation in errors in azimuth estimation from perspective views (Ellis and Grunwald 1989; Ellis et al. 1991), and Perrone and Wenderoth were able to mathematically model the underestimation of the perceived slant of rectangles (Perrone and Wenderoth 1991). Although these studies were performed with monoscopic, static images that were viewed binocularly, their results might generalize to some extent to the binocular, stereoscopic, dynamic images seen in head-mounted displays. If everything about the image generation and display were perfect, there might still be error in spatial judgement. Ellis found that the appropriate geometric distortion of the perspective image can counteract these biases and improve user performance. He suggests that head-mounted displays may require such intentional distortions for accurate spatial judgements, but also acknowledges that these might cause a loss of visual-vestibular coordination and induce motion sickness. The successful design of head-mounted spatial instruments will therefore require an understanding of the tolerable limits of conflict between visual and vestibular information.

3.6. Related work

To date there has been very little hard science reported evaluating human performance in virtual environments and with intuitive interfaces. Using Brooks's classification of human-computer interaction research results (Brooks 1988), *rules-of-thumb* and *obser-*

vations abound, but *findings* are scarce. This is unfortunate in light of Robert Eggleston's assertion that human performance assessment is critical in the development of effective virtual environment interfaces and task environments (Pausch, Chung et al. 1993). The added value of virtual environment systems must be objectively determined, which means that researchers must move beyond the current emphasis on concept demonstration and on to systematic, controlled human performance studies.

Eggleston offers three possible perspectives for such work. The first is a *task requirements satisfaction view,* which focuses on whether or not virtual environment systems offer any advantage for a specific task or class of task, such as maintenance training. Using intuitive navigation for targeting radiation treatment beams is an example of this view. The second is a *VE-unique property view,* which focuses on the performance gain of some new capability made possible by virtual environment technology, such as the spatial localization of sound. The third perspective is a *technology expansion view,* which seeks to guide development of virtual environment systems by establishing design requirements related to human sensation, perception, cognition, and motor behavior.

Eggleston's own research at Wright-Patterson's Armstrong Laboratory is multiviewed. A recently completed study compared a virtual manual control device with a physical control device. The results have not yet been published, but Eggleston has indicated that the study did point out some of the problems of running virtual environment studies, including the difficulty in selecting conditions to ensure fair comparisons. Currently, Eggleston is in the process of studying the effects of spatial-temporal registration errors in multi-sensory virtual environment systems on general human performance. What follows is a brief review of other research concerned with using head-mounted displays or intuitive navigation in a task environments.

3.6.1. Basic human performance parameters

Tom Piantanida and colleagues at the Stanford Research Institute take the technology expansion view in their research (Pausch, Chung et al. 1993). Piantanida's group has been studying field-of-view effects by varying the size of an imaged aperture through which the subject must look to find a red cube target. The subject wears a head-mounted display and his head is tracked. The aperture is fixed in head-space, so varying its size effectively varies the subject's binocular field of view. Measuring time to target detection against aperture size, initial results verify the expected effect of target detection being impaired by reduced field of view. Future experiments will vary environment and target and introduce distractors. Such studies will hopefully reveal the effects of virtual world contents and the constraints on detection of critical features.

SRI researchers are investigating the effects of system-produced distortions by decoupling visual feedback from kinesthetic and proprioceptive feedback. By systematically altering the visual feedback provided to the user they hope to examine the effects of contradictions between proprioceptive and visual information. The expectations are that the viewer will be able to adapt to the sensory discord, but the time to adaptation will differ with task and with experimental condition.

Using two high resolution color monitors (1280 by 1024 pixels) that can be programmed to simulate coarser resolutions, Piantanida intends to study the effects of display resolution on visually-guided tasks. Also proposed are studies of the effects of image lag and update rate on task performance.

3.6.2. Hand-tracking vs. head-tracking

At the University of Virginia, Randy Pausch and colleagues recently completed a small study comparing hand-tracked navigation with head-tracked navigation in a simple search task (Pausch, Shackelford and Proffitt. 1993). Twenty-eight subjects were asked to search a virtual environment for twenty target numbers and to call out each number as it was found. For each trial the subject used either a head-tracked navigation mode, in which the subject viewed the virtual room through a head-mounted display with full six degree-of-freedom tracking, or a hand-tracked mode, in which the subject viewed the virtual room through a stationary head-mounted display but was able to change the view direction by manipulating a flashlight outfitted with a six degree-of-freedom tracker. As each target number was found the subject called it out and and it was removed from the environment.

Pausch et al. report two major findings. The first is that searching for targets with head-tracking was 42% faster than searching with hand-tracking (1.5 sec. vs. 2.6 sec.), and although no significance statistics are reported for the difference, the presented distribution of data indicates very high significance. The second result is that for hand-tracked performance, subjects using hand-tracking after previously completing head-tracked trials performed 23% faster than subjects who used were using the hand-tracked navigation first. Again, however, no indication of significance is given, nor is any summary of data variability presented.

The results of this study are germane to investigating radiation beam targeting with intuitive navigation modes. Unfortunately, in the preliminary reports seen so far, the investigators have done very little to explain the cause of such performance effects. They speculate that some sort of training occurred to enable the hand-tracked performance of those subjects who had used the head-tracking first to surpass that of the subjects who used hand-tracking first. Pausch et al. also speculate that head-tracked navigation produced better mental models of the virtual environment, enabling subjects to more easily determine where they had and had not already searched. One might also guess that fatigue and physical constraints associated with using the tracked flashlight would seriously affect the performance in the hand-tracked modes. One last observation from the study is that whereas a "practice effect," in which performance improved dramatically over the first couple of trials and then leveled off, appeared in the hand-tracked trials, the head-tracked trials by and large showed no such effect. This could be taken as support for the argument that intuitive navigation modes require very little learning, but to be fair, one might instead infer that the head-tracked navigation was so difficult that no practice effect was possible. The former is more likely the case, however, given that twenty-five of the twenty-eight subjects preferred the head-tracked navigation over the hand-tracked.

3.6.3. Navigation metaphors and velocity control

Ware and Osborne (1990) defined and evaluated three navigation metaphors, based on a six degree-of-freedom mouse, that can be used when viewing a virtual environment through a stationary monitor. The *scene-in-hand* metaphor allows the user to change the view by translating and rotating the scene. This metaphor is best suited to applications that involve manipulation and examination of compact objects, but does not work well for navigating a virtual environment that surrounds the user. The *eyeball-in-hand* metaphor gives the user direct control of the position and orientation of the viewpoint. Ware and Osborne claim that this metaphor is very easy to learn, but it requires the user to maintain a mapping between the orientation of the virtual environment within which the input device operates and that of the virtual environment seen through the fixed monitor. The third metaphor is that of the *flying vehicle* whose translational and rotational velocities are controlled with the manual input device. This mode was found to perform best in navigation within surrounding environments, but did not lend itself well to external examination of discrete objects. Ware and Osborne found that for through-the-window viewing the nature of the task determines which metaphor is most appropriate to use, and it is interesting to note that the range of tasks found in a typical head-mounted display application can require all three metaphors to be in use simultaneously. In an immersive virtual environment it is possible for a user to fly to a distant destination while holding a virtual object in his hand and controlling his view with head movements.

Ware and Slipp (1991) compared three methods of controlling the velocity of a flying vehicle: a six degree-of-freedom mouse, with which translational and rotational velocities were determined by the displacement of the mouse from a "null" point, defined by a button press; a six degree-of-freedom isometric joystick; and a conventional two-dimensional mouse and monitor panel, in which cursor position on the control panel display during a button press determined flying direction, and a control panel slider controlled speed. The control panel was found to be easy to use and safe in the sense that users seldom lost control. On the other hand, the control panel allowed only one dimension of motion at a time, and required too much hand movement to operate. Perhaps because it allowed simultaneous control of translation and rotation, the six degree-of-freedom mouse was described as the most fluid and natural control. Arm fatigue and and inability to return to the null point were its major drawbacks. The isometric joystick was appreciated for its ability to return to the null point and for being easy to learn, but subjects found it difficult to exert fine control and to separate translational control (forces exerted on the joystick) from rotational control (torques). As a result, navigation with the joystick often degraded to a series of one-dimensional movements, despite the fact that six degrees of freedom were theoretically available.

One other useful finding is that users tend to adjust their flying speed in proportion to the scale of the environment through which they fly. One of the flying tasks used by Ware and Slipp was to fly through a large cornucopia as quickly as possible but without hitting the walls. The cornucopia was actually a square tube built from a sequence of successively smaller segments that randomly meandered through space. The size of each segment was 75% that of the previous segment, so over the thirty-two segments that made up the object, the scale decreased by four orders of magnitude. Ware and Slipp found that their subjects maintained a constant ratio of speed to local tube size, and the value of this ratio varied among navigation modes. The six degree-of-freedom mouse and the control panel yielded similar ratios, which were greater than the ratio produced by the isometric joystick.

3.6.4. Architectural visualization

Although there is no shortage of promising ideas for improving human performance through the adoption of head-mounted displays and intuitive interfaces, very few applications have been developed to the point of being able to provide some evaluation of the new methods. Close to the completion of its first decade, the Walkthrough project at the University of North Carolina is one of the most mature applications of intuitive navigation (Brooks 1986). As such, it has generated valuable observations regarding virtual environment navigation. For beam targeting the most relevant observation may be that users can better maintain their bearings when the orientation of the virtual building is fixed relative to the user's physical environment. Under these conditions it appears to be easier for the user to build a mental model of the virtual environment and to keep track of his situation within that environment. Interfaces that violate this principle, such as the steerable treadmill that was used for some time at UNC, greatly complicate navigation.

Related research has been conducted by Daniel Henry at the University of Washington (Henry 1993). Henry studied the appreciation of architectural space gained through four different methods. The first method was to physically walk through the real space, an art gallery. The second and third methods involved wearing a head-mounted display and exploring a virtual model of the gallery. In one method head-tracking was used to provide the user with natural view control. In the other, no head-tracking was performed. The fourth method was to view the virtual gallery on a stationary monitor. Henry found that the three virtual display conditions yielded dimension estimates significantly smaller and more inaccurate than the "real" display, and he attributed that the lack of peripheral vision when using the displays. He also found that the fully-tracked, immersive head-mounted display method produced significantly worse dimension estimates than either of the other two virtual displays. Henry suggests that this effect may be a result of a tendency for subjects, when immersed in the virtual world with a head-mounted display, to look through the left and right periphery of the head-mounted display, where distortion is the greatest, when turning their heads to study the space.

Chapter 4

Preliminary Experiment

4.1. Introduction

The key to successful beam targeting in radiation therapy treatment planning is to orient and shape the beams so that the entire tumor is covered by each beam while as little of the healthy surrounding tissue as possible is hit by the beams. Given the complex spatial arrangement of a patient's anatomy (tumors may be draped around healthy organs or have micro-extensions snaking out into the healthy tissue), this is usually not an easy task.

The experiment described in this chapter was undertaken to shed light on the relative merits of several different *navigation modes* that could be used to target treatment beams. By navigation mode, I mean the mechanism by which the human user explores the patient's anatomy—the method used to move from one prospective beam site to another in the process of finding the best beam configuration. For my purpose, I classified the navigation modes as either *head-tracked* or *non-head-tracked*. Head-tracked modes link the computer generated view to the position and orientation of the user's head. Non-head-tracked navigation modes disregard the head position and orientation, and rely on some other device for changing the view of the anatomy. As originally conceived, the study was to have addressed head-tracked navigation modes only, for since this project

---

Note: Much of the contents of this chapter was published previously (Chung 1992) and is reproduced here with the express written permission of the ACM. The statistical analyses presented herein, however, supercede those of the previous publication, which contained minor flaws.

was aimed at determining the effect of using a head-tracked navigation mode, it was important to find out which head-tracked navigation mode provided the best performance. It became evident, however, that the same investigation should be pursued for non-head-tracked modes, also, for meaningful conclusions could only be drawn if I compared the best head-tracked mode to the best non-head-tracked mode.

4.2. Experimental method

4.2.1. Design

The experiment was a one-factor, within-subject investigation, with navigation mode as the independent variable. Dependent variables measured were final score, task completion time, confidence in the final beam configuration, and rank ordering of the seven navigations modes by ease-of-use and by subject preference.

4.2.2. Subjects

Fourteen subjects were recruited from graduate students and staff members of the Departments of Computer Science, Radiation Oncology, and Radiology at the University of North Carolina at Chapel Hill. Prospective subjects were screened to eliminate those known to be color blind or highly susceptible to motion sickness. Because use of head-mounted displays has been known to induce nausea in some people, signed written consent forms were required from each subject.

4.2.3. Procedure and task

Each subject underwent seven sessions. A different navigation mode was used in each session, and the order of the navigation modes used by each subject was varied according to a 7x7 Latin square. Each session consisted of three practice trials followed by three test trials.

In each trial the subject was presented visually with a different virtual model consisting of a collection of equally sized spheres. This model was intended to serve as an analog to the relevant human anatomy in a cancer case. One sphere, in the center of the model, was the analog for the tumor itself and was known as the *target*. The other balls were analogs for surrounding healthy anatomical structures that should be avoided and were known as the *dodges*. Each model comprised one target at the center of a double-

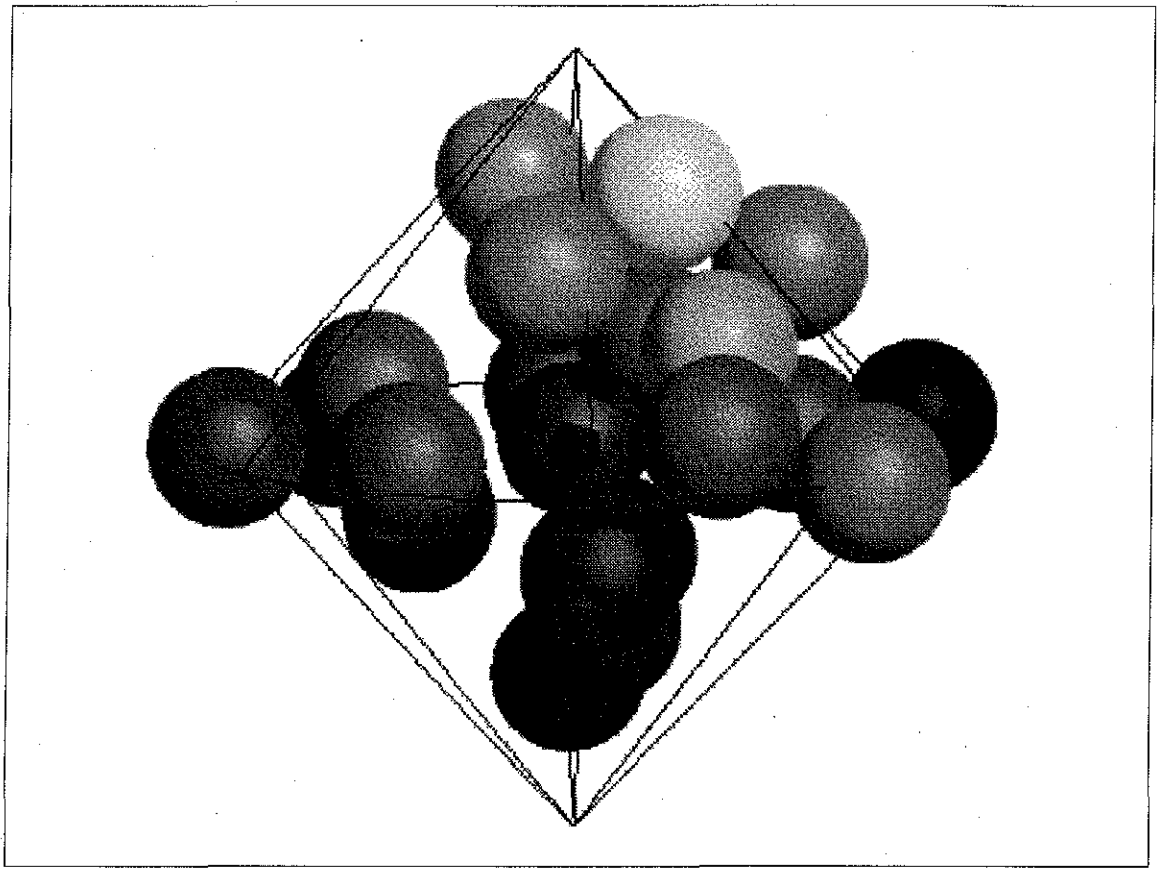


Figure 4-1. One of the models used in the idealized beam targeting task of the preliminary experiment. The central target sphere (tumor analog) is surrounded by a collection of uniquely colored dodge spheres (organ analogs). For clarity, black lines representing the extent of double-cone HLS color space, in which dodge spheres were randomly distributed and which was used to determine dodge color, are shown here, although they were not displayed for the subjects during data collection..

cone (two cones with conjoined bases) volume, and twenty dodges randomly distributed within the double-cone such that no two dodges intersected and no dodge intersected the target.

Anatomical models provide intrinsic context to anyone examining them. Normal human anatomy consists of uniquely shaped structures arranged in a fixed spatial order known well to any physician, who can, with one glance, orient himself in relation to the anatomy. In an attempt to provide some intrinsic context to the random collection of spheres composing the model, each dodge was given the color that corresponded to its position mapped into the double-cone HLS color space. Each sphere was thereby uniquely colored, since no two spheres could occupy the same space. The target sphere was actually a rhombo-icoso-dodecahedron that had its pentagon faces colored differently from its square and triangle faces. Figure 4-1 shows a typical model used in this study.

The task in each trial was for the subject to find the best beam direction through the target/dodge model. The beam was defined to be an invisible cone whose vertex (analogous to the radiation oncology beam source) was a fixed distance from the target, and whose divergence was such that the beam exactly encompassed the target. "Best" direction was defined as that beam orientation which afforded the smallest total volume of intersection between the beam and all the dodges. Exploration of the model was facilitated by one of the seven navigation modes described below. There was no time limit, nor any emphasis on task completion time—the subject was instructed to take as long as necessary to find the best beam path. A virtual *marker* (an arrow pointing through the model) was provided for the subject to use as a reference. At any time the subject could set the marker to be aligned with the current beam direction. The marker would remain fixed in the model until it was reset in a different orientation. When the subject felt that the best beam direction had been found and was represented by his current position, the trial was stopped. Task completion time for the trial was recorded, as well as a score for the designated best beam direction which was equal to the volume of intersection between that beam and all the dodges. Lower scores (smaller intersection volume) were better, and a score of 0 was optimal. Also recorded for each trial was a rating on a scale of 1 (no confidence) to 10 (total confidence) of how confident the subject was that he or she had found the best beam orientation.

After all seven sessions were completed, the subject ranked the seven navigation modes according to two criteria: ease-of-use of the navigation mode, and preference for performing the beam targeting task. Although these criteria may appear to be related they are not equivalent. Ease-of-use refers to the amount of effort required to use a particular navigation mode, whereas preference encompasses the issue of how well a task can be performed with a particular mode. For example, it is conceivable that a navigation mode can be easy to learn and use, and yet not be well-suited to a particular task because it perhaps does not provide the precise control required by the task.

4.2.4. Equipment

The display used for all navigation modes in this study was a VPL EyePhone[1] Model 2 head-mounted display. For the head-tracked navigation modes, tracking was performed by a Polhemus 3Space[2] magnetic tracker. Real-time stereoscopic images dis-

1. Eyephone™ is a registered trademark of VPL Research, Inc.

2. 3Space™ is a registered trademark of Polhemus Navigation Sciences.

played in the head-mounted display were generated by UNC's Pixel-Planes 4 graphics processor (Eyles et al. 1988). Command input was provided by a DragonWriter[3] speech recognition system running on an IBM PC XT[4].

4.3. Navigation modes

4.3.1. Head-tracked

These navigation modes depended upon movement of the subject's head to facilitate exploration of the patient's anatomy. Because these modes translated changes in the orientation and/or position of the subject's head into different views of the patient, they enabled the subject to navigate using the vestibular information provided by the three semicircular canals in the inner ear and kinesthetic information provided by the proprioceptors of the muscles, tendons, and joints.

4.3.1.1. Walkaround

When one first considers using a head-mounted display to perform some task one's first instinct is to place the user in a computer-generated virtual world and to let him move about in the virtual world in the same manner as he does in the real world. This was the approach of Walkaround mode, in which the subject moved about in the virtual world by walking and shifting his body. The Polhemus tracker reported the resulting changes in the position and orientation of the subject's head to the host computer, which in turn appropriately changed the picture seen by the subject in the head-mounted display. Because the picture changed in a manner perceived by the subject to be consistent with his movement, navigating the virtual world should have been as natural as navigating the real world. In actuality, however, several factors contributed to Walkaround mode feeling substantially less than natural. In addition to being subjected to the image lag and distortion problems discussed in Chapter 3, the subject was unnaturally restricted in his movement by the limited range of the tracking device. And the nagging fear that he may collide with or trip over some physical object that is not visible in the virtual world forced the subject to move in a timid and uncertain manner. Nevertheless, Walkaround mode

3. DragonWriter™ is a registered trademark of Dragon Systems, Inc.

4. PC XT™ is a registered trademark of IBM Corp.

offered the advantage that the subject could easily maintain his bearings in the virtual world and determine what actions are required to effect a desired change in those bearings.

For this study the virtual target/dodge model was located at eye height in the center of the working volume defined by the tracker's range. The model was scaled to be roughly 0.5 meters in diameter to allow sufficient room for movement within the bounds of the tracker's working volume (approximately 2.5 meters in diameter), and it was oriented such that its central axis (recalling that the model was built within a double-cone) was vertical with the dark end down and the light end up. While the subject moved about and explored the stationary model from different perspectives the direction of the beam was always defined by the vector pointing from the midpoint of the subject's eyes to the center of the target.

Because human beings have a very limited range of vertical movement, examination of the model from above and below in Walkaround mode was somewhat difficult. For this reason the subject was given the ability to vertically translate the model using a six-dimensional mouse. Although the mouse had three translational and three rotational degrees of freedom, only one degree of freedom, vertical translation, was used to effect a matching vertical translation of the model. When the mouse button was depressed any vertical movement of the mouse would cause similar vertical movement of the model, and with this control the subject was able to move the model to any desired height or depth by repeatedly grabbing-translating-releasing the model. No other manipulation of the model was possible.

4.3.1.2. Walkaround with Rotation

This mode was the same as Walkaround mode, with the exception that the subject used the six-dimensional mouse to also rotate the model in space about any axis through its center.

4.3.1.3. Immersion

Immersion mode placed the subject at the center of the target so that he was always looking outward through the model from the center of the target. Immersion mode ignored the head position information supplied by the tracker, but used the rotational information to change the direction and orientation of the subject's view outward through the model. The orientation of the model was fixed relative to the subject's physical surroundings, so when the subject's head turned, his view swept across portions of the model from its fixed, central vantage point. At all times the current beam direction was defined to be the subject's gaze direction, and the task of finding the best beam orientation became one of looking for the portion of the model with the biggest clear opening. But because his viewpoint was fixed at the center of the target, the subject could only see that portion of the model traversed by the beam as it exited the model after passing through the target. It was equally important for the subject to know what the beam was passing through before it reached the target, i.e. the portion of the model in back of his head. Of course the subject could have just turned his head around and looked backwards, but I felt that the accuracy with which he could align this backward view with the forward view would be too small. Instead, I gave the subject the ability to reverse his gaze direction by holding down a mouse button. No head movement is required. By depressing the mouse button the subject could instantaneously see the reverse view, essentially looking out the back of his head. Releasing the mouse button restored the normal forward view.

4.3.1.4. Orbital

Unlike Immersion mode and the Walkaround modes, Orbital mode had no real-life metaphor to aid the user in grasping how it worked. Consequently, verbal descriptions of Orbital mode sounded somewhat convoluted and confusing. And yet, most users found Orbital mode very easy to assimilate once they actually started using it.

Orbital mode interaction with the target/dodge model followed three rules. First, the orientation of the model was fixed relative to the subject's physical surroundings. Second, the model always remained a fixed distance away from the subject's eyes. Third, the center of the target always stayed in the subject's line of sight. These rules yielded interaction in which only head rotation was used to view the model from different directions. As an example, consider a subject holding his head level and looking straight ahead. He sees a particular view of the model. Now the subject turns his head 90 degrees to the left. Since the model must always stay in the subject's line of sight at a constant distance from the subject's eyes, the model is seen to move about or orbit the subject's head. And because the model's orientation is fixed relative to the subject's environment, the orbital movement of the model consists of pure translation with no rotation. Hence, after turning his head 90 degrees to the left, the subject has a new view of the model that

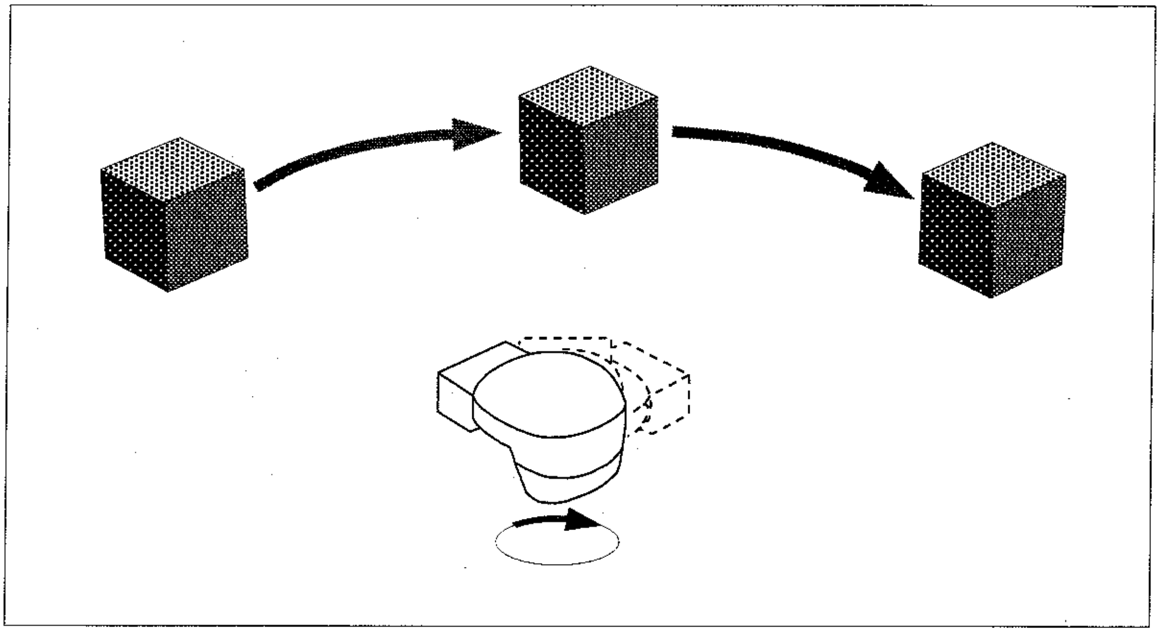


Figure 4-2. Illustration of Orbital mode navigation. Direction from which subject views model is controlled solely by head orientation. (Head-mounted display not shown here for clarity.)

is horizontally 90 degrees away from the view he had when he started. Similarly, when the subject looks down he views the model from above, and looking up yields a view from below. (See Figure 4-2.)

At all times the current beam direction was defined to coincide with the current gaze direction, and the beam source was located at the midpoint between the subject's eyes.

4.3.2. Non-head-tracked

These three modes all placed the beam source at the midpoint between the subject's eyes and aligned the beam direction with the subject's gaze vector. Head-tracking information was not used to change the subject's view of the model. Instead, exploration of prospective beam orientations was facilitated through rotation of the model in space.

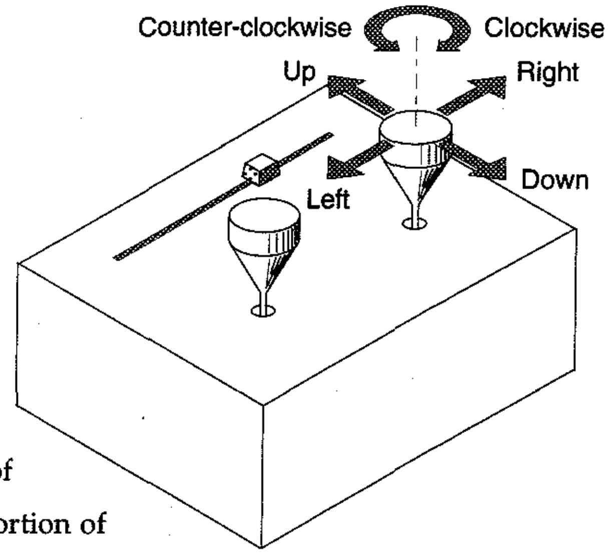
4.3.2.1. JOYSTICK

In Joystick mode the model was rotated with a velocity-control joystick[5]. (See Figure 4-3) In addition to the standard left-right/forward-backward movement of the joystick, the cap of the joystick could turn clockwise and counterclockwise, thereby pro-

5. Model 101-Z, AlphaControl Systems, Inc., Fairfield, Conn.

vide all three degrees of rotational freedom. The joystick was spring-loaded so that it automatically returned to its central rest position when not being manipulated. The mapping of the joystick's movement to model rotation was such that pushing the joystick forward caused a positive rotation (using the right-hand rule) of the model about an axis parallel to the left-vector of the subject's view. In Figure 4-3 this is labeled as "Up," because the near portion of the model moves upward while the far portion of the model moves downward. Pushing the joystick to the right rotated the model about an axis parallel to the subject's up-vector such that the near portion of the model moves to the right, and turning the joystick knob clockwise rotated the model clockwise about the subject's gaze-vector. Any combination of these three rotational controls was possible, providing the ability to rotate the model about any arbitrary axis that passed through the model's center. The joystick was used as a velocity-control device and the speed of model rotation was directly related to the magnitude of the joystick's deflection from its normal rest position.

Figure 4-3. Illustration of joybox used in the preliminary study, showing rotation control provided by joystick deflection. Labels indicate direction of movement of near portion of model resulting from indicated joystick deflection.



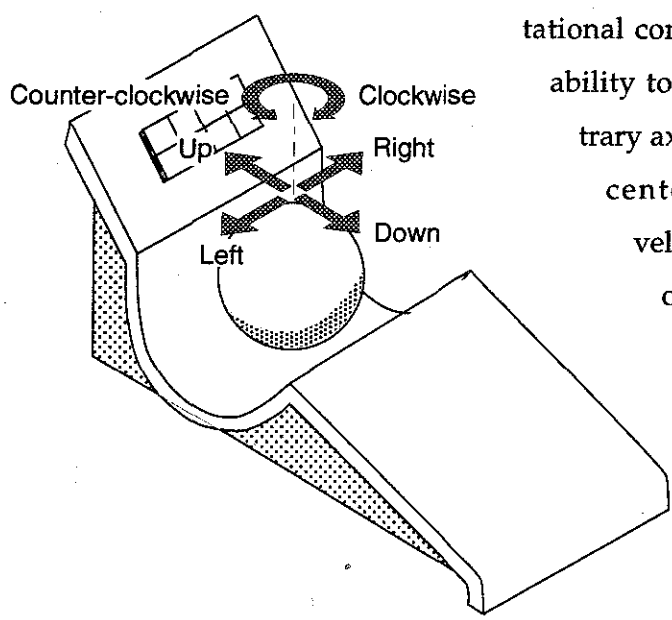


Figure 4-4. Illustration of Spaceball, showing rotational control provided by exerting torques on the control ball.

4.3.2.2. SPACEBALL

In this mode the model is rotated with a Spaceball[6], an

6. Spaceball™ is a registered trademark of Spatial Systems, Inc.

isometric, force-sensitive device that provided six degrees of translational and rotational freedom. (See Figure 4-4.) This mode, however, used only the three rotational degrees of freedom as a velocity control for rotation of the model in three-space, such that the speed of rotation was directly related to the magnitude of the torque exerted on the ball. The mapping of Spaceball action to rotational control was similar to that used with with joystick described above. Rotating the control ball forward about its stem rotated the model about the left-vector. Rotating the ball about its stem to the right rotated the model about the up-vector, and twisting the ball clockwise rotated the model clockwise about the gaze-vector. A major difference between operating the Spaceball and operating the joystick was that the Spaceball deflected very little, perhaps imperceptibly, and responded to the forces exerted on it by the user's hand, whereas the joystick responded to the magnitude of its deflection.

4.3.2.3. MOUSE

In Mouse mode the orientation of the model was controlled with a six degree-of-freedom mouse, constructed in our shop by simply embedding a tracker sensor in a hollowed-out pool ball and mounting two buttons on its surface. (See Figure 4-5.) This is the same mouse as that used in the Walkaround modes for manipulation of the model, but in this mode mouse translation was ignored and only its rotation was used. When either mouse button was held down, the mouse became a position-control for model orientation, as opposed to the velocity control provided by the joystick and the Spaceball. When a button was depressed, the rotational component of the mouse movement was directly linked to model rotation, and the subject saw the model rotate about its center in the same direction as the mouse in his hand was rotating. Releasing the button uncoupled model rotation from mouse movement.

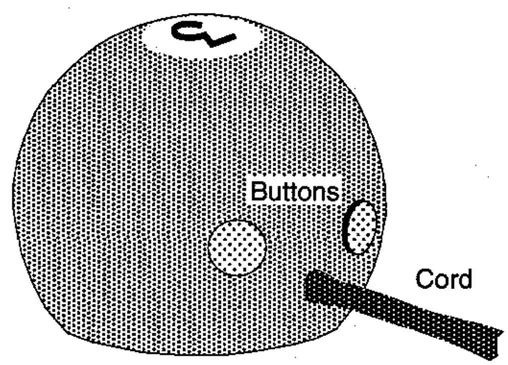


Figure 4-5. Illustration of six degree-of-freedom mouse constructed by embedding a tracker sensor in a billiard ball and mounting two buttons.

4.3.3. Beam's-eye view

In radiation treatment planning the term *beam's-eye view* refers to the perspective view seen by an eye that is coincident with the beam source and whose gaze vector coincides with the beam's central axis. (Goitein, Abrams, et al. 1983) Since the diverging rays of the beam follow the lines of sight that make up the perspective view, the shape of the beam can be represented by a single closed contour in the perspective view. The advantage of this representation is that it becomes very simple to determine which anatomical structures are being hit by the beam. Those structures whose silhouettes intersect or are enclosed by the beam boundary contour will be hit by the beam. Those structures whose silhouettes are completely outside the beam boundary contour will not be hit by the beam. Such a determination is much more difficult to make without the beam's-eye view.

The impact of using a beam's-eye view was important in this study, as four of the seven navigation modes studied (Orbital, Joystick, Mouse, and Spaceball) provided beam's-eye views, while the other three (Walkaround, Walkaround/Rotate, and Immersion) did not. With beam's-eye view it was very easy to determine which dodges intersected the beam, for since the beam was defined to diverge just enough to exactly enclose the target, the silhouettes of those dodges would have overlapped with the silhouette of the target. Technically, neither eye's view was strictly a beam's-eye view, because the beam source was assumed to be located at the midpoint between the eyes. Even so, the work of Linksz (1952) suggests that the two flanking views seen by the individual eyes are combined by the subject's perceptual system to create a percept equivalent to the true beam's-eye view that would be seen by a "cyclopean" eye located midway between the subject's eyes.

In Walkaround and Walkaround/Rotate modes beam's-eye views were possible, but somewhat difficult to achieve because they required the subject to look at the target while holding his head at the same distance from the target as the beam source was. Since there were no physical restraints on the subject's head holding him in that position, beam's-eye views were more the exception than the rule with these modes. It could be argued, however, that being slightly off in distance to the target would still provide a view close to a true beam's-eye view, and make possible reasonably good estimates of what was being hit by the beam.

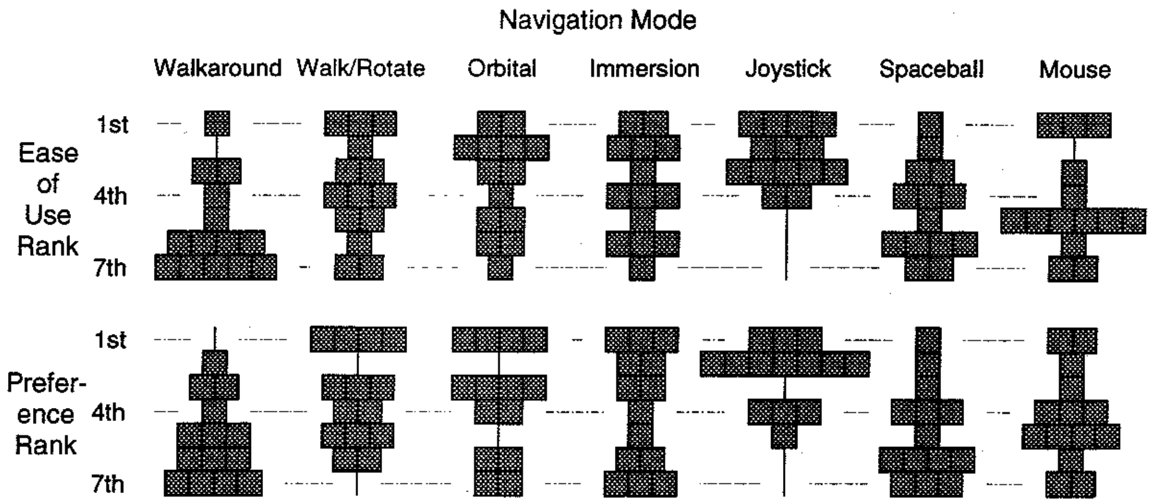


Figure 4-6. Vertical centered histograms of rank assigned to each navigation mode for ease-of-use and preference. Each square represents one subject assigning that particular rank to that particular mode.

Immersion mode, on the other hand, presented the subject with a much more difficult situation. There was no beam's-eye view, since the subject's eyepoint was constrained to stay at the target's center, and in the view seen by the user there was no indication of the location of the beam boundary. Prior experimentation with several beam representations revealed all to be confusing, and I felt that the subject was better off using his own judgement. Under these conditions, the task of finding the clearest beam path was approached by finding the biggest holes in the model, with little attention paid to actually estimating the volume of intersection between the dodges and the beam.

4.4. Results

Figure 4-6 presents histograms showing for each navigation mode, the number of times it was ranked 1st, 2nd, ... 7th by ease-of-use and by preference. The arrangement and shape of the histograms permits easy comparison between the navigation modes. The bottom-heavy shape of the Walkaround mode ease-of-use histogram shows that most subjects found it to be one of the more difficult steering modes to use. The ability to manipulate the model, however, definitely made things easier, for the ease-of-use histogram of Walk/Rotate mode has its bulk shifted up to the middle and upper rankings. Immersion mode has a somewhat uniform histogram, suggesting no general consensus on how easy it was to use. Orbital mode has a slightly top-heavy histogram, indicating a bias toward it being one of the easier navigation modes to use. The consensus for Joystick mode appears to be that it was very easy to use, judging from the extremely top-heavy shape of its histogram. Spaceball mode, on the other hand, appears to be a not-so-easy mode to use, as does Mouse mode.

The preference rank histograms clearly show that Joystick mode was widely preferred by the subjects, whereas Walkaround Mode was widely disliked. For Spaceball and Mouse modes, subjects' opinions were on the less-preferred sides, slightly more so for Spaceball mode than for Mouse mode. On the other hand, the histograms for Walk/Rotate and Orbital modes are biased toward the more-preferred side of the scale. Immersion mode shows an interesting bimodal distribution, which suggests that subjects either liked it or disliked it relative to the other modes.

Figures 4-7 through 4-9 summarize the three dependent variables measured for each trial: score (equal to the volume of intersection between the beam and all the dodges), task completion time, and subject confidence. For each subject, performance in the three trials for each navigation mode were averaged together to produce one data point. Figure 4-7a shows the distribution of the scores grouped by navigation mode. The best possible score is 0, indicating no intersection between beam and dodges, and for comparison, the score resulting from the beam completely enclosing one and only one dodge sphere is 305.4. The order of the seven modes from best (lowest score) to worst (highest score) is shown in Figure 4-7b, and statistically significant differences revealed by paired Student's t-tests performed on pairwise mode comparisons are indicated by the arcs. Taken as a group, the head-tracked navigation modes (Walkaround, Walk/Rotate, Immersion, Orbital) did not differ significantly from the non-head-tracked modes (Joystick, Spaceball, Mouse).

Figure 4-8a shows the distribution of the task completion rates grouped by navigation mode. The order of navigation modes from best (shortest time) to worst (longest time) is shown in Figure 4-8b, and significant differences are indicated by the arcs. As with the scores, no significant difference was found between head-tracked and non-head-tracked modes.

Figure 4-9a shows the distribution of the confidence ratings grouped by navigation mode. The order of navigation modes from greatest confidence to lowest confidence

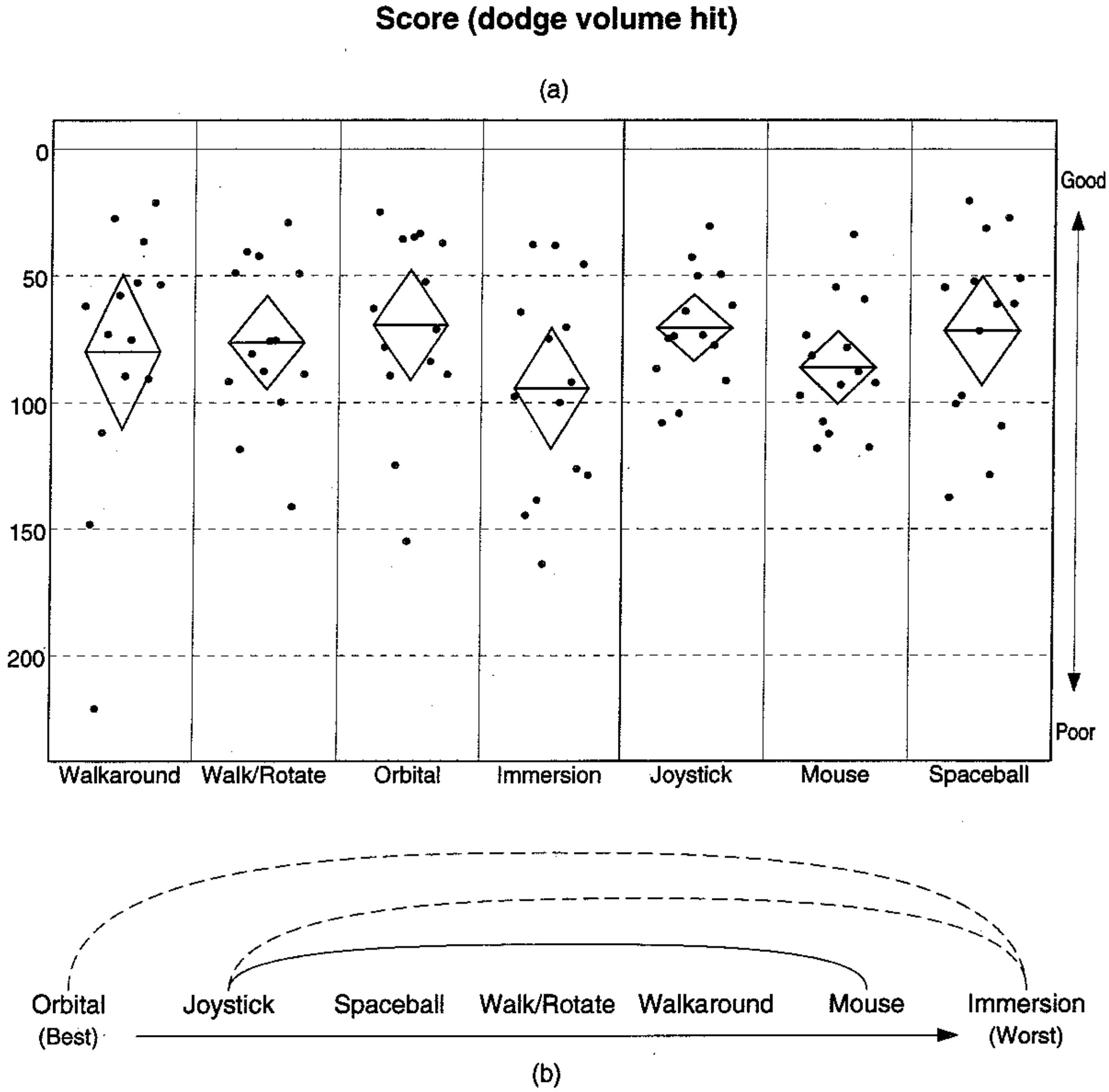


Figure 4-7. Analysis of scores obtained in preliminary study. Part (a) displays distribution of scores grouped by navigation mode, with means and 95% confidence interval represented by diamonds. Part (b) shows resulting order of navigation modes from best to worst. Solid arcs connect navigation modes whose paired t-test produced differences significant at the α=0.05 level. Dashed arcs represent differences significant at the α=0.10 level.

is shown in Figure 4-9b, and the only significant difference revealed by paired t-test is indicated by the arc. As with the scores and task completion rates, no significant difference was found between head-tracked and non-head-tracked modes.

Table 4-1 presents correlations between the dependent variables. Not surprisingly, ease-of-use and preference rank are highly correlated. Significant correlations were also found between subject confidence and ease-of-use, preference, score, and task rate.

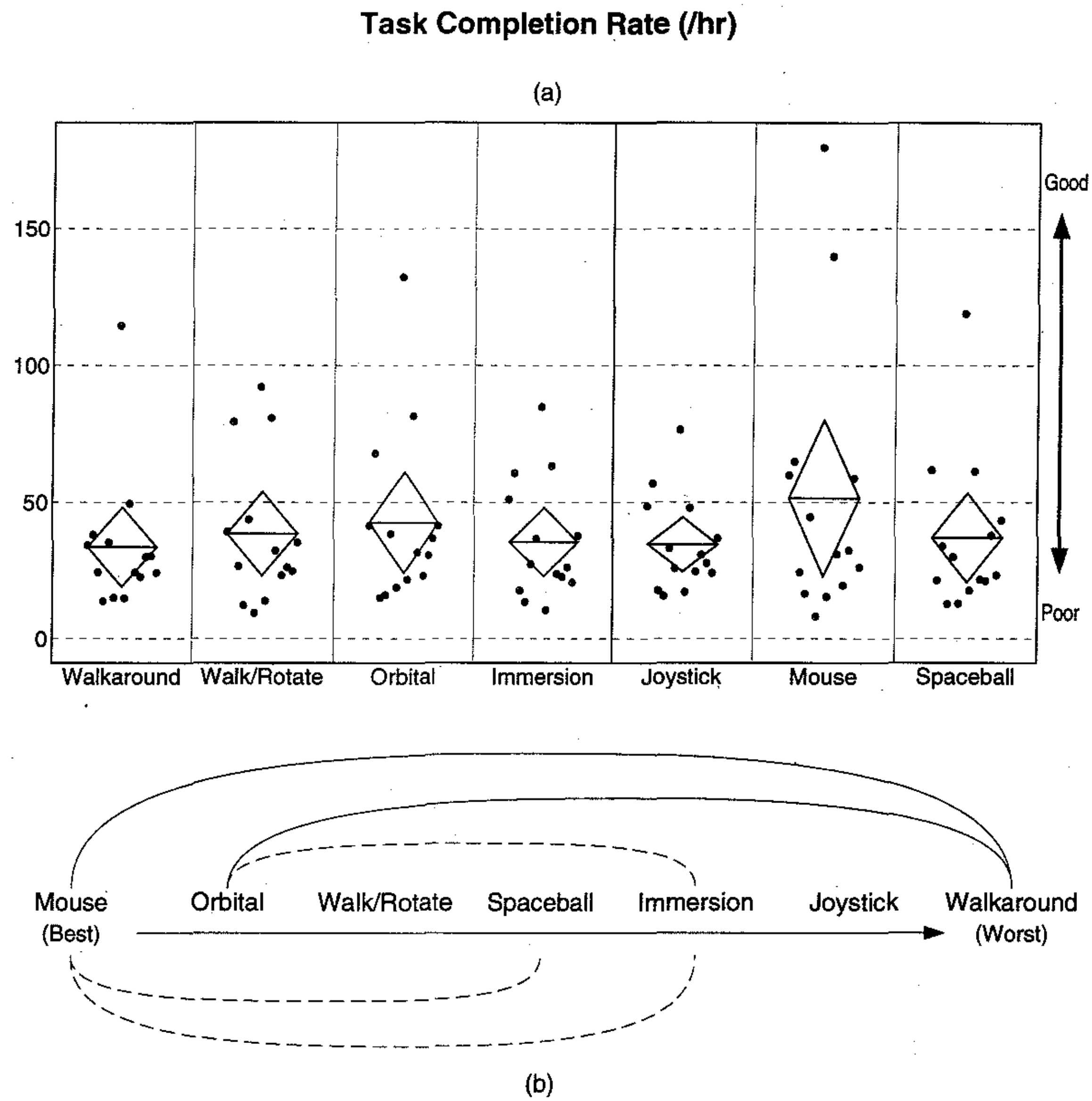


Figure 4-8. Analysis of task completion rates obtained in preliminary study. Part (a) displays distribution of rates grouped by navigation mode, with means and 95% confidence interval represented by diamonds. Part (b) shows resulting order of navigation modes from fastest to slowest. Solid arcs connect navigation modes whose paired t-test produced differences significant at the α=0.05 level. Dashed arcs represent differences significant at the α=0.10 level.

These coefficients suggest that subjects had lower confidence in their performance when using difficult navigation modes (higher ease-of-use ranking) or modes they did not like (higher preference ranking). Greater confidence appears to accompany better performance (lower score) and faster performance (higher rate).

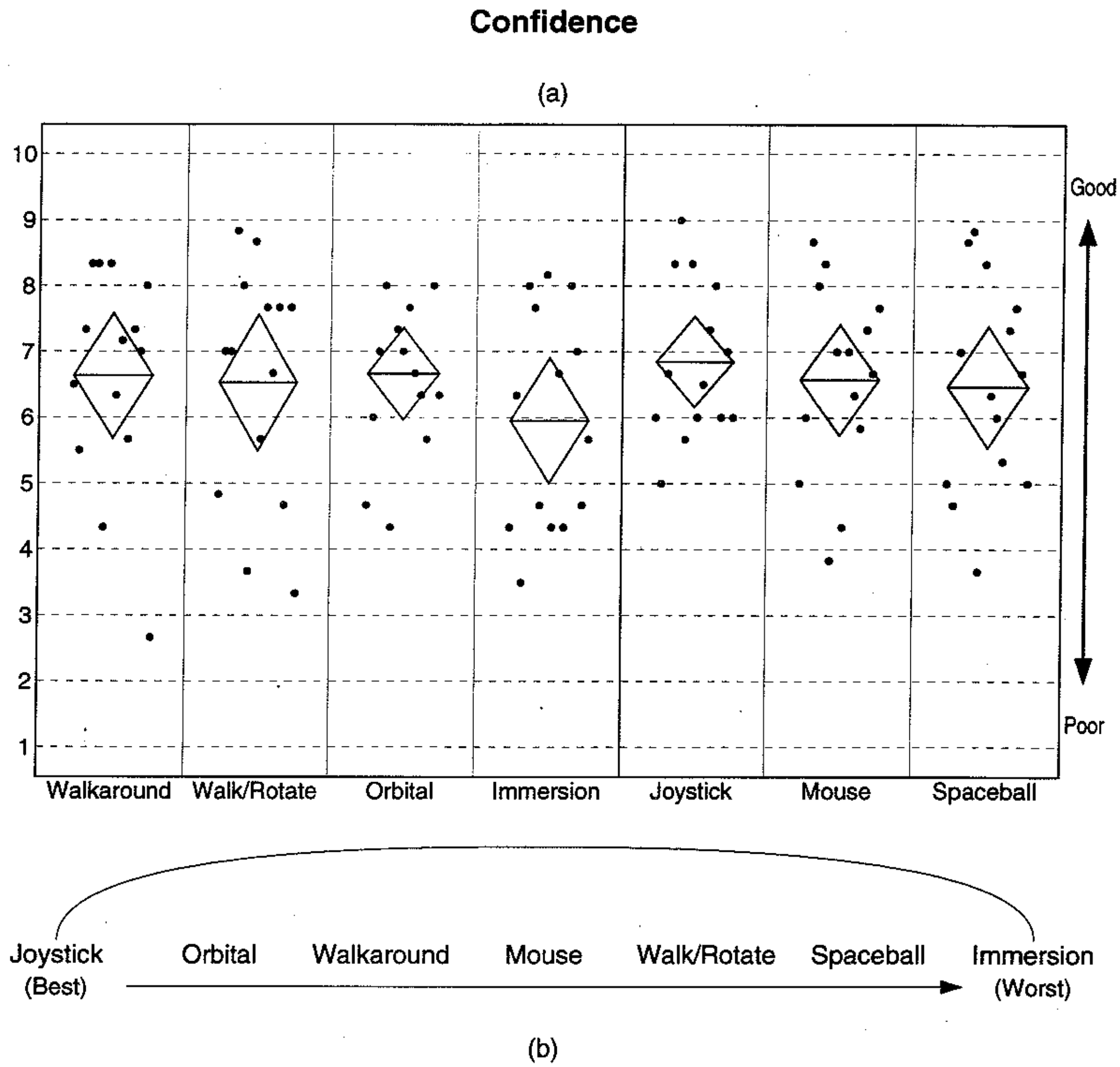


Figure 4-9. Analysis of confidence scores obtained in preliminary study. Part (a) displays distribution of confidence scores grouped by navigation mode, with means and 95% confidence interval represented by diamonds. Part (b) shows resulting order of navigation modes from greatest to lowest confidence. Solid arcs connect navigation modes whose paired t-test produced differences significant at the $\alpha=0.05$ level. Dashed arcs represent differences significant at the $\alpha=0.10$ level.

4.5. Discussion

4.5.1. Trial replay

In addition to the statistical analysis presented above, a subjective review of each trial was conducted by replaying the subject's actions as recorded in a log file at the time of the trial. These logs contained status information for the subject's head, the model, the marker and the beam at roughly half-second intervals. I replayed each trial and observed the movements of the subject and the model while wearing a head-mounted display,

	Confidence	Task Rate	Score	Preference Rank
Ease of Use Rank	**-0.1950**	-0.0332	0.0256	**0.8680**
Preference Rank	**-0.2258**	-0.0287	0.0093	
Score	**-0.2101**	-0.0250		
Task Rate	**0.1993**			

Table 4-1. Correlation coefficients for pairs of dependent variables. Coefficients in bold indicate correlations significant at the α=0.05 level.

which afforded me a "god's-eye-view" of the trial. For those navigation modes in which the model's orientation was fixed relative to the subject's environment (Walkaround, Immersion, Orbital), the trial playback consisted of a stationary model, about which moved a representation of the subject's head. For the non-head-tracked navigation modes (Joystick, Spaceball, Mouse) the subject's head was held stationary, while the model rotated under the subject's control. The replay for Walk/Rotate mode contained elements of both model rotation and head movement.

For all modes the trial replay also displayed the path of the beam source as it developed through the trial by connecting successive source positions with a line segment. This trace quickly revealed which beam directions were considered, and perhaps more important, which directions were not considered. If a subject had considered all possible beam directions, then the resulting trace of the beam source should resemble a spherical surface surrounding the target. Holes in this spherical shell would be indicative of possible beam directions that were not visited by the subject.

In spite of being instructed to find the best possible beam direction, subjects usually terminated the trial before considering all possibilities. This was evidenced by the incomplete spheres traced out by the beam source. Trials in which the model had been completely covered were usually of extremely long duration, with subject movement suggesting confusion and disorientation. In only a few cases did subjects follow a systematic search strategy, and these systematic searches would usually be abandoned after the first good candidate beam direction had been found. For the most part, subjects followed what might be called a "greedy" steering strategy, moving about the model in a manner dictated by their current view of the model, and not by some global plan. As a result, in most trials the traces of the beam source showed large holes that were never considered. Some of these areas corresponded to beam directions that were obviously bad, while other holes contained prospective beam directions that were good enough to deserve consideration. In either case it was impossible to determine from just watching the replay whether the holes represented areas that were deliberately skipped or unintentionally missed. In the case of the former, the subject made a conscious decision, wisely or unwisely, to not investigate a particular area. In the case of the latter, the subject was not aware that he had missed that area, indicating a deficient mental model of his search process. In most cases the beam directions that required the subject to look straight up or straight down were not covered, as the head-mounted display would exert uncomfortable torques on the subject's neck in these positions.

Most subjects relied very heavily on the marker to provide a reference point in the model. One common use of the marker was to set the marker at a good prospective beam location, and then move to the opposite end of the marker to investigate the opposite direction. It was reasonable for the subjects to assume that the opposite of a good prospect might be a good prospect itself, although they wouldn't be equivalent because of beam divergence. The marker made possible quick and efficient evaluation and comparison of such opposed candidates. The marker was also typically used as a "best-beam-direction-so-far" placeholder, from which the subject would venture in search of a better prospect. If such an alternative were found, the marker could be reset to mark it. Otherwise, the subject could return to the marked position. Many subjects expressed a desire to have the use of more than one marker, and, in retrospect, multiple markers or unlimited markers might have contributed to better performance.

As discussed above, the dodges were uniquely colored to provide some intrinsic context to the model. The subjects were not told that the dodges were uniquely colored, nor that the dodges' colors were dictated by the HLS color space. Consequently, most subjects did not make use of the color context. Although many did comment on the visually pleasing colors used in the model, most subjects did not see them as an aid to navigation. Instead, more reliance was placed in the context provided by the geometric arrangement of the model. Only one subject, whose own research is concerned with the use of color, found the colors useful—so useful, in fact, that she never used the marker.

4.5.2. Navigation mode summaries

4.5.2.1. Walkaround

Walkaround mode yielded the lowest task completion rate, but was undistinguished in score and subject confidence. The low task completion rate is not surprising, given the difficulty of walking about in a virtual world while wearing an immersive head-mounted display that seals off any view of the real world. Most subjects found this mode very awkward and time-consuming, and ranked Walkaround low in ease of use and preference. Interestingly, this mode more than any other was used for systematic searches. One subject repeatedly circled around the model, inspecting the model at a different height with each loop—effectively performing a latitudinal scan. Another subject opted to walk less and walk around the model just once. At regular intervals in his trip around the model this subject would stop translate the model up and down, scanning the model longitudinally. Perhaps the awkwardness of Walkaround mode instilled in these subjects a need for a disciplined, efficient approach.

An interesting observation is that most subjects preferred to walk around the virtual model to get to the other side, rather than take the shorter route directly through the model. For some, the option of walking through the model just never occurred to them. Could this be evidence that perhaps the feeling of presence in the virtual world was strong enough that the subject subconsciously accepted the existence of the model hanging in the middle of the room, and thereby precluded any notion of walking through them. Comments from a handful of subjects indicated that they walked around the model to continue exploring possibilities, and that they preferred to keep the model at a comfortable distance, where they could make better use of it as a visual reference for maintaining balance and orientation.

4.5.2.2. WALKAROUND WITH ROTATION

Walk/Rotate mode did not perform much differently than Walkaround mode in terms of score, task completion rate, and confidence, but the addition of the model rotation capability produced better ease-of-use and preference rankings. Model rotation was used to different degrees by the different subjects. Most subjects walked very little and spent most of their time standing still and using the mouse to rotate the model. Small head movements were used to fine tune the beam direction after gross model rotation by hand. On the other hand, some trials showed no rotation at all, perhaps indicating a reluctance in those subjects to lose the navigational advantage provided by a fixed model reference frame.

4.5.2.3. Immersion

Immersion mode produced significantly worse scores than Orbital and Joystick modes, and it appeared to have instilled less confidence in the subjects than the other modes. This may be a result of the subjects' being able to see only a small portion of the model at any time, which, combined with the lack of any head-motion parallax, could have hindered the subject's development of a complete mental picture of the model. In addition, subjects were required to evaluate prospective beam orientations by looking in one direction and then in the other direction, with no clear indication of where the boundary of the beam was. Immersion mode did, however, have the advantage of providing the ability for the subject to use muscle memory in navigation. Even without a complete global understanding of the model, subjects knew how they had to orient their heads to get back to a particular beam direction.

4.5.2.4. Orbital

Despite the fact that there is no real-world metaphor for Orbital mode, this navigation mode produced significantly better scores than Immersion, Mouse, and Walk/Rotate modes. This may have been due to the unique combination of several factors. Orbital mode provides a beam's-eye view of the model, which at once gives the subject an external global view of the model and allows the subject to easily determine which dodges intersected the beam. In addition, the fixed orientation of the model relative to the subject's environment provided kinesthetic information that aided navigation. As with Immersion mode, muscle memory could be used in Orbital mode to provide reference landmarks in the exploration of the model.

4.5.2.5. Joystick

Joystick mode ranked very high in ease-of-use and preference, which is not surprising in light of the fact that most of the subjects worked with computers and were somewhat familiar with video games. In addition, Joystick mode performed well in terms of score and subject confidence, but second to Walkaround mode as the slowest navigation mode. Trial replay revealed that most subjects used only principal axis rotations, i.e. they rotated models mostly vertically and horizontally and very little diagonally. This was probably due to the mechanics of the joystick used, which in addition to returning automatically to a centered, rest position, also exhibited a secondary preference to remain in one of the cardinal axes. Deflecting the joystick in an oblique direction required more effort than and was less precise than simple deflection along an axis. The effect of this joystick behavior is unclear, for while it encouraged subjects to decompose their movements into a series of principal axes rotations, it provided a precision of movement not available with the other non-head-tracked modes.

Rotation about the gaze direction, invoked by twisting the joystick cap, was not used to any great extent. Such rotation did not change the subject's beam direction, and the model contained no apparent vertical which a subject might prefer to align with his own vertical. One subject complained that the mapping of joystick movement to model rotation was counter-intuitive. Tilting the joystick so that it pointed at the subject and was anti-parallel to the gaze direction, rather than pointing vertically off a tabletop, alleviated this problem. In such a configuration the side-to-side movement of the joystick better matched the resulting model rotation about the vertical. Likewise, the rotation of the joystick cap better matched the model rotation about the gaze direction.

4.5.2.6. SPACEBALL

The performance of Spaceball mode was relatively undistinguished. Its preference rankings, however, were heavily weighted toward the low end, because many subjects found the Spaceball's isometric action fatiguing and difficult to use for precise movements. McKinnon and Kruk (1991) have suggested that a proportional displacement, multi-axis controller would provide better performance than an isometric device such as the Spaceball. Their evaluation of controllers for the Shuttle Remote Manipulator System revealed greater inter-subject variance with isometric devices than with displacement controllers. Unlike the joystick used in Joystick mode and discussed above, the Spaceball had no inherent preference for rotation about the cardinal axes, and this was evident in the beam source trace, which showed much diagonal movement.

4.5.2.7. MOUSE

Compared to subjects' preference for Joystick mode and dislike for Walkaround mode, preference response to Mouse mode was relatively flat. This mode boasted the fastest task completion rate, and yet averaged the second lowest scores of all the navigation modes. Trial replays showed that this mode suffered greatly from system latency, which greatly hindered both precise alignment and large-scale rotation that required more than one grab-release cycle. Fine alignment was difficult because the subject's releasing of the mouse button to release the model depended upon an image that was not current. Instead of releasing the model at the intended point, the button release would not be acted upon until the next program cycle, which would also obtain new information from the tracker. If this tracker report was not identical to the previous one, which is often the case with a hand held, magnetically-tracked device, then the subject would see the model rotate slightly beyond his intended release point before stopping. Only with difficulty and concentration could the subject hold the mouse still enough to rotate the model exactly as he desired. For large rotations that could not be comfortably handled by grabbing and releasing the model once, the subject had to repeat the (grab model—rotate mouse—release model—rotate mouse back) cycle several times. Trouble with this maneuver arose from the discrepancy between information provided by the tracker and information provided by the analog/digital controller for the mouse buttons. While the A/D controller provided almost instantaneous response to changes in the states of the mouse buttons, the information from the tracker was not nearly as up-to-date. The result was that the system would think that the subject pressed the mouse button earlier relative to the hand movement than the subject intended. As an illustration, consider the action of the subject rotating his wrist counterclockwise, grabbing the model by pressing the mouse button, and then rotating his wrist clockwise. The subject may think that he did not press the button until the counterclockwise wrist rotation had stopped, but because of the latency within the tracker, the system saw the button pressed before the counterclockwise rotation was completed. The resulting images that the user sees show the model briefly rotating counterclockwise before beginning a clockwise rotation. Both latency-based behaviors were considered very annoying by the subjects, and they manifested themselves in the beam source traces as jagged paths resulting from large direction changes separating relatively small rotations.

4.5.3. Inter-subject variability

One of the most notable characteristics of the data collected in this study is the large inter-subject variability, which can be found not only in the measured dependent variables of score, task completion rate, and confidence, but also in the ease-of-use and preference rankings. Figures 4-10 through 4-12 show the measured dependent variables grouped by subject, and significant differences between subjects are plainly evident. As an example, compare subjects 4 and 5. Subject 4 typically worked more quickly than subject 5, and generally felt more confident in his results than subject 5. Even so, subject 4's scores tended to be worse than those of subject 5. Subject 2's scores are significantly greater (worse) than subject 9's, even though they both worked at about the same pace and felt similarly confident about their performances. Such gross differences between subject performances make it difficult to detect any effect due to the different navigation modes. Perhaps standardized tests of spatial orientation and spatial visualization (McGee 1979) could be used to provide a normalizing factor to reduce this variability. For example, one might find that subjects with poor spatial abilities took unusually long times to complete the task, or that people who scored well on the standardized tests were more confident in their solutions. Such information might prove useful in identifying those who would benefit most from using a head-mounted display.

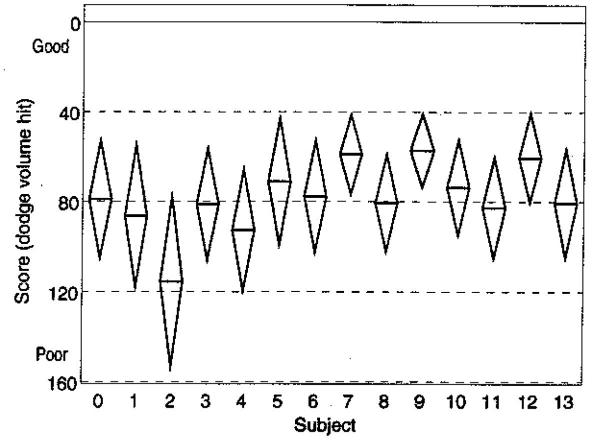


Figure 4-10. Score means by subject with 95% confidence intervals.

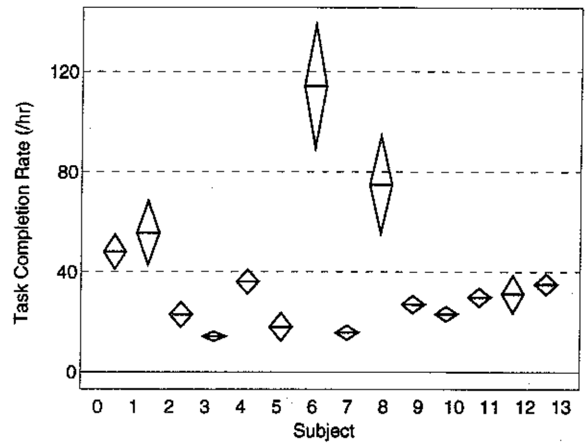


Figure 4-11. Task completion rate means by subject with 95% confidence intervals.

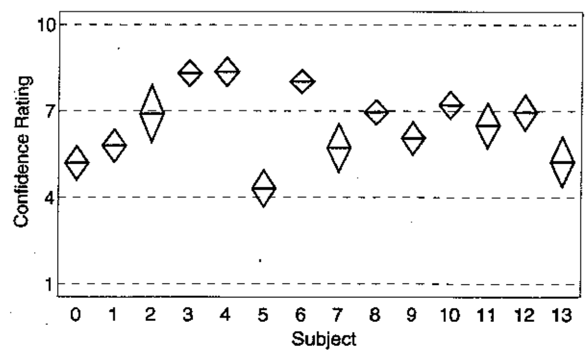


Figure 4-12. Confidence rating means by subject with 95% confidence intervals.

4.6. Conclusion

The goal of this experiment was to identify which head-tracked navigation mode and which non-head-tracked navigation mode were best suited to the task of targeting radiation treatment beams. These two modes would then be carried forward into the next stage of this research, a comparison study of the two modes using real radiation oncologists working on real cancer cases. The data was by no means conclusive, and yet, I do not feel that I went too far out on a limb in selecting Orbital mode and Joystick mode to use in the next study. Joystick mode was widely preferred by the subjects, contrasting sharply with Spaceball and Mouse modes, both of which generated a significant number of complaints. Joystick mode also averaged lower (better) scores and greater confidence than the other non-head-tracked modes. Orbital mode was easier to use and more preferred than the other head-tracked navigation modes, and it yielded better scores, faster task completion and greater confidence than any of the other head-tracked modes. Additionally, it was the only head-tracked mode that provided the user with a beam's-eye view of the model. Because of the obvious advantages it provides in the assessment of a beam's efficacy, beam's-eye view has become an indispensable component of modern radiation treatment planning systems.

Chapter 5

User Study: Method

Though interesting, the information provided by the preliminary experiment described in Chapter 4 did not answer our question of whether a head-tracked navigation mode makes for better treatment plans. The task was only an idealized analog of targeting beams in radiation treatment planning, and the people performing the task were not radiation oncologists. Moreover, the results did not show a significant difference between head-tracked and non-head-tracked modes, even for this task. The answer to this question had to come from a full user study involving real radiation oncologists working on real cancer cases. The results of the preliminary experiment, though, were useful in narrowing the focus of this user study. Given the results of the preliminary experiment discussed above, the full user study compared beam configurations produced by subjects only using either the Orbital navigation mode, the best of the head-tracked modes, with those produced using the Joystick mode, the best of the non-head-tracked modes. As with the preliminary experiment, I expected that the intuitive steering and navigation provided by the head-tracking Orbital mode would enable better beam configurations to be designed.

It should be noted that full treatment plans were not produced and compared. The main concern here was the effect intuitive steering has on the basic geometric problem of beam targeting, i.e. finding the best position and direction for each beam. For this reason, the term *configuration* is used here to mean the collection of beam positions and orientations that would normally undergo further development to produce a full radiation treatment plan. There was no beam modification through filters, wedges, collimators or whatever, and no dose computation and examination, as these aspects of treatment plan design are not relevant to the question of intuitive steering.

5.1. Subjects

Volunteer subjects were recruited from the readily accessible pool of radiation oncologists, dosimetrists, and radiation physicists practicing in the Radiation Oncology Department of the University of North Carolina Hospitals. All thirteen subjects were familiar with the principles of radiation oncology and the process of planning radiation therapy. The subjects received no immediate benefit from participation in the study, nor were they provided with any tangible inducement for their participation. Prospective subjects were rejected if they were known to lack stereoscopic vision, or if they were known to be susceptible to motion sickness.

5.2. Design

The study was designed as a single-factor within-subject investigation with repeated measures. The independent variable was navigation mode, which took the nominal values of Orbital mode and Joystick mode. Dependent variables measured were subjective and objective quality scores (see Section 5.4) and task completion rate for each beam configuration, all of which were interval[1] quantities.

5.3. Procedure

5.3.1. Case studies

The six case studies used in this experiment were drawn from the radiation oncology archives of the UNC Hospitals Department of Radiation Oncology. All cases involved lung tumors, because the targeting of treatment beams for such tumors was

1.A variable must belong to one of three basic types of measurement scales. *Nominal* scales simply label observations so that they fall into different categories. With *ordinal* scales, observations are ranked in terms of size or magnitude. In *interval* scales, differences between observation values reflect differences in magnitude.

deemed to be sufficiently challenging, and lung cancer cases were numerous enough to find the required number of cases that were unfamiliar to the subjects. The experiment was designed to use six cases to improve the generalizability of the results to the domain of lung tumor treatment planning. This strategy reduced the power of the experiment by introducing variability among the different cases, and traditional experimental design principles would have dictated that only two cases be used to minimize this variability and to improve the chances of obtaining significant results. I was advised, however, by Dr. Keith Muller of UNC's Department of Biostatistics that it was more important to adequately represent the inherent stimulus domain variability. I therefore traded off sensitivity to navigation mode effects for increased generalizability.

Per UNC Hospitals's policy, all patient identification was removed from the case data. All that was required for this study was the three-dimensional CT data for the patient, which was processed for use in this study in the following manner.

First, closed boundary curves, called *contours*, were constructed for each anatomical structure pertinent to radiation treatment of lung cancer. These were the skin, heart, trachea/bronchi, lungs, spinal cord, and treatment volume. The treatment volume was defined as the radiation oncologist's best estimate of the extent of the tumor, surrounded by a safety margin to account for errors in tumor location judgment and unseen extensions of the tumor into the surrounding tissue. Upon completion, the collection of contours constituted a stacked "wire-loop" representation of the patient's anatomy.

In addition to the three-dimensional CT dataset, the archived case data also included anatomical contours which had been used in the original treatment planning for the patient. Although this greatly reduced the amount of data preparation required for the study, extensive editing of the contours was still required to remove duplicate vertices and contours, repair self-intersecting contours, and smooth out unnecessary jags. In addition, none of the archived datasets included heart contours, even though the heart is deemed an important radiosensitive organ, so heart contours had to be generated for each case. Contour editing was performed with *imex*, an X-Windows-based tool developed by the Department of Radiation Oncology.

Once generated, the contours for each anatomical structure were tiled to produce surface representations, again using programs developed by the Radiation Oncology Department. In addition, the treatment volume contours were used to compute the centroid of the treatment volume, which served as the *isocenter*, or center of focus, for the subsequent exploration of the anatomy by the subjects during beam-targeting.

5.3.2. Criteria survey

The first step in the subject's participation in the study was the completion of a survey concerning the criteria used by the subject in the evaluation of beam configurations. The subjects were presented with five criteria that may be used in assessing the quality of a beam configuration. These criteria were selected after discussion with Dr. Julian Rosenman of the Radiation Oncology Department. For each criterion, the subjects were asked to assign an importance weight, ranging from "Unimportant" to "Very Important." The subjects were also given the opportunity to add any additional criteria they felt were appropriate along with weights for these criteria. Figure 5-1 presents a sample Criteria Survey form.

5.3.3. Beam targeting

The second step of each subject's participation was the design of a treatment beam configuration for each of two cases using *beattl*, a test platform program written for this study that enabled the subject to use any of a variety of navigation modes to explore the patient's anatomy and search for optimal beam orientations. In this stage of the study the subjects were asked to design the optimal beam configuration for the given lung cancer case. "Optimal" beam configuration was defined to be that configuration which provided the best prospect for further development into a complete treatment plan. Note that this is not equivalent to finding all the best individual beam directions, for a an optimal beam configuration must also account for interactions between different beams. For example, there may exist two good beam directions that are separated by some small angle. Although each beam may do a good job of avoiding radiosensitive healthy tissue, the two beams would probably not both belong to an optimal beam configuration, because the small separation between the two might result in excessive exposure of healthy tissue.

Table 5-1 presents the navigation mode and case used in each trial by each subject. Each subject alternated use of the two navigation modes through their four trials, but were given a practice trial before the first trial of either mode. The emphasis of this practice run was on familiarization of the subject with operation of the navigation mode and with the program's commands and capabilities. The practice trial was terminated when

Beam Targeting Experiment

Criteria Survey

Subject:____________________________________

A) Please list and describe below the criteria you would use to evaluate a given treatment beam configuration. (Where *treatment beam configuration* refers to the geometric arrangement of a collection of treatment beams, as defined by their source positions and directions of their central rays, from which a complete treatment plan can be designed. There is no notion of beam modification or dosimetry associated with a beam configuration.) Some possible criteria are already listed. Use the blanks to add any other criteria you think are important.

B) For each of the criteria listed, please indicate the weight or importance that that factor has in your overall judgment of a beam configuration.

Criterion	Weight: Unimportant		⇔		Very important
1) **Simplicity**--Simple beam configurations are prefererable to complex configurations.	0	1	2	3	4
2) **Avoidance**--Beam configurations should try to avoid surrounding healthy tissue as much as possible.	0	1	2	3	4
3) **Beam Length**--Treatment beams should traverse as little of the patient's anatomy as possible.	0	1	2	3	4
4) **Reproducibility**--Beam configurations should be easily reproducible to facilitate repeated treatments.	0	1	2	3	4
5) **Originality**--The beam configuration should reflect a creative approach to beam targeting.	0	1	2	3	4
6) ____________________	0	1	2	3	4

Figure 5-1. Sample of Criteria Survey given to each subject at beginning of participation in study.

the subject indicated that he was comfortable with the navigation mode and with the program, and that he was ready to attempt the real, "measured" trial. There was no objective criteria for termination of the practice trial, but subjects were often encouraged to take a few more minutes of practice. For this study it was important that each subject not work on the same case more than once to avoid the strongly confounding effect of learning. A subject's behavior during the second exposure to a particular case will greatly differ from his behavior during the first exposure, for he will remember where the good and bad beam prospects are and this memory will mask any effect due to navigation

Subj	Practice 1	Trial 1	Practice 2	Trial 2	Trial 3	Trial 4
1	Orb - 5	**Orb - 1**	Joy - 6	**Joy - 2**	**Orb - 3**	**Joy - 4**
2	Orb - 6	**Orb - 2**	Joy - 5	**Joy -1**	**Orb - 4**	**Joy - 3**
3	Joy - 5	**Joy - 1**	Orb - 6	**Orb - 2**	**Joy - 3**	**Orb - 4**
4	Joy - 6	**Joy - 2**	Orb - 5	**Orb - 1**	**Joy - 4**	**Orb - 3**
5	Orb - 1	**Orb - 3**	Joy - 2	**Joy - 4**	**Orb - 5**	**Joy - 6**
6	Orb - 2	**Orb - 4**	Joy - 1	**Joy - 3**	**Orb - 6**	**Joy - 5**
7	Joy - 1	**Joy - 3**	Orb - 2	**Orb - 4**	**Joy - 5**	**Orb - 6**
8	Joy - 2	**Joy - 4**	Orb - 1	**Orb - 4**	**Joy - 6**	**Orb - 5**
9	Orb - 3	**Orb - 5**	Joy - 4	**Joy - 6**	**Orb - 1**	**Joy - 2**
10	Orb - 4	**Orb - 6**	Joy - 3	**Joy - 5**	**Orb - 2**	**Joy - 1**
11	Joy - 3	**Joy - 5**	Orb - 4	**Orb - 6**	**Joy - 1**	**Orb - 2**
12	Joy - 4	**Joy - 6**	Orb - 3	**Orb - 5**	**Joy - 2**	**Orb - 1**

Table 5-1. Navigation mode and case used in each trial for each subject. Orb=Orbital mode. Joy=Joystick mode. Shaded trials were not run due to time limitations.

mode. Table 5-1 also shows the counterbalancing of navigation mode order and case order to minimize possible confounding effects.

5.3.3.1. Rules of the Game

In designing their optimal beam configuration, the subjects worked under a set of ground rules that were consistent throughout all their trials. Perhaps the most important of these was the specification that the spine had already been irradiated to its tolerance level and should be avoided at all costs, as further irradiation would cause severe, unacceptable damage. In radiation oncology terminology, the subjects were designing what is called a *boost field*. Typically for lung tumor cases such as these, direct anterior-posterior beams are used initially, because they afford the shortest path through the patient's body and the spine can tolerate a certain amount of radiation. Once the spine has reached its tolerance, however, subsequent boost fields must be carefully designed to avoid the now highly radiosensitive spine. Subjects were instructed to use their own discretion regarding the irradiation of other tissue.

The subjects were also instructed to assume the beams they set would have a cross-sectional shape that would conform exactly to the silhouette of the treatment volume as seen from the beam source. Therefore, any anatomical structure whose silhouette overlapped that of the treatment volume as seen in the beam's-eye view were understood to be impinging on the beam.

Subjects were given limits of 20 minutes for beam configuration design and 13 beams or pairs of opposed beams. The time limit was intended to be a generous allowance that would discourage overly-obsessive efforts. The beam count limit arose from a fixed selection of beam colors, but would probably be removed if the software were to be rewritten.

5.3.3.2. Software and User Interface

5.3.3.2.1. Navigation modes

beattl, the beam-targeting program used in this study, provided the user with a stereoscopic beam's-eye view of the patient's anatomy. This beam's-eye view was anchored to, but free to rotate around, an isocenter located at the centroid of the treatment volume. Such rotation of the beam's-eye view had three degrees of freedom—rotation about the gaze direction and about two orthogonal axes that lie in the plane containing the isocenter and perpendicular to the gaze direction—and enabled the subject to dynamically study the anatomy from different directions. This experiment studied two methods of rotating the beam's-eye view, or navigation modes. These were orbital mode, the best performer of the head-tracked navigation modes in the preliminary experiment, and joystick mode, the best non-head-tracked mode.

As discussed in the previous chapter (see Section 4.3.1.4), Orbital mode allowed the subject to rotate the beam's eye view about the isocenter by rotating his or her head. The translational component of head movement was ignored. The Orbital mode view was constructed by translating the anatomy so that the isocenter lay on the gaze vector at a fixed distance, known as the *source-axis distance (SAD)*, from the eyepoint. Because the anatomy was only translated onto the gaze vector and not rotated with respect to world space, as the gaze vector rotated (due to rotation of the subject's head) the subject viewed the anatomy from different orientations. For example, consider a subject starting with a straight-on anterior view of a standing patient with the patient's head above and the patient's feet below. Then a 90° rotation of the subject's head to the right yields a lateral view of the patient from the patient's right and a 90° rotation to the left yields a lateral view from the patient's left. Likewise, looking down provides a view from above the patient and looking up provides a view from below. Subjects stood while using Orbital mode to take full advantage of the kinesthetic and proprioceptive input provided by moving their feet and changing their stances in maintaining their bearings. I had originally opted to not have the subjects sit on a revolving stool because I felt such information would be too valuable to ignore. After observing the subjects, I changed my mind and decided that it would have been better for the subjects to have the stability that would have been provided by a stool.

With Joystick mode the subject used a custom-built joystick to rotate the anatomy about the isocenter (see Section 4.3.2.1). Pushing the joystick to the right or left rotated the anatomy about an axis passing through the isocenter and parallel to the up vector of the subject's view. Similarly, pushing the joystick away from or pulling the joystick toward the subject rotated the anatomy about a horizontal axis. Lastly, the end of the joystick contained a knob that could be twisted clockwise or counterclockwise to rotate the anatomy about the subject's gaze vector.

For both navigation modes subjects used the head-mounted display to view the patient anatomy. Of course, in Joystick mode, head movement had no effect on the beam's-eye view as it did in Orbital mode, but the head-mounted display was still used so that image quality would be equalized across the two modes.

5.3.3.2.2. Commands

beattl provided a handful of commands to use in designing treatment beam configurations. Because operation of a standard keyboard is virtually impossible while wearing a head-mounted display, the user interface was originally intended to be based upon a speech recognition unit receiving input from an HMD-mounted microphone. However, poor performance of the speech recognizer, probably due to interference from the magnetic tracker, in addition to speech recognizer training time seriously infringing on limited subject time, necessitated the replacement of the speech recognizer with a human operator who translated the subject's desired actions into appropriate keyboard entries. This also served to reduce subject anxiety as they were able to operate in a more natural manner, conversing with the operator rather than speaking in the fragmented phrases required by the speech recognizer.

5.3.3.2.2.1. BEAMS

First was the *set beam* command, which embedded in the patient anatomy a beam aligned with the subject's current beam's-eye view. The set beam was represented by an arrow passing through the isocenter and fixed in the anatomical model such that it rotated in conjunction with the anatomy, and each beam's arrow was uniquely colored so that arrowheads could be easily matched up with their corresponding tails. As radiation oncologists often use pairs of opposed beams that approach the tumor from opposite directions, a *set opposed pair* command was also provided to automatically plant a beam and its opposed counterpart simultaneously. Opposed pairs were represented by double-tailed arrows. Subjects were limited to thirteen single beams or opposed pairs in their beam configurations.

It may have been more useful to the subjects for the full volumes of the beams to be displayed rather than just indicating their orientations with arrows. Experimentation with several different possible full volume representations demonstrated that they were unsuitable for use in the head-mounted display. The low resolution of the displays made it impossible to discern the fine detail of the beam shapes and much important anatomical information was obscured by the beams. I decided that with the current technology, display of direction only was preferable. With improvements in display technology, however, displaying beam volumes may become more feasible, but related work concerning the display of volumetric dose distributions effectively in the context of the anatomy has indicated that finding a good representation for the beam volume may be a difficult task.

Subjects were also provided with the ability to delete beams from the current configuration. To do this the subjects first had to select the beam to be deleted by aligning the current beam's-eye view with it, at which point the selected beam would turn white. Once the beam was selected, the subject issued a *delete beam* command to remove it. *beattl* did not have the facility to change the status of a beam once it had been set in the anatomy. If the subject wanted to change a single beam into an opposed pair, or move a beam a few degrees, he had to delete the old beam and set a new one with the desired characteristics.

beattl did not permit beams to be set if they would enter the patient's body at a point that was not represented in the anatomical model, e.g. through the bottom plane of a truncated torso. Such beams would pass through unrepresented tissue, and their true effect would be difficult to assess. *beattl* displayed the international "forbidden" symbol (see Figure 5.2) to warn the subject that the current beam's-eye view was entering or exiting the body at points not represented in the anatomical model. The display of the symbol with its top half solid and its bottom half outlined meant that the current beam's-eye view was entering the body at an invalid point. A symbol with its top half outlined and its bottom half solid indicated a beam's-eye view that exited the body at an invalid point. And when the entire symbol was solid, the beam's-eye view both entered and exited the body at invalid points. Invalid entry points precluded the setting of both single beams and opposed pairs. Invalid exit points did not prohibit the setting of single beams, implicitly relying on the subject's discretion to determine whether the unrepresented tissue would cause problems when irradiated with a beam that had already passed through most of the patient's body and presumably had lost much of its strength. Since invalid exit points of the current beam's-eye view are equivalent to invalid points for beams opposed to the beam's-eye view, opposed pairs were not allowed to be set when an invalid exit point was evident.

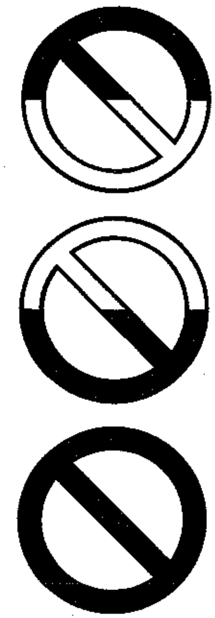


Figure 5-2. Symbols used to indicate invalid beams.

5.3.3.2.2.2. ANATOMICAL STRUCTURES (ANASTRUCTS)

Each anatomical structure in the patient model could be displayed with one of four representations: invisible, shaded surface, line segment mesh, and points. (See Figure 5-3.) Switching between representations gave the subject crude control over the transparency of a particular anastruct, which is very important when anastructs obscure each other. Each trial began with the patient's skin represented as points, lungs as meshes, and treatment volume, spinal canal, trachea, and heart represented as surfaces. Representations of individual anastructs could be separately changed at any time by the subject.

5.3.3.2.2.3. VIEWS

The subject's beam's-eye view could be magnified and minified in fixed increments by the subject. In addition, a *global view* was available, which gave the subject a view of the anatomy from 2.5 meters away, instead of the usual 1.0 meter beam's-eye view.

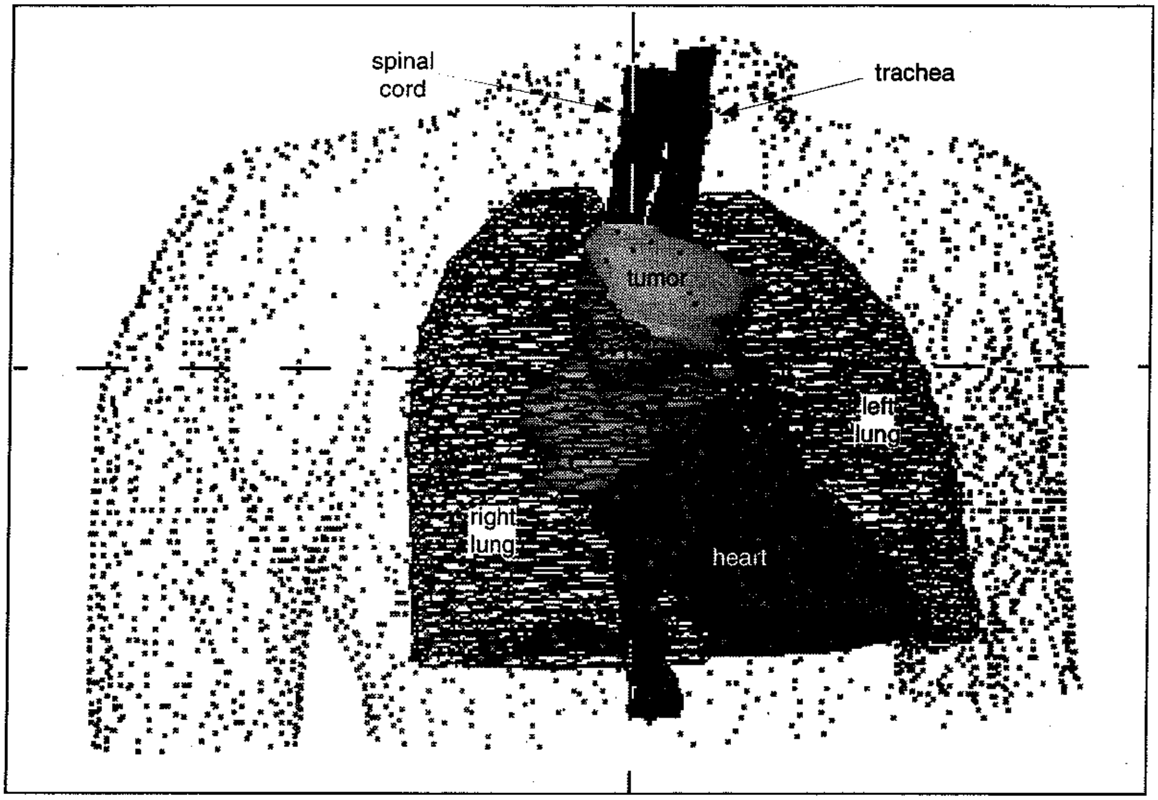


Figure 5-3. Typical anatomical model used in study. Anatomical structures have been labeled for clarity. Skin is represented as dots, lungs as meshes, and all other anastructs as surfaces. Dashed crosshair indicates position of isocenter.

5.3.3.3. HARDWARE

Pixel-Planes 5, a high-performance, scalable multicomputer for three-dimensional graphics developed at UNC—Chapel Hill (Fuchs et al. 1989), was used to generate the real time images viewed by the subjects. The head-mounted display used was a Virtual Research Flight Helmet[2]. A Polhemus Fastrak[3] magnetic tracker was used to track the head-mounted display for Orbital Mode, and an Alphacontrol Systems Model 101-Z joystick mounted in a custom-built box was used to rotate the anatomy in Joystick Mode.

5.3.4. Subjective evaluation

After all subjects had completed the beam-targeting phase, each subject then became a judge and reviewed and scored beam configurations produced by other subjects. In the course of designing experimental procedure, an important question arose as to how to present a particular beam configuration to reviewer to best promote full comprehen-

2. Flight Helmet™ is a registered trademark of Virtual Research.

3. Fastrak™ is a registered trademark of Polhemus Navigation Sciences.

sion of the relationship of the beams to the patient anatomy. The need to present the beam configurations produced with the two navigation modes in an unbiased manner precluded the use of a head-mounted display-based or joystick-based visualization tool. Presentation through static views or non-interactive video sequence was deemed to limiting. The best option was to use a tool already familiar to the subjects, the UNC Department of Radiation Oncology's virtual simulation program, *xvsim*.

xvsim was already in regular clinical use at UNC, and the subjects, all of whom were intimately involved in the treatment planning process, were accustomed to making judgments of the quality of a beam configuration from the information displayed by *xvsim*. This information included beam's-eye views from the perspective of any beam in the configuration; "skinprints", which are the outlines of the beams where they enter the body; digitally reconstructed radiographs computed from the three-dimensional dataset, which sometimes provide more useful beam's-eye views of the patient anatomy than the normal wireloop representations; and a slice-by-slice display of the three-dimensional dataset showing not only the patient anatomy at each slice, but also the intersection of each treatment beam with that slice. The last item was most useful in precisely determining what structures were hit by the beams and which were not.

The subjects used *xvsim* to examine each beam configuration. No indication was given of which steering mode was used to produce the configuration. From this presentation, each subject assigned a numerical score from 1 to 7 for each of the five criteria presented in the Criteria Survey, as well as a subjective overall score from 1 to 7.

The user interface of *xvsim* provided for exploration of the patient anatomy through rotation controlled by virtual knobs on the screen. Although all the subjects were familiar with using *xvsim* to design and evaluate beam configurations, all were not equally facile in manipulating the anatomy and using the various features of the program. For those subjects who did not feel comfortable operating *xvsim* on their own, the investigator served as "driver" and ran the program under the subject's direction.

5.4. Analysis

After all the subjects had completed beam targeting, a three-pronged approach was taken to evaluate the beam configurations and compare those produced with head-tracking information (Orbital mode) with those produced without head-tracking (Joystick mode). The first approach was to use the subjects themselves as evaluators of the beam configurations to yield subjective scores. The second was to perform objective calculations based on the intrinsic geometry of the beam configuration. The third was to analyze observational data in order to extract behavioral effects of the two navigation modes. Chapters 6 through 9 present and discuss the results of these analyses.

5.4.1. Subjective scoring by judges

After all subjects had completed the beam targeting phase of the study, each subject was brought back for a second session in which he or she was asked to score a number of beam configurations according to the criteria he or she had specified in the initial survey. For each criterion the subject assigned a numerical value from 1 to 7 indicating how well the given beam configuration fulfilled the criterion (1=poorly, 7=very well). In addition to these criteria, the subject was also asked to assign an overall score (from 1 to 7) for the beam configuration. Each subject typically graded six or seven beam configurations, all of which were for the same case. Subjects worked on the configurations in a randomized order, and they were permitted to re-examine any configurations they desired. This was to allow the subjects to more fully compare and contrast the different beam configurations for their particular case.

For each beam configuration reviewed by each subject, a *computed overall score* was calculated by summing the individual criterion scores after each score had been weighted by the appropriate factor as indicated on the Criteria Survey. This computed overall score was compared with the assigned overall score.

5.4.2. Objective measures

The objective measures computed for each beam configuration included the *number of beams* used (which corresponds roughly to the notion of simplicity of a beam configuration), and the volume of spinal cord that was irradiated (which ideally should be zero). These measures were statistically compared to find any effect due to navigation mode.

5.4.3. Observational analysis

The preliminary experiment demonstrated that replaying the actions of the subject can yield strong insights into the relative merits of the navigation modes, and so I also used that technique in this study. Not only did I replay and study the logs, but I also reviewed videotapes of the user sessions and analyzed graphical traces of the subjects' movements.

Chapter 6

User Study: Results and Discussion —Criteria Survey

At the beginning of each subject's targeting session, he or she was asked to complete a survey indicating how important each of five criteria are in the evaluation of a beam configuration. The five criteria were:

Simplicity — Simple beam configurations are preferred over more complex configurations that produce the same result. Simplicity is basically the number of beams used in the configuration.

Avoidance — Beam configurations should avoid radiosensitive critical structures as much as possible.

Beam length — Treatment beams should traverse as little of the patient's anatomy as possible.

Reproducibility — Beam configurations should be easily reproducible to facilitate repeated treatments.

Originality — How much does the beam configuration reflect an unusual, creative approach to beam targeting?

For each criterion, subjects circled a number from 0 to 4, where 0 indicated the criterion played no part in the evaluation of a beam configuration, and 4 meant the criterion was very important.

6.1 Simplicity and reproducibility

Reproducibility is related to, but not equivalent to simplicity. Reproducibility reflects the real-world implementation aspects of a beam configuration. Complex configurations will tend to be less reproducible because the greater number of beams provides greater opportunity for mistakes to be made in patient setup or machine setup. It is possible, however, for a more complex configuration to be more reproducible than a simpler one. Because radiation therapy technicians are accustomed to setting up treatments that use basic, cardinal angle beam orientations, a standard four-field plan (anterior-posterior beam, posterior-anterior beam, right lateral beam, and left lateral beam), may be less prone to error in daily setups than a two-field plan that requires unusual gantry, collimator, and table angles.

Similarly, although simple, single opposed-pair, two-field plans are usually considered very reproducible, their reproducibility can be greatly diminished if they pass through parts of the body that are difficult to immobilize and for which radiation-tissue interaction is complicated. For example, beams that pass through the armpit are nearly impossible to reproduce without some form of patient immobilization, and treatment consistency is very important in that area to control adverse skin reactions. The quality of the patient immobilization, then, determines the reproducibility of the beam. Therefore, simpler beam configurations do not guarantee higher reproducibility, nor are all complex configurations difficult to reproduce.

6.2. Importance scores

Figure 6-1 shows the distribution of survey responses for each criterion. Avoidance of healthy tissue was considered very important by almost all subjects. Reproducibility was nearly as important, on the average, followed by beam length and simplicity. Originality was rated very low, almost of no consequence in beam configuration evaluation. This is not surprising. The goal of radiation therapy is to improve the patient's condition, and therefore the medically relevant criteria of simplicity, avoidance, beam length, and reproducibility all carry significant importance. In and of itself, the originality of a beam configuration has very little value, except for where it improves the therapeutic aspects of the beam configuration. In those cases, however, what is important

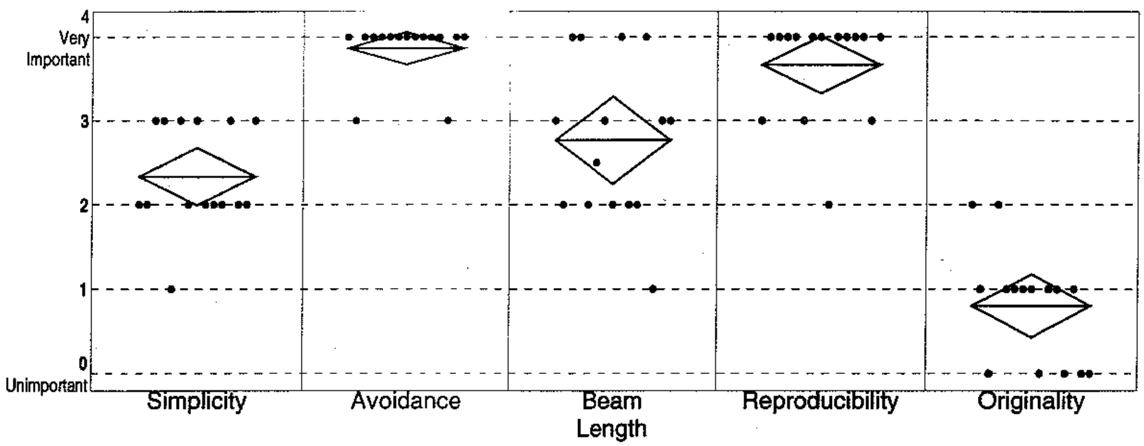


Figure 6.1. Responses to Criteria Survey. Diamonds show sample mean with 95% confidence interval for estimate of population mean.

is the improvement in avoidance or reproducibility or whatever, not the fact that nobody had ever thought of such an approach before. This is not to downplay the role of creativity in the process of designing a treatment plan. Creativity is required to discover new solutions that may turn out to be improvements over the standard approaches with their known track records. The originality criterion was included here to shed light on whether or not the head-tracked navigation provided completely new insights that would not have been seen with the non-head-tracked navigation.

Each subject's importance scores were converted into relative weights by dividing each criterion's importance score by the sum of the five importance scores. Therefore, the sum of all five relative weights for each subject was 1.0, and the relative weights for a subject who responded "very important" for all five criteria were the same as those for a subject who responded "unimportant" for all the criteria. Figure 6-2 shows the distribution of the relative weights for each criterion. The relative weights for each subject were used later when the subjects acted as judges and evaluated beam configurations. Overall scores were computed from the judges' individual criterion scores using each judge's relative weights as determined by the Criteria Survey.

6.3. Other criteria

The list above omits the most important criterion—that all the beams should hit the target volume. In this experiment it was not necessary to specifically mention this, because the isocentric exploration used in this study guaranteed that all beams would hit

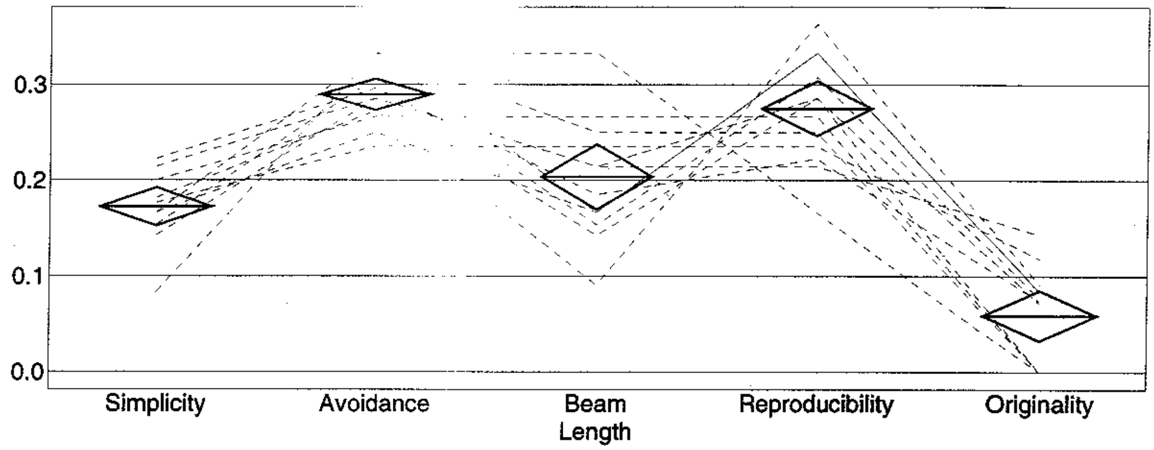


Figure 6-2. Relative weights computed from subjects' responses to Criteria Survey. Dashed lines represent collections of weights for individual subjects. Diamonds indicate means with 95% confidence intervals.

the tumor, and all beams were assumed to exactly conform to the shape of the tumor. In clinical practice, however, it is very important to ensure that the beams do hit the entire tumor.

The survey also asked the subject to describe any additional criteria, if any, he or she felt to be important and to assign an importance score to it. A handful of subjects did respond to this, bringing to light two other important aspects of beam targeting. The first was that the dose distribution produced by a beam configuration should be homogeneous and should tightly conform to the treatment volume. This criteria can only be addressed after dose calculations are performed for the beam configuration and was not included in this study, which focused only on the targeting aspect of treatment planning.

The other interesting criterion mentioned was patient condition. What was the stage and grade of the tumor? Was the treatment intended to be curative or palliative? How comfortable will the patient be in the proposed treatment position? How well are the patient's organs functioning? These types of questions affect the perspective taken in the evaluation of a beam configuration. For curative treatment, treatment planners will be more willing to try more complex configurations to get a better dose distribution. If the patient is not comfortable in the treatment position, he will not be able to tolerate lengthy treatment sessions. And a treatment planner may be willing to expose a larger lung volume in a patient with fully functioning, healthy lungs than in a patient with emphysema. A beam configuration must be designed and judged in the light of the individual patient's complete situation, for there are important constraints other than merely the spatial arrangement of the anatomical structures. These additional constraints can be important in beam targeting, but they were not included in this study because I did not learn about them until data collection was well underway. I do believe, however, that future research in radiation treatment planning should provide more complete patient histories to better simulate the design process.

Chapter 7

User Study: Results and Discussion —Targeting

The targeting session for each subject took about an hour. During the session the subject reviewed the instructions, underwent a practice session and a test trial for one navigation mode, and then a practice session and a test trial for the other navigation mode. If there was sufficient time remaining, the subject repeated the first navigation mode and then the second navigation mode. Each practice session and test trial used a different lung tumor case.

The statistical analyses reported below were performed only on the first two trials for each subject to compare initial exposures of each subject to each navigation mode. The qualitative discussions below are based on observations made during all trials.

7.1. Subject approaches and philosophies

For treatment-beam targeting the subjects were instructed to design the optimal beam configuration for the given case. Optimal configuration was defined to mean that configuration that provided the best prospect for further development into an optimal treatment plan. The subjects were told to think in the same manner as they would in real practice—to generate configurations that they would really use in the clinic. As a result, the beam configurations produced varied widely in their complexity, reflecting the range of points of view brought to the task by the subjects.

At one end of the spectrum are the minimalist planners who refused to use any more than one pair of opposed beams in their configurations. The concerns of these subjects were the more practical aspects of radiation treatment. Subject 13 acknowledged that in general, using more beams will produce more tightly conforming dose distributions, but he[1] questioned whether the gains in terms of reduced morbidity truly outweighed the disadvantages of the concomitant increase in treatment time. Fewer beams require shorter treatment time for the patient, and provide fewer opportunities for error. Subject 13 felt these considerations outweigh dosimetric benefits at this time, and never uses more than four beams in a treatment plan. Subject 5 concurred with Subject 13, and added that the particular machine settings also figure into a configuration's reproducibility. The technicians who actually deliver the treatment are not involved in the treatment planning, and as a consequence they do not have as strong an understanding of the treatment as the physician and the dosimetrist. When unusual table and gantry angles are specified in the treatment plan, this lack of understanding prevents the intuitive double-checking that the technicians can usually perform with cardinal angle treatment plans. This problem increases with the number of beams used.

At the other end of the spectrum are the subjects who used large numbers of beams in their configurations. These subjects each had their own reasons. Subject 1 assumed that ideal clinical conditions were available in which patient setup time and treatment delivery time were negligible. With these constraints removed, he felt free to use a large number of beams in an attempt to optimize the dose distribution and minimize the effect on the normal tissue. Subject 11 followed his regular practice of targeting a generous number of beams with the intention of perhaps assigning some of the beams a weight of 0 after reviewing the dosimetry. As an example, he explained that there may be a questionable beam that hits part of the liver. "But that may be acceptable. It'll depend on the dosimetry." Subject 7 simply did not perform the desired task in his trial using Orbital mode, in which he set twenty-six beams, twice the number used by the next most complex configuration. He later explained, "I was just doing that to see if I could do as many beams as I could do. I don't think I had any medical rationale for it. It was more that I was just playing. There was not any medical or any good sense explanation. It was

1.For simplicity and readability subjects will be referred to as "he," regardless of the subject's actual gender.

like Mt. Everest. Because it was there, and because I could spin around in this chair[2], I just tried to do as much as I could."

Between these two extremes are the majority of the subjects—those who try to more equitably balance the dosimetric advantages of complex configurations with the practical advantages of simple ones. In general, they appreciate that more beams can be desirable, but only if they are truly warranted. Each beam added to the configuration must have a good reason for its existence—it must add to the overall quality of the configuration. Beams that are separated by small angles should be avoided, for they are essentially redundant. These subjects typically designed beam configurations of medium complexity, using 4, 6, or 8 beams, but they were not above using a single opposed pair if the situation called for it. After a long trial that produced a beam configuration consisting of only two opposed beams, Subject 6 commented with tongue in cheek, "An opposed pair. Jeez, how humiliating. I can come up with better than that, usually."

In designing this experiment, I had not anticipated the practicality factor in treatment planning. I had thought that a radiotherapist always strove to produce the absolute best dose distribution. My intention was to compare the two navigation modes in the designing of dosimetrically optimal beam configurations, but I found I was actually getting something else by instructing the subjects to produce the beam configuration they would actually use in clinical practice. Such variability in clinical approaches surely has a confounding effect on the results of this experiment, for if a therapist intends to target only one opposed pair of beams, then extensive exploration of the patient's anatomy is probably not necessary to find the best orientation for that pair. If, however, the situation requires the therapist to go all out to design a beam configuration that will produce a tightly fitting dose distribution, then extensive exploration of the patient's anatomy and experimentation with different prospective beams will be required to achieve the goal. I speculate, therefore, that intuitive navigation with a head-mounted display will have greater effect when designing complex configurations than when designing simple two-field configurations. As with the patient history, future research might do well to establish firmer conditions to reduce variability. Subjects could be instructed to ignore practical, clinical considerations, and to concentrate on achieving the best dose distribu-

2. Because of balance problems, Subject 7 sat in a swiveling chair for Orbital mode. All other subjects stood for Orbital mode.

tion, as Subject 1 did. This would increase the sensitivity of the study to the effects of intuitive navigation, but reduce the applicability of the results to clinical practice.

The dosimetry-guided approach taken by Subject 11 illustrates another unanticipated factor—the tight linkage between the geometric problem of beam targeting and dosimetry. Evaluation of the effect of intuitive navigation is complicated by the fact that the geometric beam configurations cannot be evaluated without considering the resulting three-dimensional dose field. Calculation of the dose distribution requires the specification of numerous parameters, e.g. beam weights and beam modifiers, that are not related to the navigation mode used to target the beams. So, without dosimetry it is difficult to say that the beam configurations produced with navigation mode A are better than those produced with mode B. With dosimetry, we could more certainly say that the treatment plans produced from the beam configurations designed with navigation mode A are better than the plans produced from the configurations designed with mode B, but it would be difficult to determine whether the difference is due to the quality of the beam configurations or the other aspects of treatment planning. This study deals with the beam configurations only, without dosimetry, because I wanted to isolate the effect of the navigation mode. But in the designing of beam configurations, and in their evaluation, therapists are always thinking about the dose distribution that could result from the collection of beams. The differing abilities of different subjects to imagine dose distributions are a source of considerable variability that may mask the effect of navigation mode.

One last possibly confounding factor stems from the observation that subjects used old strategies in targeting beams that may not be appropriate with new navigation modes. As a result of training that emphasized evaluation of a treatment plan based on the examination of transverse slices through the patient, many subjects used superior (straight down) or inferior (straight up) axial views of the patient to evaluate their beams. One wonders if using this old strategy makes the best use of currently available three-dimensional navigation tools, and if perhaps a different strategy that is more appropriate for today's tools could be developed. Ideally, to determine the effect of intuitive navigation in beam targeting, subjects who had been trained to think in terms of the new tools should be used. This is not possible, so the old way of thinking must be used as a starting point for the development of new tools and techniques.

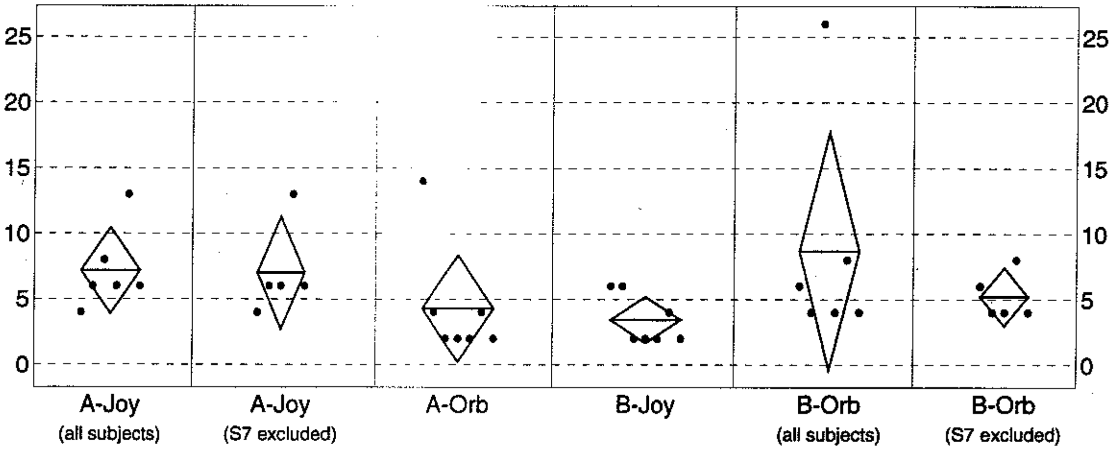


Figure 7-1. Number of beams used by sequence and navigation mode. A|B indicates sequence: A=1st trial, B=2nd trial. Joy|Orb indicates navigation mode used. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample I	vs. Sample II	*t*	*df*	Prob > |t| *p*	Est'd. Diff. of Means (I–II)	Lower 95% Conf. Limit	Upper 95% Conf. Limit
A-Joy (S7 excl.)	vs. A-Orb	1.147	10	0.278	2.714±5.274	-2.560	7.988
B-Joy	vs. B-Orb (S7 excl.)	1.628	10	0.135	-1.771±2.424	-4.195	0.653
A-Joy (all)	vs. A-Orb	1.340	11	0.207	2.881±4.732	-1.851	7.613
B-Joy	vs B-Orb (all)	1.571	11	0.145	-5.238±7.338	-12.576	2.100

Table 7-1. Comparison of number-of-beams samples using Student's t-test of null hypothesis $H_0:\mu_I-\mu_{II} = 0$. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

7.2. Number of beams

While there was great inter-subject variability in the number of beams used, no significant difference across navigation modes was found. Figure 7-1 presents the distribution of sample values broken down by navigation mode and sequence. Because Subject 7 admittedly did not actually attempt the desired task, I feel justified in excluding his data from this and all subsequent analyses. For completeness, analyses with Subject 7 included are presented in the figures and the tables, but I consider the analyses without Subject 7 to be more meaningful. Table 7-1 presents the results of independent-measures analysis performed on this data. Comparison of first trials that used Joystick mode (A-Joy) with first trials that used Orbital mode (A-Orb) yields *p*-values of 0.278 when Subject 7 is excluded and 0.207 when all subjects are included. Therefore the differences between the navigation modes are not significant at the α=0.10 level. Similarly, the differences across navigation modes are not significant for the second trials (B-Joy vs. B-Orb), with *p*-values of 0.135 and 0.145.

While statistically significant results are lacking, an interesting effect is somewhat apparent in this data. Recall that these data represent the first two trials of each subject, in which the subject first used one navigation mode and then the other mode for the second trial. The six data points in the A-Joy(all) sample represent the six subjects who used Joystick mode first. These are the same six subjects who used Orbital mode second, for which the data is presented in the sample labeled B-Orb(all). Similarly, the two samples A-Orb and B-Joy represent the same group of seven subjects, which is disjoint from the A-Joy/B-Orb subjects. Figure 7-1 suggests that the A-Orb/B-Joy subjects tended to use fewer beams than the A-Joy/B-Orb subjects. I believe this to be a result of the practicality factor discussed in Section 7.1, as the A-Orb/B-Joy group contained more of the practicality-minded subjects than the A-Joy/B-Orb group. The practicality factor will be evident in many of the statistical analyses that follow.

Figure 7-2 and Table 7-2 presents the repeated-measures analysis for the number of beams using both trials for each subject. The repeated-measures analysis reduces the

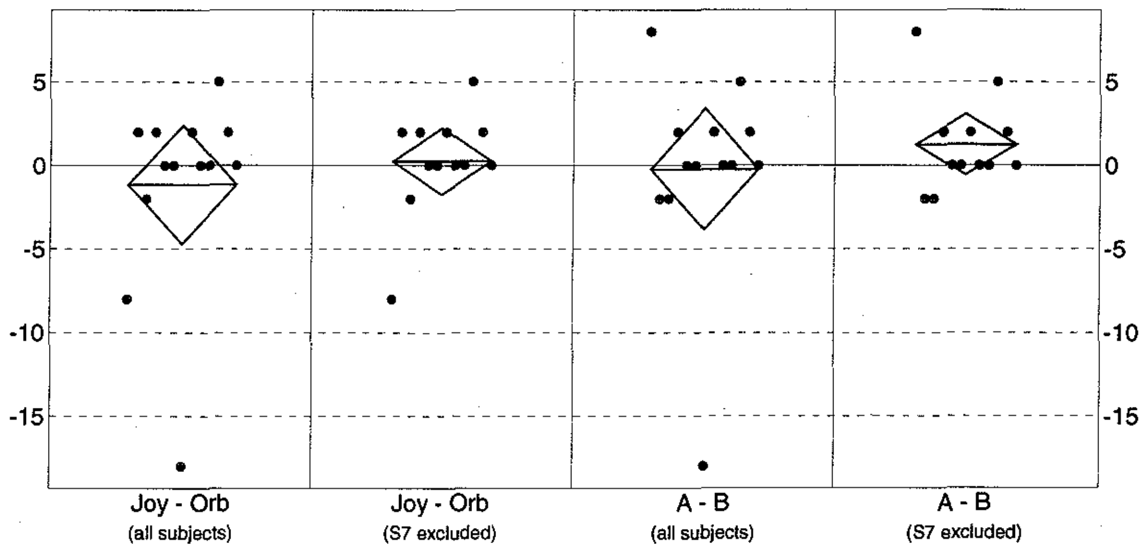


Figure 7-2. Repeated-measures analysis for number of beams used. Joy-Orb shows difference between Joystick mode performance and Orbital mode performance for each subject. A-B shows performance difference between Trial A (1st trial) and Trial B (2nd trial) for each subject. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample	t_o	*df*	Prob > |t| *p*	Est'd. Mean	Lower 95% Conf. Limit	Upper 95% Conf. Limit
Joy-Orb (S7 excl.)	0.276	11	0.276	0.250±1.993	-1.743	2.243
A-B (S7 excl.)	1.512	11	0.159	1.250±1.820	-0.570	3.070
Joy-Orb (all)	-0.707	12	0.493	-1.154±3.557	-4.711	2.403
A-B (all)	-0.139	12	0.892	-0.231±3.627	-3.858	3.396

Table 7-2. Student's t-test of null hypothesis, H_0: μ_D=0 for each number-of-beams repeated measures sample. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

error introduced by inter-subject variability by comparing performance under different treatments for each subject individually. For each subject the difference between the two trials' performances was computed. First, the number of beams used in the Orbital mode trial was subtracted from the number used in the Joystick mode trial (Joy-Orb), and then the number of beams used in the second trial was subtracted from the number used in the first trial (A-B). Figure 7-2 shows the distributions of these differences, and the analyses presented in Table 7-2 tested the null hypothesis that the population mean difference estimated by each sample is equal to zero (H_0: $\mu_D = 0$). None of the four analyses yielded a *p*-value small enough to reject the null hypothesis H_0 at the α=0.10 level, so we are unable to conclude that the population mean difference is non-zero, which means that we are unable to detect any significant effect of either navigation mode or trial sequence on the number of beams used. It is possible to see, however, that Subject 7's Orbital mode performance is an influential outlier that substantially alters the sample statistics.

7.3. Preference

At the completion of their trials and having used both navigation modes, each subject was asked which mode they preferred. There was a nearly even split in the responses, with 7 of the 13 subjects preferring Orbital mode and 6 of the 13 preferring Joystick mode. Generally speaking, the Orbital mode advocates liked the ease of control provided by the head movement. On the other hand, those who liked Joystick mode felt that turning the head and body was too unstable and preferred the precision of the joysticks. These are two sides of the same coin. Head-tracking provides an easier way of grossly moving about and exploring the anatomy, but at the same time feels less stable and less precise. There was no correlation between the order in which the subjects used the navigation modes and their preferences.

Below are presented some comments to provide insight into the subject's thoughts. First, the subjects who preferred Orbital mode:

Subject 2 exclaimed, "Oh cool!" when his Orbital mode practice run began.

Subject 3 felt that Orbital mode was easier, provided better control, and was faster.

Subject 5 explained his preference for Orbital mode, "It's easier to look around and see how much volume you're getting. You have a better perception of what you're doing." Subject 5 also felt the joysticks were too sensitive, but adjusting the sensitivity to his liking would not have changed his vote.

Subject 6 liked the better perspective provided by the head-movement. Head motion was easier for him, and he had a much better feel for the treatment machine consequences of different beams. Subject 6 claimed not to be very dextrous manually, and anticipated not doing well with joysticks. However, he did perform "not as badly as I expected."

Subject 7 said used Orbital mode after Joystick mode and commented, "This is much better, much funner. Feels better. Much more pleasurable interactively." Recall that Subject 7 then went on to set 26 beams, apparently having a lot of fun.

Subject 11 liked being able to move his head and felt that Orbital mode was more intuitive. He didn't like having to stand up, though, and suggested increasing the gain of the head-control, so that head turns of only ±90° would cover the entire azimuth range of ±180°. Previously dubious about the use of head-mounted displays in beam targeting, after the trials Subject 11 conceded that "The technology's promising. [Orbital mode] worked better than I expected."

Subject 12 felt Orbital mode was a little more intuitive, but found it very difficult to look straight down vertically in Orbital mode. Looking down vertically gives the user a superior axial view of the anatomy, allowing him to examine to some extent the transverse plane. This was an important part of Subject 12's targeting procedure, but he admitted that he had been trained to refer to the transverse plane, a practice which may not be necessary with new technology and techniques.

Comments of the Joystick mode proponents:

Although Subject 1 said, "That's fun. That's neat." after completing his Orbital mode trial, be preferred the stability of sitting in a chair with Joystick mode. "Turning around in a circle was silly. If I were walking through a museum, then it would have been okay, but for this task joysticks were more appropriate."

Subject 8 found the head-mounted display heavy and uncomfortable. He preferred Joystick mode because he was able to sit down, and maintaining his balance in Orbital mode required too much effort.

Subject 10 also complained about the discomfort of the head-mounted display. He found Joystick mode more comfortable, because it was better to not have to move the body. The joysticks required some getting used to, but "once you get the hang of it, you can move fast."

Subject 13 expressed only a slight preference for Joystick mode, which was due to it providing more control. He experienced more overshooting and misjudged movement with Orbital mode, explaining that it was "not as easy to be sensitive to how far to turn your head."

Subject 9 did not like the body movement required by Orbital mode. He moved his feet very little, which forced him to twist his body a lot to explore the patient anatomy. He preferred the stability of sitting with the joystick.

In addition to preference between navigation modes, subjects were also asked if they felt one mode enabled them to do a better job than the other. None of the subjects felt that the navigation mode had an effect on their beam targeting. They all felt they would have ended up with the same configuration regardless of which navigation mode was used. Subject 11, however, suggested that prostate cancer cases might provide a more discriminating arena than the lung cancer cases used in this study. Treatment planning for prostate cases typically involve more nearby radiosensitive structures and higher doses. Under those conditions, navigation mode might have an effect on performance.

7.4. Task completion rate

Interesting effects appear in the statistical analysis of task completion rate. Task completion time, measured in seconds, was converted to task completion rate (cases/hr.) by inverting and multiplying by 3600. The independent-measures analysis is presented in Figure 7-3 and Table 7-3, and no significant effects at the α=0.10 level can be found in these tests. There does appear, however, to be a tendency for the A-Joy/B-Orb subjects to work more quickly than the A-Orb/B-Joy subjects.

The repeated-measures analyses, shown in Figure 7-4 and Table 7-4, reveals that when using Orbital mode, subjects are able to work about 3 cases/hr. faster than when using Joystick mode. This difference is significant at the α=0.10 level, and the 95% confidence interval for the difference mean roughly spans 0.5 cases/hr. (Joystick faster than Orbital) to -6.5 cases/hr. (Orbital faster than Joystick). No significant effect appears in the trial sequence analysis (A vs. B).

Two possible explanations for this effect are suggested by the subjects' comments

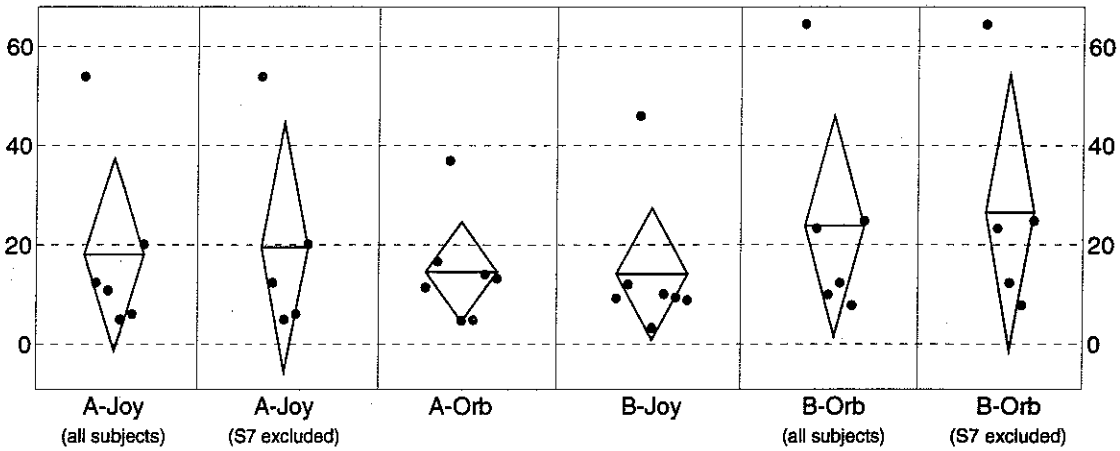


Figure 7-3. Distribution of task completion rate (cases/hr) by sequence and navigation mode. A|B indicates sequence: A=1st trial, B=2nd trial. Joy|Orb indicates navigation mode used. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample I	vs. Sample II	*t*	*df*	Prob > |t| *p*	Est'd. Diff. of Means (I–II)	Lower 95% Conf. Limit	Upper 95% Conf. Limit
A-Joy (S7 excl.)	vs. A-Orb	0.551	10	0.594	4.928±19.939	-15.011	24.867
B-Joy	vs. B-Orb (S7 excl.)	1.177	10	0.266	-12.407±23.485	-35.892	11.078
A-Joy (all)	vs. A-Orb	0.424	11	0.680	3.481±18.082	-14.601	21.563
B-Joy	vs B-Orb (all)	0.977	11	0.350	-9.648±21.746	-31.394	12.098

Table 7-3. Comparison of task-completion-rate samples using Student's t-test of null hypothesis $H_0: \mu_I - \mu_{II} = 0$. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

presented above. The first is that the mechanics of moving in Orbital mode permitted more efficient gross movement than in Joystick mode. Many subjects liked the natural movement control provided by Orbital mode. Although the joystick could actually rotate the anatomy faster than any human would find comfortable doing in Orbital mode, it provides more difficult control. When using Joystick mode to explore the anatomy, subjects were observed to use primarily rotations about the principal eye-space axes and to avoid rotating about oblique axes. This tendency was probably due to the joystick's construction, which permitted off-axis joystick deflection but required slightly greater effort to do so. This reliance on principal axis rotations required arbitrary rotations to be decomposed into a sequence of principal axis components. This decomposition resulted in more distance being covered than necessary and took more time. Overshooting was also a problem with the velocity control joysticks. Subjects tended to misjudge the timing of a rotation and would often rotate past the intended destination and have to reverse direction to correct. Orbital mode tended to make gross movement easier. Movement to a destination was more direct and corrections were more immediate. Subjects 10 and 13 are

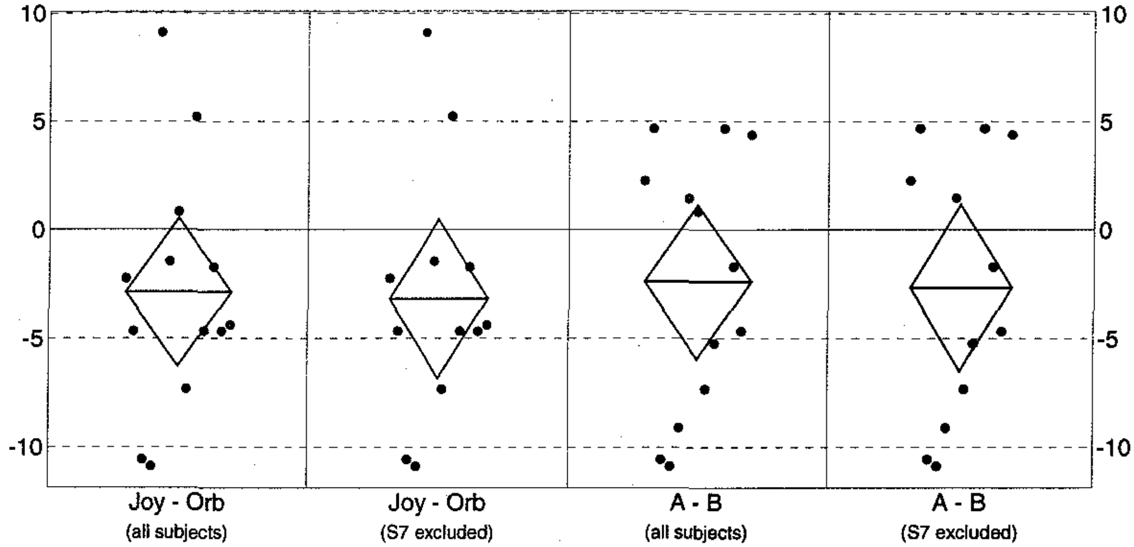


Figure 7-4. Repeated-measures analysis for task completion rate (cases/hr). Joy-Orb shows difference between Joystick mode performance and Orbital mode performance for each subject. A-B shows performance difference between Trial A (1st trial) and Trial B (2nd trial) for each subject. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample	*t*	*df*	Prob > |t| *p*	Est'd. Mean	Lower 95% Conf. Limit	Upper 95% Conf. Limit
Joy-Orb (S7 excl.)	-1.913	11	0.082	-3.190±3.670	-6.860	0.480
A-B (S7 excl.)	-1.531	11	0.154	-2.675±3.847	-6.522	1.172
Joy-Orb (all)	-1.841	12	0.091	-2.880±3.410	-6.290	0.529
A-B (all)	-1.475	12	0.166	-2.405±3.552	-5.957	1.147

Table 7-4. Student's t-test of null hypothesis, H_0: μ_D=0 for each task-completion-rate repeated measures sample. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

exceptions to this explanation. They felt that Joystick mode provided faster, more accurate movement.

Another possible explanation is that Orbital mode provided a better understanding of what the subject was doing, and thus enabled the subject to make decisions more quickly than in Joystick mode. According to some subjects, making a judgement on the amount of lung volume being irradiated by a beam is easier to do in Orbital mode. Apparently, the head movement helps to build a stronger mental model of the situation.

Further research is required to better understand these two possibilities, and I would suggest that more basic studies involving narrowly defined tasks would yield valuable insight. The speed and accuracy with which the simple task of moving to a specified destination can be completed would help us understand the effects of navigation mode on movement efficiency. Similarly, speed and accuracy of intersection volume judgements may test the better-mental-model hypothesis.

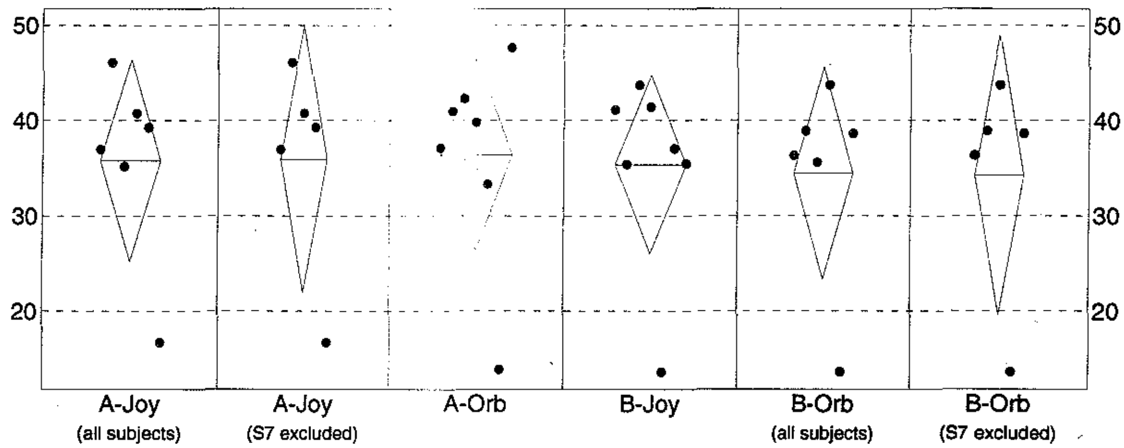


Figure 7-5. Distribution of average beam length (cm) by sequence and navigation mode. A|B indicates sequence: A=1st trial, B=2nd trial. Joy|Orb indicates navigation mode used. Analyses with Subject 7 excluded and include are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample I	vs. Sample II	*t*	*df*	Prob > |t| *p*	Est'd. Diff. of Means (I–II)	Lower 95% Conf. Limit	Upper 95% Conf. Limit
A-Joy (S7 excl.)	vs. A-Orb	0.075	10	0.942	-0.482±14.394	-14.876	13.912
B-Joy	vs. B-Orb (S7 excl.)	0.181	10	0.860	1.148±14.134	-12.986	15.282
A-Joy (all)	vs. A-Orb	0.103	11	0.920	-0.604±12.884	-13.488	12.280
B-Joy	vs B-Orb (all)	0.160	11	0.876	-0.918±12.658	-11.740	13.576

Table 7-5. Comparison of average beam length samples using Student's t-test of null hypothesis $H_0{:}\mu_I - \mu_{II} = 0$. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

7.5. Average beam length

For each beam configuration the lengths of the beams, from entry point to exit point, were computed and averaged together. Figure 7.5 and Table 7.5 show the distributions of average beam length broken down by sequence and navigation mode. Notice that most of the configurations are clustered together with average beam lengths ranging from 35 to 48 cm., but each sample shows one configuration with an average length less than 18 cm. This outlier represents a case (Case 6) that differed from the other five in that it could be treated with antero-posterior-postero-anterior (AP-PA) beams, which traverse the shortest possible distance through the patient's body. Lateral beams, which have longer traversals through the body, were not permitted by this Case 6's arrangement of tumor and spinal column, whereas the situations in the other cases precluded the use of AP-PA beams and essentially required lateral beams to be used. The repeated-measures analysis of average beam length presented in Figure 7-6 and Table 7-6 show no significant effect due to navigation mode or sequence. In fact, the two navigation modes appear to

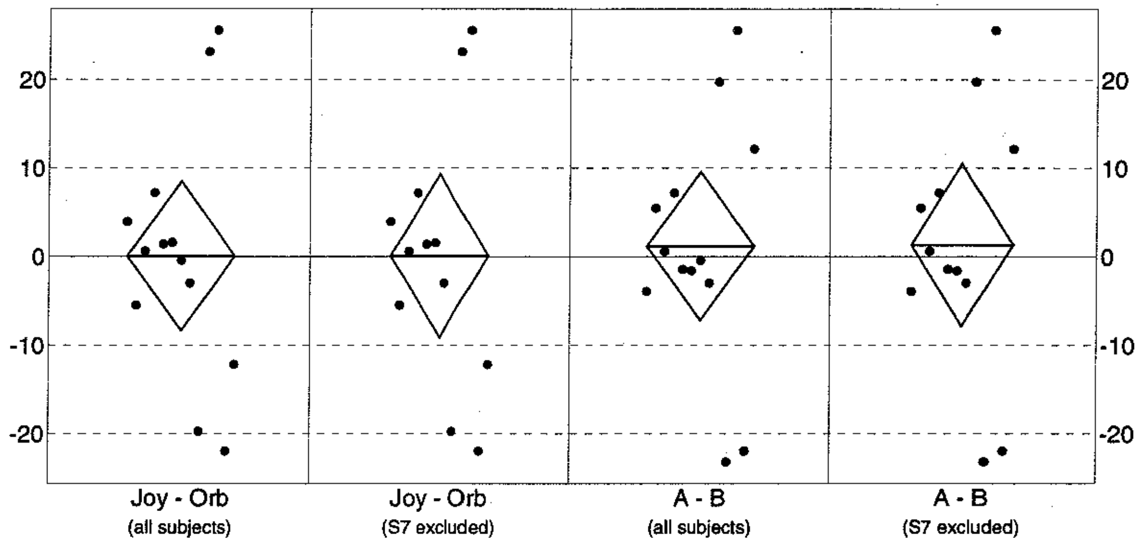


Figure 7-6. Repeated-measures analysis for average beam length (cm). Joy-Orb shows difference between Joystick mode performance and Orbital mode performance for each subject. A-B shows performance difference between Trial A (1st trial) and Trial B (2nd trial) for each subject. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample	*t*	*df*	Prob > |t| *p*	Est'd. Mean	Lower 95% Conf. Limit	Upper 95% Conf. Limit
Joy-Orb (S7 excl.)	0.025	11	0.980	0.105±9.231	-9.126	9.336
A-B (S7 excl.)	0.314	11	0.759	1.312±9.190	-7.878	10.502
Joy-Orb (all)	0.017	12	0.987	0.065±8.406	-8.341	8.472
A-B (all)	0.307	12	0.764	1.179±8.374	-7.194	9.553

Table 7-6. Student's t-test of null hypothesis, H_0: $\mu_D=0$ for each average beam length repeated measures sample. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

be equivalent, although the large confidence interval produced by Case 6's configurations makes this less certain.

7.6. Ergonomics

Many of the comments made by the subjects after their targeting trials were concerned with the discomfort experienced when wearing a head-mounted display. The effects of such ergonomic problems is not completely understood. The most common complaint was that subjects did not feel stable when standing and wearing the head-mounted display for Orbital mode. Adding a reference ground plane to the virtual environment might help ameliorate the situation, but image lag would still be present to contribute to instability. More helpful would be to let the subjects sit down in a chair so that they feel more secure. Sitting in a chair, however, would make looking at the back side of an object more difficult. This could be remedied by using a swivel chair to allow unrestricted horizontal rotation. Another solution would be to double the gain of the horizontal rotation so that a 45° rotation to the right actually rotates the view 90°. This would maintain a one-to-one mapping of head orientation to view that would enable muscle memory to still be used. It would not be possible, however, to smoothly circle around the anatomy as discontinuities would exist at the ±90° head orientations. I would also worry that the perceived rotation of the anatomy in laboratory space would be more conducive to developing motion sickness. A swiveling chair would permit smooth circumnavigation, but since the subject is no longer in direct contact with the floor, it would be more difficult for the subject to maintain bearings in lab space and muscle memory would be less reliable. This would probably be more of a problem when exploring unfamiliar objects than when viewing the familiar human anatomy.

Subjects with smaller heads, including all the female subjects and some of the males, had trouble getting the headband tight enough to stabilize the head-mounted display on their heads. They found it necessary to hold up the front of the head-mounted display with one or both hands to prevent it from tipping forward. It is unclear what effect on performance this would have, but it would more strongly affect Orbital mode, where head movement requires a stable helmet. Another negative aspect is the user's hands, which might have been used for some other useful function, are now engaged stabilizing the head-mounted display.

Joystick mode was less susceptible to ergonomic problems. In fact, problems with weight and instability of the head-mounted display may have been less severe in Joystick mode than in Orbital mode. The characteristics of joystick mapping and sensitivity are not really ergonomic concerns, but can affect performance. Subject 10 had some difficulty adjusting to the mapping from joystick deflection to anatomy rotation. Subject 4 thought he understood the mapping, but during his practice run the joybox accidently became rotated without his knowing it., and he did notice that the anatomy was not rotating the way he had intended. The joystick was too sensitive for some people, who found fine adjustments to be difficult. Other subjects were able to master the joysticks with little trouble. Future work should employ more formal training in the navigation modes to some specified level of competence.

Only two subjects reported any motion sickness symptoms as a result of participating in this study. Subject 5 felt fine until he left Sitterson Hall (where the experiment was conducted) to walk back to his office in the Radiation Oncology Department of the UNC Hospitals. He experienced mild disorientation that lasted about an hour. Subject 11 felt mildly queasy at the end of his session. He waited 5 minutes or so before returning to his office, and the symptoms disappeared soon thereafter.

Glasses fogging up, heat build-up, and excessive weight were other common complaints with the Virtual Research Flight Helmet. In fact, these complaints were the most immediate reaction to using the head-mounted display at the end of a session. I feel that these ergonomic concerns are the major obstacles standing in the way of mainstream use of the head-mounted display in the clinic. Image lag did not seem to be as big a problem as anticipated, possibly because the interaction provided by Orbital mode, which did not attempt to immerse the user in a virtual world, was artificial enough to reduce the negative impact of image lag on performance. From my observations, I believe clinicians are more averse to donning uncomfortable, tiresome head units on a regular basis than they are to the image lag.

7.7. Movement and coverage

During each targeting trial a log file was created which recorded the parameters of the current beam's-eye view at each image update and any commands executed by the subject (e.g. set beam). To study the behavior of the subjects, I played back these logs through a program that allowed me to switch back and forth between seeing what the subject saw and a god's-eye view of the subject and the anatomy model. The observations from these log replays are discussed below.

7.7.1. Polar projection trace

For this discussion it will be useful to use polar projections of the subject's movement as two-dimensional illustrations. To understand these diagrams, recall that the subject never changes the source-axis distance. Therefore, the beam source, which also coincides with the eyepoint used to generate the beam's-eye view, always remains a constant distance away from the isocenter, and the locus of all beam's-eye viewpoints visited by the subject will lie on the surface of a sphere whose radius is equal to the source-axis distance (the *solution sphere*). The polar projection trace is constructed by projecting each

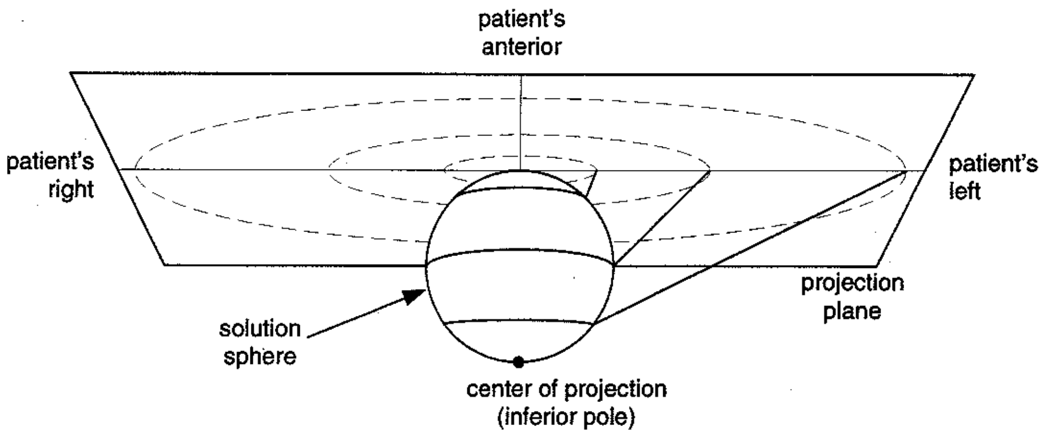


Figure 7-7. Construction of polar projection trace from solution sphere. Illustration shows projections of 45° inferior latitude circle, equator (transverse plane), and 45° superior latitude circle of solution sphere projecting onto concentric circles, the smallest of which corresponds to the 45° superior latitude.

beam source point in the subject's path from a center of projection located at the inferior pole of the solution sphere onto a horizontal projection plane tangent to the solution sphere at the superior pole. Figure 7-7 illustrates the projection. Keep in mind that the polar projection trace is a view looking up from the patient's feet with the patient's front above and the patient's back below.

Every point on the solution sphere except the inferior and superior poles can be uniquely specified by a pair of azimuth and elevation angles. Elevation values can vary from -90° for a straight inferior view (looking up from the patient's feet), through 0° for views on the transverse plane, to 90° for straight superior views (looking down from the patient's head). Azimuth values range from 0° for a posterior view through 90° for a right lateral view, through 180° view for an anterior view, through 270° for left lateral, up to 360° for the posterior view.

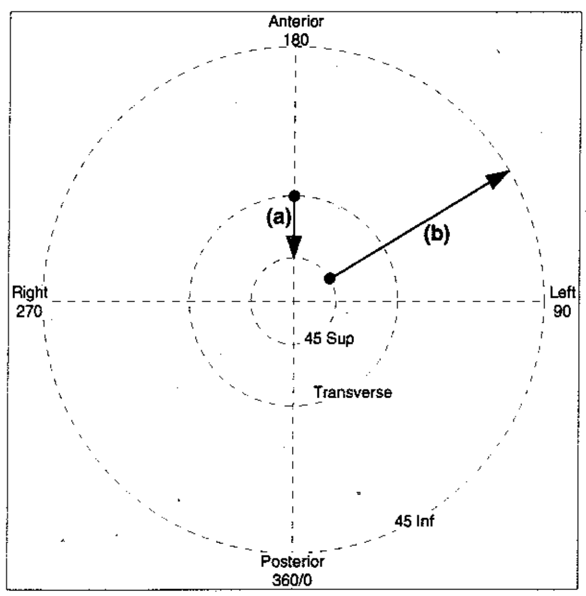


Figure 7-8. Polar trace of elevation changes.

Figure 7-8 shows the polar projection traces that would result from two simple movements. The projection is oriented such that the patient is facing up, and the cardinal azimuth angles are labeled for reference. Latitude circles of the solution sphere project into circles, and the projections of the 45° superior and inferior circles and the transverse plane (equator) are shown for reference. Segment (a) shows movement starting from a direct anterior view of the patient and moving up until the beam's-eye view is anterior 45° superior. Segment (b) shows movement from 45° superior to 45° inferior at a right posterior direction. This figure illustrates that radially-oriented trace segments indicate changes in elevation only and no change in azimuth. Both navigation modes permitted elevation changes that carried the beam's-eye view over the poles of the solution sphere. Figure 7-9 shows the polar traces of two such maneuvers. Segment (a) shows a crossing of the superior pole. Segment (b) shows the beginning of an inferior pole crossing. Since the inferior pole is also the center of projection, it cannot be represented on the polar trace. All that can be seen is the trace moving outward to the edge of the diagram and then reentering from the opposite side (Segment (c)). This inability to represent traces that pass through the inferior polar cap should be only mildly inconvenient, because subjects tended to avoid the polar regions where their beam's-eye views represented invalid beams that could not be targeted. In the few instances when subjects did venture into the polar regions, they were more likely to do so at the superior pole, which is completely represented in the polar trace, than at the inferior pole.

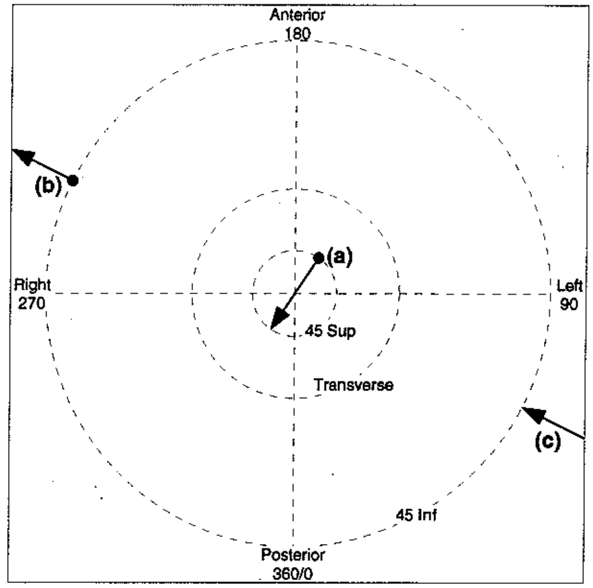


Figure 7-9. Movement across the superior (a) and the inferior (b-c) poles.

Sideways movement varies between the two navigation modes. In Joystick mode deflecting the joystick in any combination of left-right and forward-back movement without any twisting of the joystick's cap rotates the model about some axis that passes through the isocenter and is perpendicular to the gaze direction. Because of this, movement of the beam's-eye view resulting from a given joystick deflection will always be along a great circle of the solution sphere. Figure 7-10 shows three such joystick mode traces resulting from pushing the joystick directly to the left. Trace (a) begins with a direct anterior view of the patient. When the joystick is deflected to the left the model appears to rotate to the left about the vertical axis. The beam source remains on the transverse plane as it moves to a left lateral view, to a posterior view, then to a right lateral view and then back to the anterior view. Trace (b) starts with an anterior 45° superior view, and this time as the model rotates to the left the great circle path of the beam source crosses the transverse plane on the patient's left, passes through a posterior 45° inferior view, crosses the transverse plane again at the patient's right and returns to the anterior 45° superior view. Note that like trace (a), trace (b) is also circular, although it is not centered on the superior pole since its elevation is not constant. Apparently, polar projections of great circles are also circles. Trace (c) starts with an anterior 80° superior view, and its circular path extends beyond the bounds of this diagram. We see then that the trace of any Joystick movement should appear as a connected sequence of circular line segments, where the connections of different segments represents a change in the direction of joystick deflection.

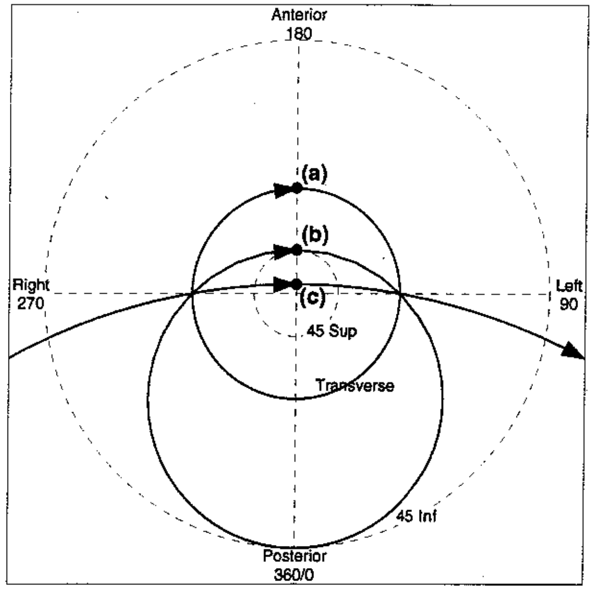


Figure 7-10. Three examples of lateral movement with Joystick mode.

In Orbital mode, vertical movement traces are identical to those in Joystick mode. When a subject tilts his head up or down the resulting trace will be a radial line segment. Sideways movement, however, differs in that it is effected by turning the body while keeping head tilt constant. This produces a beam source path that approximates a latitude circle of the solution sphere, rather than the great circle of Joystick mode. This is

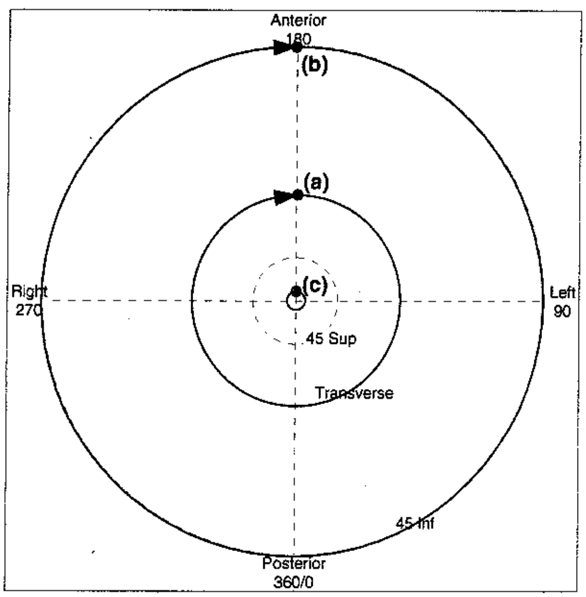


Figure 7-11. Three examples of lateral movement with Orbital mode.

shown in Figure 7-11, which illustrates sideways movement circuits at three different elevations. Therefore, circular segments in an Orbital mode polar trace that are centered on the superior pole will indicate horizontal sweeps in which head tilt was held constant. We will not expect to see circular segments not centered at the pole, as we do in Joystick mode, because there is no preference in Orbital mode to move along great circles.

As an example let us examine the short Orbital mode trial represented by Figure 7-12, which is simply a time series plot of the azimuth and elevation of the subject's view-point relative to the isocenter. The course of the 30 second trial has been broken up into labeled segments. Segment A is the beginning of the trial in which the subject is viewing the patient anteriorly and has not yet moved. In Segment B the subject turns his head about 100° to the left, which moves his view of the patient to the left lateral. He then tilts his head up in Segment C, which moves his view off the transverse plane to about a 25° inferior view. Staying inferior (keeping his head tilted up), the subject turns his head (Segment D) to the right until his view is right inferior. Then he tilts his head down to move to a right superior view in Segment E. Finally, keeping his head tilted down to maintain a superior orientation, the subject turns his head to the left (Segment F) and stops the trial with a left-superior view (Segment G).

Figure 7-13 is a polar projection of the path of the subject's beam's-eye view during the trial. In Figure 7-13 the head-turning movements of Segments B, D, and F, which appear as circular arcs centered at the superior pole, are easily distinguished from the head-tilting movements of Segments C and E, which appear as radial line segments. The polar trace also concisely displays the subject's coverage of the solution sphere. For example, we know from Figure 7-13 that the subject never examined the posterior approaches to the tumor.

The polar traces of the different targeting trials clearly reflect the different strategies and techniques used by the different subjects. Most of the interesting characteristics of the traces are effects of inter-subject differences and do not reflect any influence of navigation mode. They are nevertheless discussed here because they provide insight that is valuable in designing user interfaces for this task.

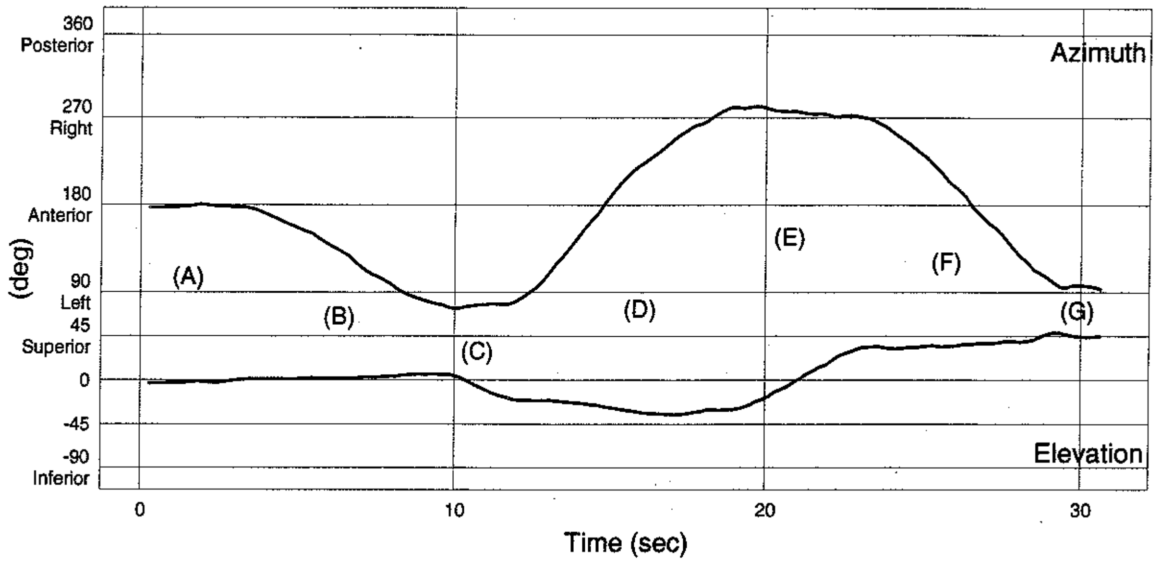


Figure 7-12. Example time series plot of azimuth and elevation of subject's beam's-eye view during exploration of patient's anatomy. See text for explanation.

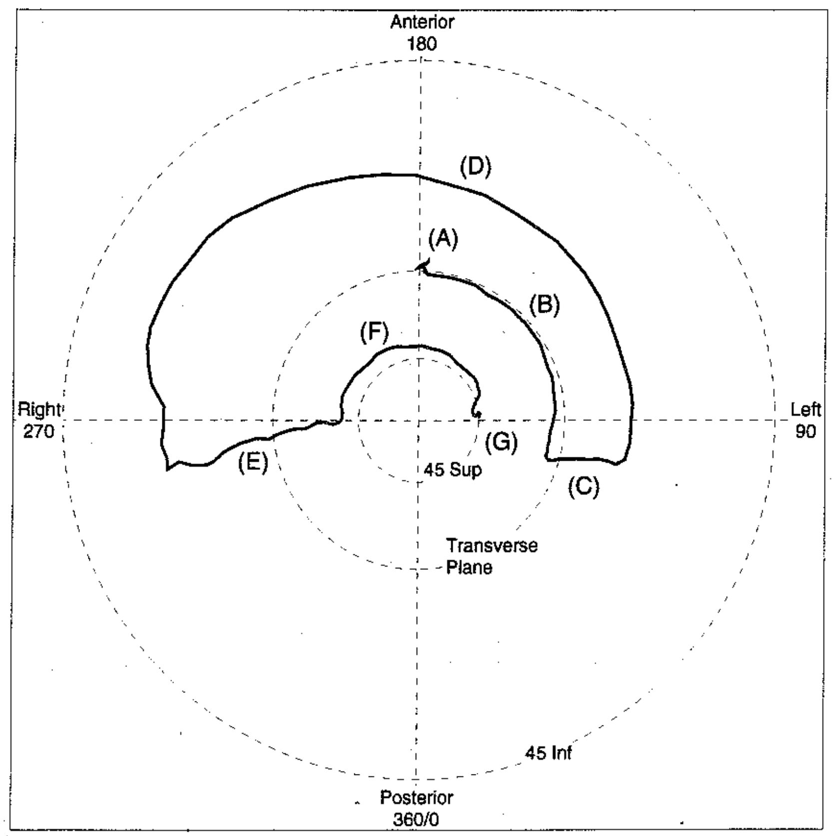


Figure 7-13. Example polar trace of same trial represented in Figure 7-12. See text for description.

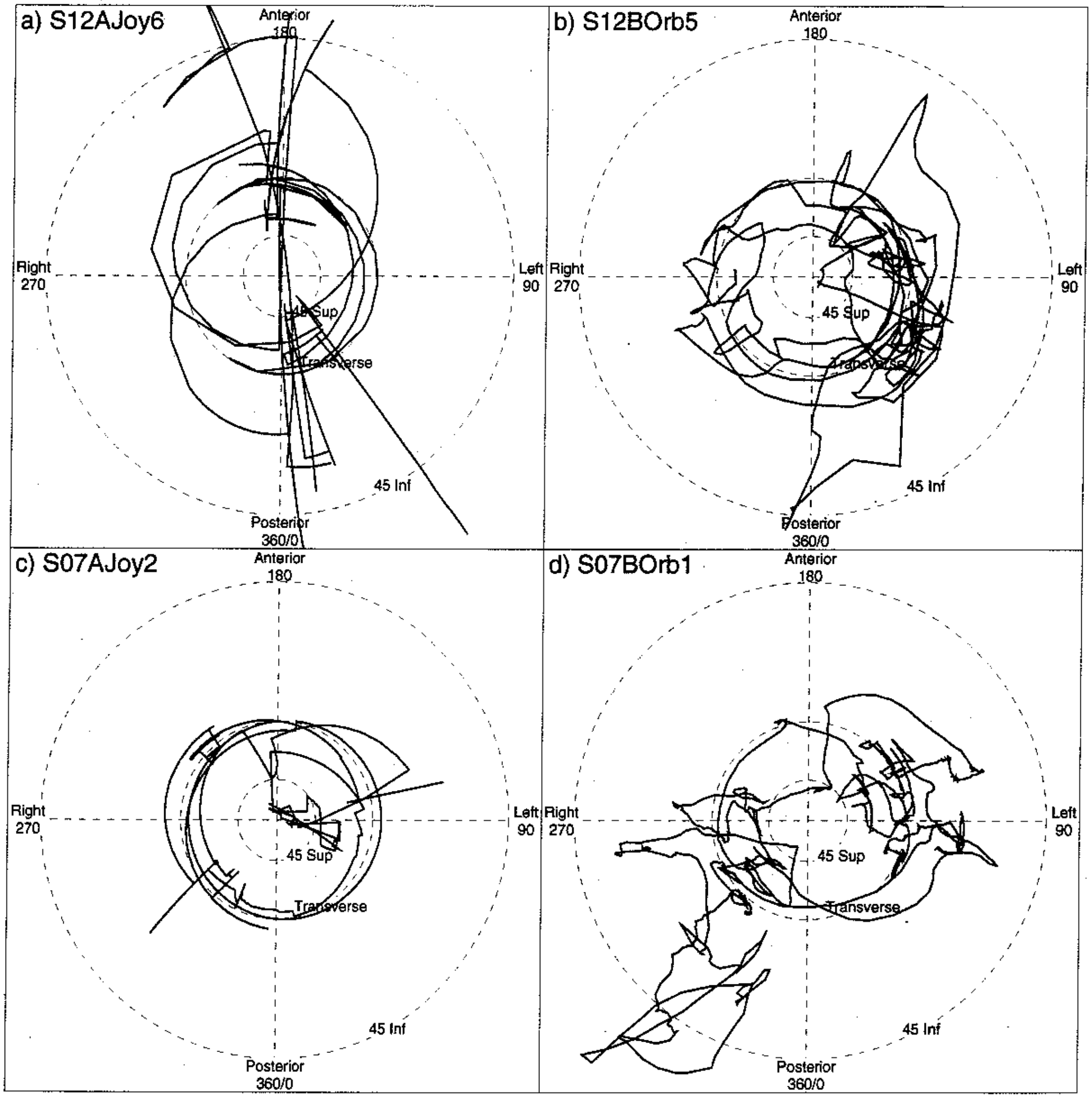


Figure 7-14. Comparison of patterns of movement between Joystick mode (parts a) and c)) and Orbital modes (parts b) and d)). Parts a) and b) taken from two of Subject 12's trials, parts c) and d) from Subject 7.

7.7.2. Freedom of movement

Figure 7-14 illustrates the one striking difference between Joystick mode navigation and Orbital mode navigation. Traces for Subjects 7 and 12[3] using both modes show the freer exploration of Orbital mode compared to the constrained exploration of Joystick mode. The tendency to use only principal axes rotations in Joystick mode is clearly shown

3. The labels used with data or illustrations for individual trials is formatted as Snnsmmmc, where nn=subject number, s=sequence (A=subject's first trial, B=second trial ...), mmm=navigation mode, and c=case number.

in Figures 7-14a and 7-14c. The pattern of movement for Joystick mode can be characterized as large, sweeping, rotations followed by 90° direction changes that were the result of changing the rotation from left-right rotation to up-down, or vice-versa. Log replays showed that most direction changes that do not appear to be 90° were the result of twisting the knob of the joystick, which rotated the anatomy about the gaze direction (perhaps to align the anatomy vertical with the view vertical—see Section 7.9), before continuing with a left-right rotation or an up-down rotation. Only very rarely did a subject actually deflect the joystick obliquely. Segmented arcs visible in the traces are the result of inadequate sampling of fast rotations. The Orbital mode trace shows direction changes that are more unconstrained than in Joystick mode, and a larger presence of small scale movement. Some of the small scale movement can be attributed to intentional fine adjustments of the beam's-eye view, but a significant proportion of them reflect inadvertent movement resulting from the imprecision of head and neck control. This imprecision was a major reason for preferring Joystick mode, for the joystick with its null position enabled the subject to maintain a constant view indefinitely without the fatigue and jittering produced by the head-mounted display's weight and moment of inertia.

An interesting effect on direction of movement due to navigation mode becomes discernible in Figures 7-15 through 7-18. These figures summarize the movement of the beam source relative to the subject's view for each image update. Figure 7-15 presents the analysis for Joystick mode. The scattergram in part (a) reveals a very strong tendency to move horizontally or vertically along the principal axes. There are some off-axis points, but the histogram in part (b) shows that they do not represent any significant level of activity. As discussed in Chapter 4, this axial tendency is most likely due to the joystick mechanism, which had a preference for movement along the axes. A curious feature of Figure 7-15a is that the down-left quadrant is essentially unpopulated, while the other three quadrants show greater activity. I have no explanation for this.

Figure 7-16 presents the same analysis for Orbital mode, and comparison with Figure 7-15 above reveals similarity and difference. The difference is that there is much more off axis activity with Orbital mode. This is not surprising since Orbital mode does not impose a principal axis preference on its users. In light of this it is somewhat surprising, however, that the Orbital mode plots also show greater activity in the horizontal and vertical directions. I believe this characteristic is a result of a search strategy commonly

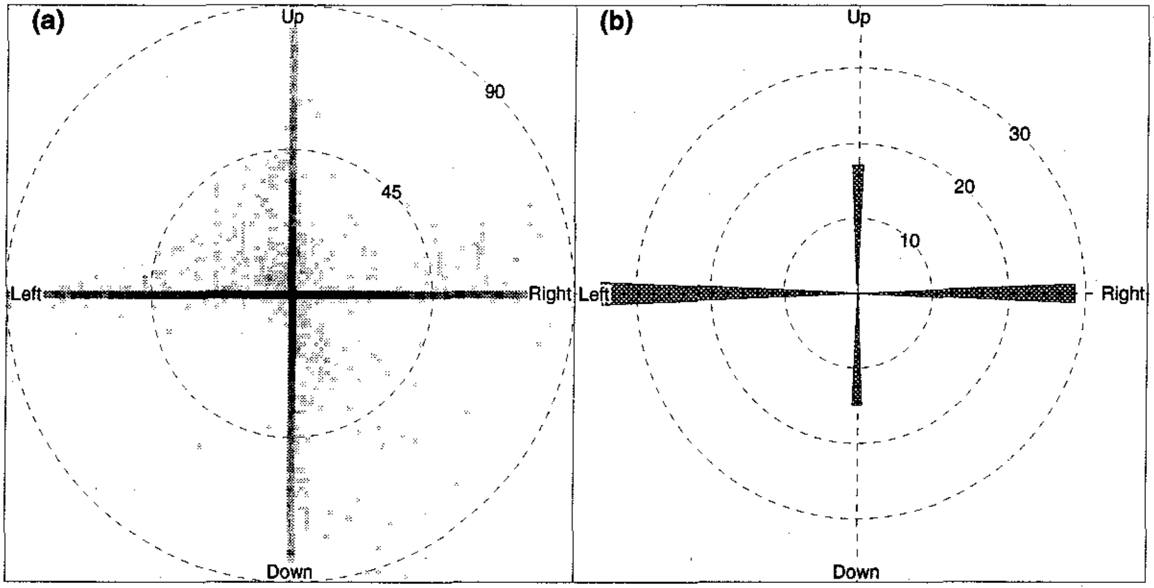


Figure 7-15. Analysis of movement direction relative to subject's view for Joystick mode trials. Part (a) is scattergram in which each point represents a movement velocity vector for the beam source. Darkness of a given point is positively correlated with frequency of use of that velocity vector. Dashed circles represent speed contours in °/sec. Part (b) is a direction histogram showing frequency of movement for 5° increments. Dashed circles represent percentage contours.

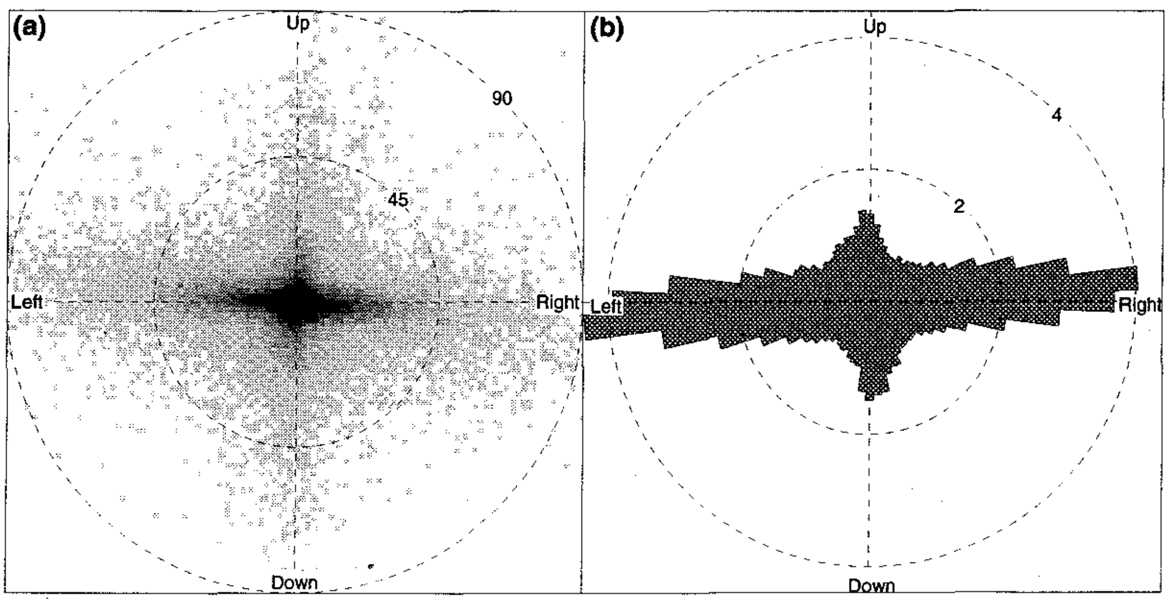


Figure 7-16. Analysis of movement direction relative to subject's view for Orbital mode trials. Part (a) is scattergram in which each point represents a movement velocity vector for the beam source. Darkness of a given point is positively correlated with frequency of use of that velocity vector. Dashed circles represent speed contours in °/sec. Part (b) is a direction histogram showing frequency of movement for 5° increments. Dashed circles represent percentage contours.

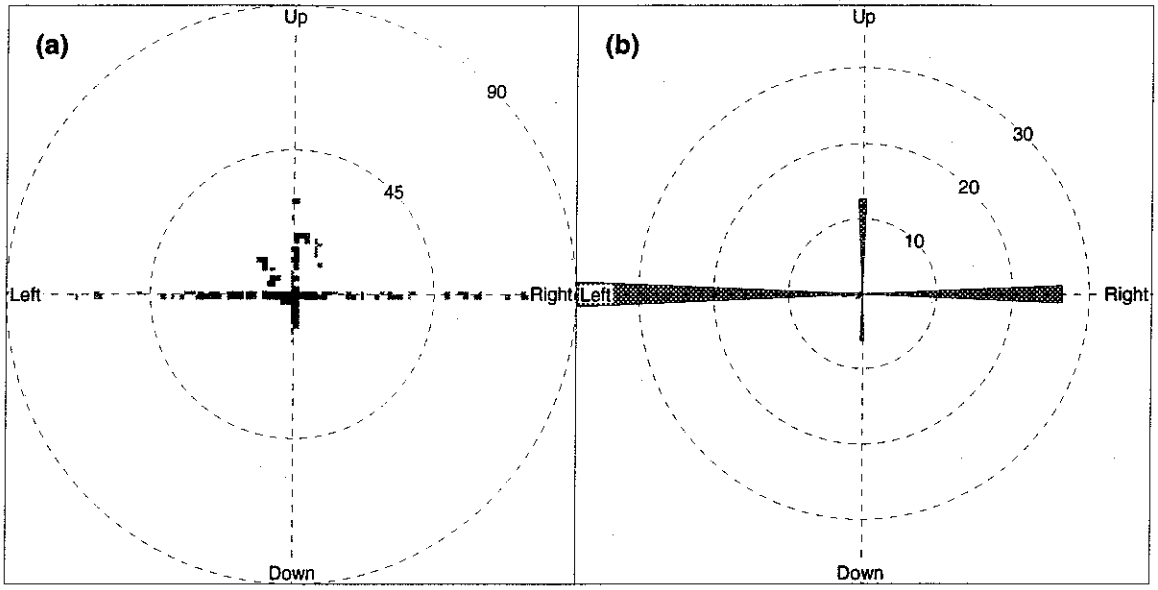


Figure 7-17. Analysis of movement direction relative to subject's view prior to beam deletions for Joystick mode trials. Part (a) is scattergram in which each point represents a movement velocity vector for the beam source. Darkness of a given point is positively correlated with frequency of use of that velocity vector. Dashed circles represent speed contours in °/sec. Part (b) is a direction histogram showing frequency of movement for 5° increments. Dashed circles represent percentage contours.

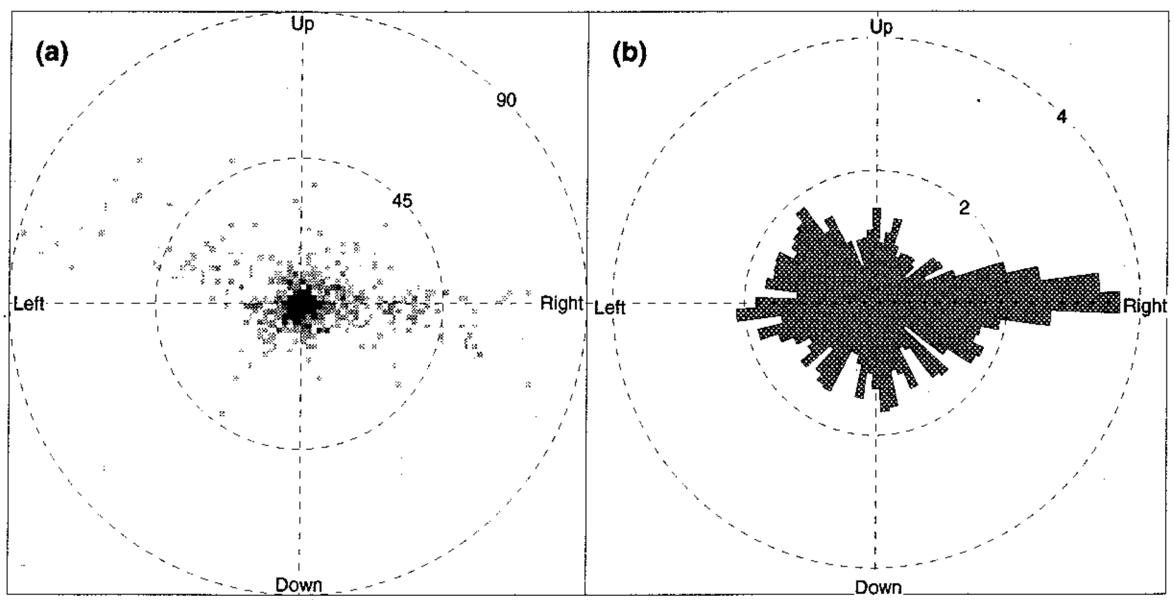


Figure 7-18. Analysis of movement direction relative to subject's view prior to beam deletions for Orbital mode trials. Part (a) is scattergram in which each point represents a movement velocity vector for the beam source. Darkness of a given point is positively correlated with frequency of use of that velocity vector. Dashed circles represent speed contours in °/sec. Part (b) is a direction histogram showing frequency of movement for 5° increments. Dashed circles represent percentage contours.

used by subjects in Orbital mode, which consisted of scanning horizontally to find good azimuth prospect and then scanning vertically to find the best elevations.

On the other hand, if we examine directed movement to a particular destination instead of free exploration we get a different picture. Almost the only times I could be sure that the subject was moving to a specific target were when the subject deleted a beam he had set previously, because to delete a beam the subject had to align his view with the beam to select it for deletion. I therefore identified all beam deletion events in the log files and extracted from them the 5-second segments preceding the beam deletions. Figures 7-17 and 7-18 show the analyses for this collection of 5 second segments. Joystick mode, presented in Figure 7-17, shows the same basic structure for this directed movement as for the overall movement, i.e. a very strong tendency to move along the principal axes. This means that movement to a target would be decomposed into a series of principal axis movements. Orbital mode, presented in Figure 7-18, shows a more omnidirectional character. Unlike the overall Orbital movement presented in Figure 7-16, this analysis of directed Orbital movement shows very little preference for axial movement, suggesting that when subjects had a particular target in mind they moved directly toward it.

Kilpatrick (1976) studied the control of a virtual end-effector with a force-display manipulator, and conducted a study which required subjects to pick up virtual blocks from a virtual tabletop using the virtual end-effector. He observed that the subjects decomposed the three-dimensional task of placing the center of the end-effector's tongs close to the center of the target block into a sequence of two-dimensional fits and one-dimensional fits. The fact that the two-dimensional fits (e.g. sliding the end-effector along the tabletop to surround the block) were not themselves decomposed into one-dimensional fits suggests that the principal axes tendencies exhibited in my subjects' joystick use were artifacts of the mechanical behavior of the joysticks. I would expect that an unbiased joystick would produce a more uniform direction histogram. It also appeared, however, that the joystick's "encouragement" of its users to employ only principal axes rotations was not found to be objectionable by the users, and may in fact have provided them with more precise control.

It is appropriate to note here the results of Jagacinski and Monk (1985), who studied two-dimensional target-capture movement controlled by either a position-control joystick or a helmet-mounted sight. In both cases Jagacinski and Monk found that diagonal movement was slower than horizontal or vertical movement, although they were unable to explain the phenomenon. In this experiment, movement to alignment with a particular beam can be thought of as two-dimensional target capture, but the scattergram of Figure 7-18 shows no cardinal direction preference other than greater speeds used when moving to the right. This difference may be due to the subjects in this study not being under the same time pressure as the subjects in Jagacinski and Monk's experiment. The parallels that can be drawn between their experiment and this research are limited because Jagacinski and Monk used a position-control joystick rather than the velocity-control joystick used here.

7.7.3. Trial duration

Figure 7-19 presents the polar traces for the shortest trial of all subjects on all cases, in which Subject 3 took 49 seconds to target the beams for Case 3, and the longest trial, in which Subject 6 took 1099 seconds to work on Case 3. Subject 3 started anteriorly and essentially performed a single sweep most of the way around the patient with a couple of excursions into slightly superior and inferior views. On the other hand, Subject 6 repeatedly swept back and forth between areas of interest located roughly at 280° and 100° azimuth and extensively studied a wide range of elevations. It is interesting to note that Subject 3's beam configuration contained three opposed pairs, and while Subject 6 had as many as five beams set a one time, most of them were deleted and he finally ended up with just one pair.

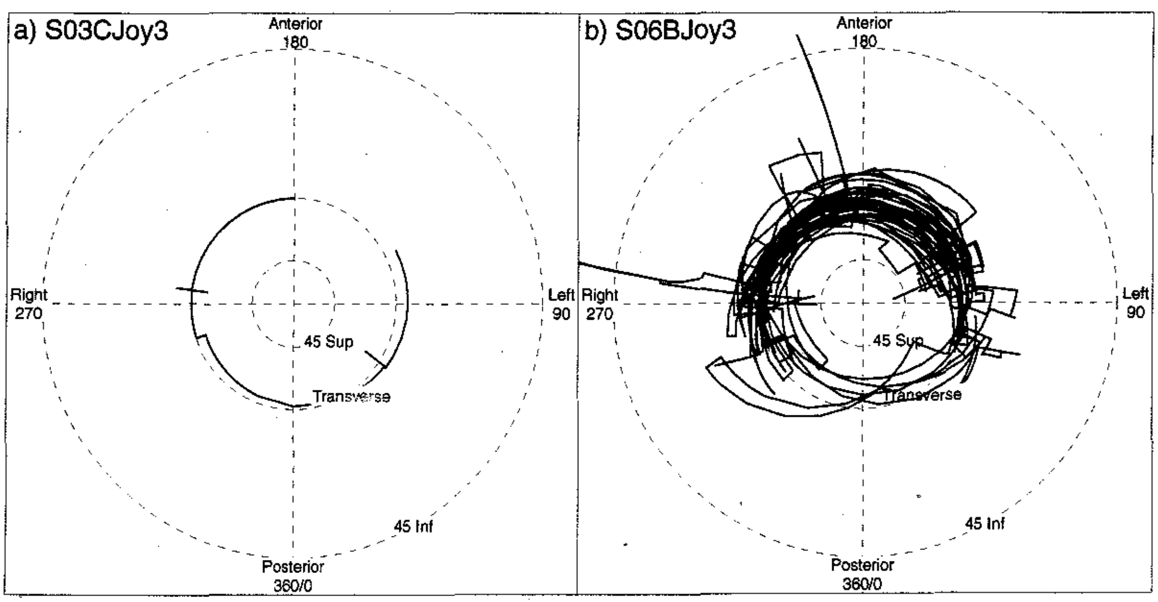


Figure 7-19. Comparison of shortest targeting session and long targeting session, both for Case 3.

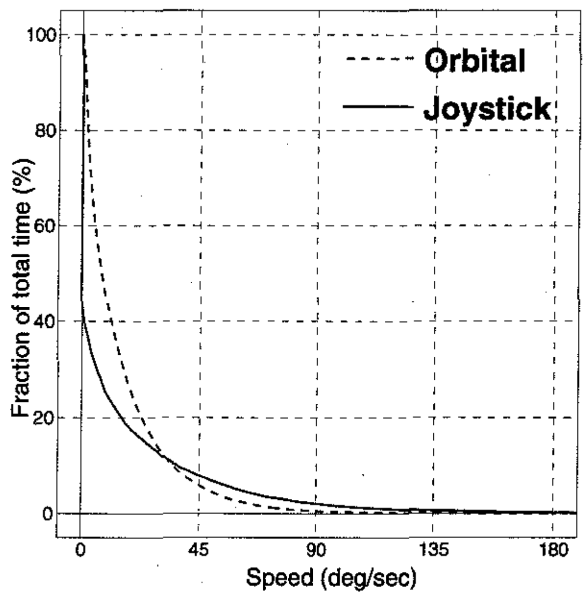


Figure 7-20. Speed time histograms showing fraction of time spent moving at a given speed or faster.

A fundamental difference between Joystick mode and Orbital mode is that the joystick has a stable null position to which it will return in the absence of any manual deflection. This null position produces no movement and requires no exertion on the part of the subject. Orbital mode, on the other hand has no stable null position, and requires a conscious and possibly demanding effort from the subject to hold a constant position. We see evidence for this difference in the speed-time histograms for the two modes, which tells us what fraction of the total time was spent moving at a given speed or faster. Figure 7-20 presents the two histograms corresponding to the data previously presented in Figures 7-15 and 7-16. The Joystick mode plot shows that 60% of the subjects' time was spent not moving. Being much more difficult in Orbital mode, staying perfectly still was almost non-existent, although a great deal of time was spent moving slowly as evidenced by a median speed of 7.5°/sec. The general shapes of the two plots are similar, especially at the high ends. In both modes, almost no time was spent moving 180°/sec. or faster, but below that, Joystick mode users tended to spend more time at speeds greater than 30°/sec. than did Orbital users. I suspect that there is a phenomenon here that parallels Ware and Slipp's finding of subjects maintaining a constant ratio of flying speed to diameter of the tunnel through which they were flying. (Ware and Slipp 1991) Perhaps there is some optimal maximum rotation rate at which people feel comfortable moving, and which depends upon the characteristics of the scene. I would suspect that just as Ware and Slipp's subjects varied their ratios depending on the flying control being used, the optimal rotation rate will vary according to the characteristics of the rotation mode, and according to system performance parameters such as lag and update rate. Informal observations of users in our lab have shown that most users will slow down their movements in the presence of large lag or low update rate.

7.7.4. Off-transverse movement

Figure 7-21 shows four targeting trials involving Case 4. In the upper two polar

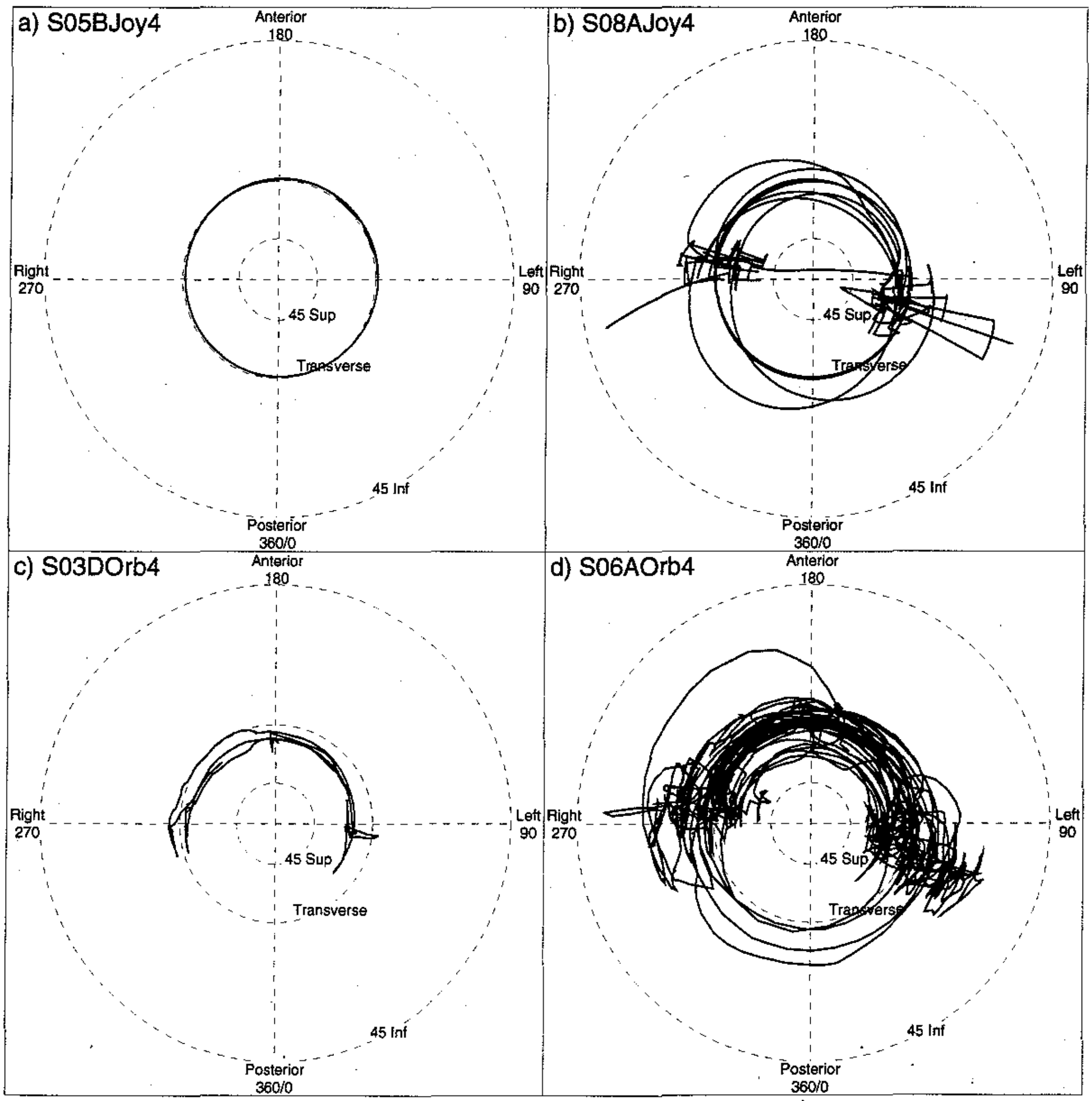


Figure 7-21. Comparison of movement away from transverse plane for four subjects targeting beams for case 4. Subjects 5 and 8 (parts a) and b)) used Joystick mode, and Subjects 3 and 6 (parts c) and d)) used Orbital mode.

traces Subjects 5 and 8 used Joystick mode, and the lower two traces show Subjects 3 and 6 using Orbital mode. This figure is a study of movement away from the transverse plane and reflects large inter-subject variability that appears to be independent of navigation mode. Subject 5 never strayed from the plane and used only left-right rotation to explore the patient. Subject 3 also did very little exploration in the superior and inferior directions. What little variation there is in elevation is most likely due to imprecise head control. Subjects 8 and 6, however, show good coverage of the vertical dimension, especially around azimuths of 80° and 260° where the most viable beam candidates could be found. Replaying the trials showed that Subjects 6 and 8 conscientiously explored the full range of allowable beams, moving superiorly and inferiorly until they crossed into the illegal beam region. Just as no correlation with navigation mode is apparent, there appears to be no connection between elevation coverage and the complexity of the resulting beam configuration, as Subjects 5 and 6 favored simpler configurations and each used two beams and Subjects 3 and 8 each used 6 beams.

The lack of vertical exploration by some subjects is confusing. It is understandable that there are certain azimuth angles which provide good beam prospects and other which are simply not worth considering, given that the subjects had to try to avoid the roughly vertical spinal cord. For those good azimuth prospects, however, changing the elevation of the beam can have a significant effect on the beam's efficacy, as the amount of irradiated lung volume can change dramatically. For subjects to ignore this possibility implies that training and/or experience has developed in them an approach to treatment planning that relies primarily on coplanar, transverse beams. Whether this philosophy is based on practical or dosimetric considerations, it confounds the issue of whether intuitive navigation will yield better treatment plans. Searching only the transverse plane is not a demanding enough navigational problem to evoke an effect from different navigation modes. The question again arises as to whether new techniques can be fully effective without new strategies designed to make the best use of those techniques.

Although not shown in Figure 7-21, the polar traces for several Orbital mode trials showed subjects exploring only the superior hemisphere of the solution sphere. The entire trace lay inside the Transverse plane reference circle, indicating that the subject never tilted his head back to view the anatomy from below. This was probably an inadvertent result of the head-mounted display being somewhat front-heavy, which would tend to pull the subjects' heads down, perhaps imperceptibly by the subjects, and turning the head would left and right would sweep along a superior latitude circle on the solution sphere. More effort would be required to look up, and apparently not all subjects were willing to exert themselves. I believe this is supporting evidence for the assertion that intuitive interfaces will encourage more complete examination of the solution space. Difficult navigation modes will discourage the user from fully exploring all possibilities, and if the effect is subtle enough, the subjects may not even be aware of what is happening.

7.7.5. Polar viewing

Another strategy used by subjects was to study the situation by viewing from either the superior pole or the inferior pole. This practice was as close as the subjects could get to viewing transverse CT slices of the patient, which is a common technique in treatment planning. When outlines of the areas of intersection between the treatment beams and the particular slice are included, the transverse CT slice becomes a valuable tool in accurately determining what parts of the anatomy are hit by the beam. The transverse slice is also useful in evaluation angles between transverse beams. The transverse CT slice is an important tool emphasized in the training of radiation oncologists and incorporated into treatment planning CAD tools such as the *virtual simulator* program used in the UNC Hospitals' Radiation Oncology Department. In light of this, it is not surprising that the subjects would want to try viewing the patient from one of the poles. Figure 7-22 shows the traces of two subjects who did use this technique. Subject 2 in Figure 7-22a moved to the superior pole a couple of times and also visited the inferior pole, which is evidenced by the traces extending beyond the square limit of the figure. Figure 7-22b shows Subject 10 repeatedly visiting the superior pole. As might be expected, the weight and balance of the head-mounted display made polar views more difficult to attain in Orbital mode than in Joystick mode. Looking straight down for a superior pole view was very demanding of the subject's neck muscles, and looking straight up for an inferior pole view was nearly impossible. Unfortunately, this technique was thwarted by two aspects of the experiment software. Because the polar views represented illegal beams in that they did not enter and exit the patient through the skin, the big red *forbidden* symbol was displayed, distracting the subjects and obscuring parts of the anatomy. In addition, polar views of the treatment volume were often obscured by other anatomical structures. I would suggest that further development of this tool use more subtle and less obtrusive illegal beam alarms and pro-

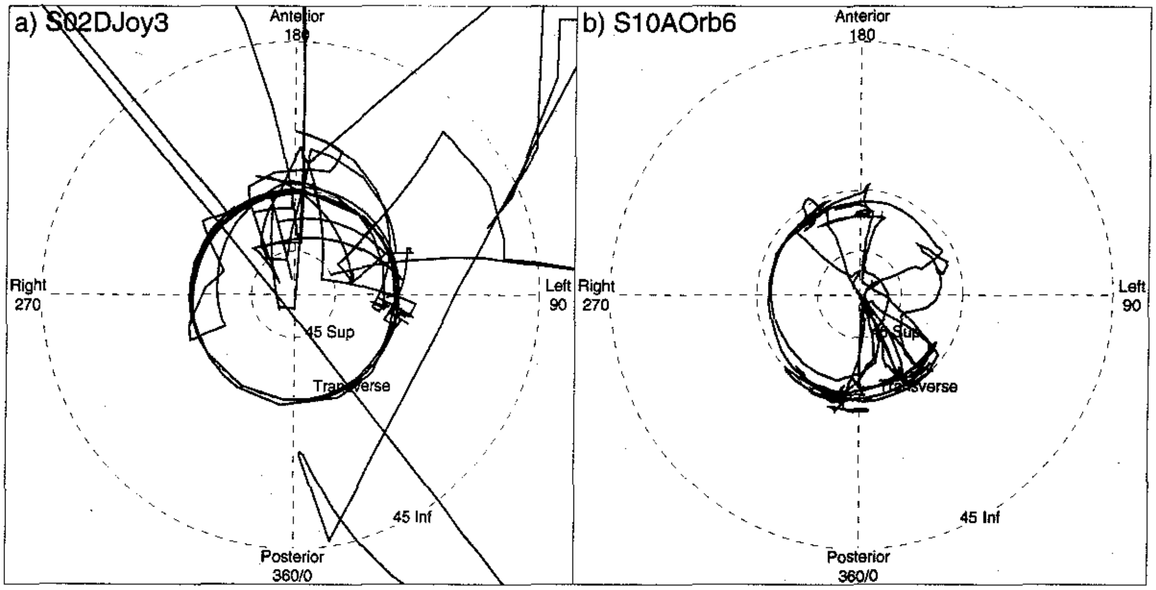


Figure 7-22. Examples of polar viewing to examine transverse plane. Traces extending beyond limit of diagram in part a) are evidence of Subject 2 using the joystick to move to inferior pole to view anatomy. Part b) shows Subject 10 repeatedly visiting superior pole.

vide a clipping plane with which the user could eliminate obstructing objects. And with modern treatment plans that are not transversely coplanar, it is inappropriate to view the anatomy from the poles to examine inter-beam angles. Instead, it would be helpful to provide a snap-to-normal feature which would automatically provide users with a view normal to the plane containing the two beams in question. Also displaying the slice through the CT dataset along this plane would provide an additional benefit.

7.7.6. Azimuth coverage

There was great variation in the amount of azimuth coverage used in the trials. Figure 7-23 compares an Orbital trial of Subject 3, in which he explored only 120° of azimuth, with a Joystick trial of Subject 12, who covered all 360° of azimuth. In general, Joystick mode users tended to cover more azimuth than Orbital mode users, often circumnavigating the patient's anatomy three or four times in the course of targeting their beams. Orbital mode users, on the other hand, managed at most one circumnavigation, and this was due to the head-mounted display's bundle of power and video cables complicating extensive head turning, and a lack of stability discouraging the users from moving their head and feet very much.

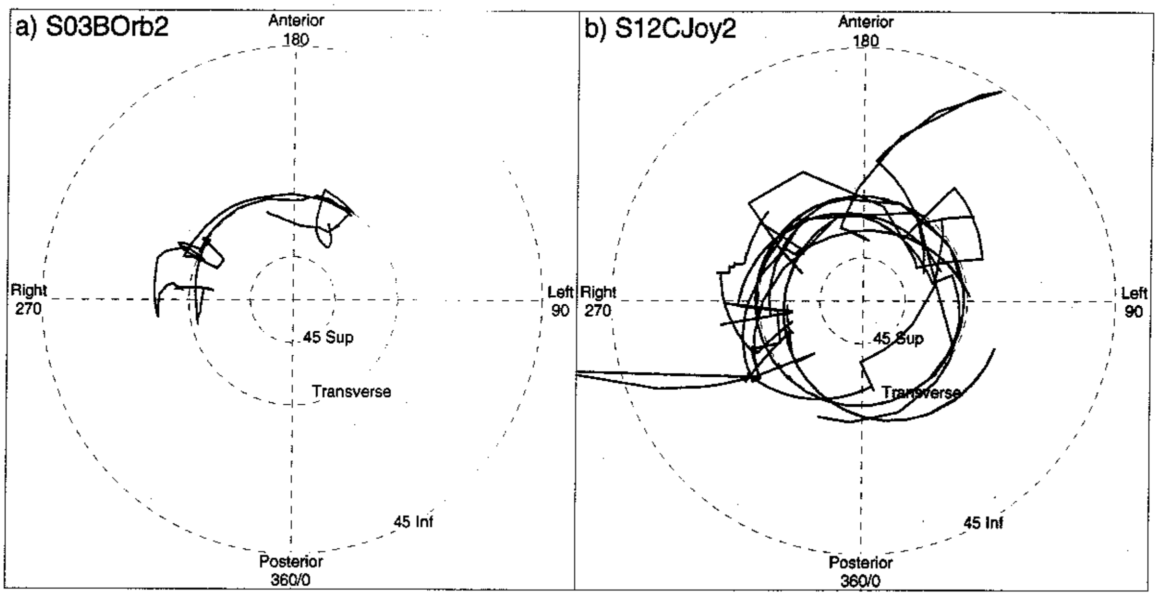


Figure 7-23. Comparison of azimuth coverage for two subjects targeting Case 2.

7.8. Techniques

In addition to the polar viewing discussed in the previous section, several other common techniques were observed in the targeting trials. These techniques were not restricted to one navigation mode or the other, and they present user behaviors that should be supported in treatment planning systems.

Most common was the practice of setting a beam and then immediately moving to the opposite end to examine the opposite direction. If a certain beam path appears to be good in terms of missing critical structures and providing good access to the treatment volume, then the opposite path will very likely also be good. Sometimes subjects would set a single beam, swing around to and examine the opposite direction, and then change that single beam to an opposed pair of beams. More often subjects would initially set an opposed pair and then swing around to verify that the opposite direction was good. One subject preferred to set a beam and then rotate 90°, because he felt that the sideways view of the beam was more helpful in determining the quality of the beam than the beam's-eye view. Viewing from the side yielded a more complete impression of what was happening along the whole length of the beam. For example, the length of the beam's passage through the lung can only be determined from a side view, and not from a beam's-eye view.

Moving to the opposite end of a beam is a special case of the more general technique of aligning oneself with some target direction. Subjects often revisited and reevaluated beams set earlier in the session, sometimes deciding to delete or alter them. Several subjects expressed desires for being able to align with the anterior-posterior axis and with the right-left lateral axis. In general, it may be helpful to provide some form of snapping either to fiducial directions, to previously set beams, and/or to some user-specified reference.

7.9. Anatomical representation

The four different representations of anatomical structures (shaded surface, mesh, points, invisible) provided a "poor man's" way of providing transparency with very little performance penalty. The anatomical representations at the beginning of each trial was skin as points, lungs as mesh, treatment volume, trachea, spinal cord, heart as shaded surfaces. Many subjects made the skin invisible during targeting to concentrate on the internal structures. Only one subject, however, who was very concerned with skin effect, turned the skin back on to check the beams' interaction with the skin. Similarly, most subjects reduced the visibility of the lungs, by changing them to points or making them invisible, to get a clearer picture of the relationship between the treatment volume and the spinal cord. Most of those who made the lungs invisible eventually turned the lungs back on after targeting to check the irradiated lung volume. Those subjects who did not examine their beams with respect to the skin and lungs were concentrating on other aspects of beam targeting, such as missing the cord or having adequate separation between beams, and apparently had an adequate mental impression of where those structures were to satisfy any concerns they had regarding skin effect and lung volume.

Because it is important to see and understand the internal situation in the full anatomical context, from my observation of the subjects I feel that more effective methods of providing transparency are required in this application. The mesh representation was too dense to really be able to see through it—a problem compounded by the coarse resolution of the liquid crystal displays. The point representation was too sparse to be able to understand the shape of the anatomical structure without gross movement. Finer gradation in transparency level and higher resolution displays would help in this respect.

Targeting with Joystick mode revealed another interesting aspect of how the subjects liked to view the anatomy. In Joystick mode it was possible, and very likely, to rotate the anatomy so that its vertical was not aligned with the view vertical. For example, there are an infinite number of ways to move from a vertically-aligned anterior view to a posterior view. Horizontal rotation will move the view along the equator to arrive at a posterior view that is also vertically-aligned. Vertical rotation will bring the view over one of the poles of the solution sphere and result in a posterior view that is upside down. Subjects preferred to have the patient's anatomy aligned vertically, and were often tweaking the knob of the joystick to twist the anatomy into alignment. One exception was Subject 8, who oriented the patient horizontally to help him visualize the patient on the treatment machine table. Some subjects were more tolerant than others of misalignments, but in general none of them was able to work with misalignments greater than 90°.

Vertical misalignment was not such a problem in Orbital mode, because the patient anatomy was aligned with the world-space vertical, and subjects tended to keep their heads also aligned with the world-space vertical. In fact, Subject 8 would have found it quite difficult to attain a horizontal-patient view with Orbital mode. In designing the experiment I feared that this would yield an unfair advantage to Orbital mode, as I envisioned subjects spending a lot of effort realigning the patient anatomy with their view or making bad judgements from misaligned views. I designed an alternate Joystick mode that used the two axes of the joystick to control changes in view elevation and azimuth. Deflecting the joystick right or left changed azimuth only, so that the view moved along a latitude circle of the solution sphere. Deflecting the joystick forward or back moved the view along a meridian, with movement bounded by the inferior and superior poles. Informal experimentation with this mode demonstrated that it was difficult to use at the larger elevations. Subjects became confused as they neared the poles, where azimuth changes resulting from left-right joystick movement appeared as rotation about the gaze vector. I felt it was safer to stay with the more familiar joystick function of rotation about eye-space axes, even with the misalignment problem, than to introduce the new joystick function which always kept the verticals aligned, but was confusing to use. The use of a positional joystick would probably do much to alleviate this problem of maintaining alignment while still providing intuitive control of the view.

7.10. Volumetric analysis

A volumetric analysis was conducted in which each beam configuration was represented by a three-dimensional grid of voxels. For each configuration, the contours of each anastruct were scan-converted into the grid and filled, so that each voxel in the grid was labeled according to the anatomical structure or structures that contained it. Because the contours of different anatomical structures sometimes overlapped, it was possible for a particular voxel to be labeled with more than one structure. For example, treatment volume contours often overlapped lung contours.

After this segmentation of the grid according to anatomical structure, each beam of the configuration was examined, and those voxels that projected into the treatment volume from the beam source center of projection were labeled as belonging to that beam. The result was a volumetric representation of a perfectly conforming beam. A given voxel could belong to more than one beam, and all the voxels in the treatment volume were contained in each of the beams of the configuration.

7.10.1. Spinal cord irradiation

One aspect of the beam configuration that can be easily examined with the volumetric representation is the irradiation of the spinal cord. Subjects were instructed that the spinal cord had already been irradiated to tolerance, and that the configuration they were designing should avoid the cord at all costs to prevent severe irreparable damage. Other objectives, however, such as minimizing the lung volume irradiated and minimizing the distance traversed by the beams in the body, encouraged subjects to bring their beams as close as possible to the spinal cord while not irradiating the cord. If the volumetric representation contained voxels that belonged to both the spinal cord and a treatment beam, then the subject failed at avoiding the cord.

Figure 7-24 and Table 7-7 present the volume of spinal cord irradiated for each beam configuration, broken down by navigation mode used and sequence. Of the twenty-six beam configurations, nine (33%) of them irradiated the cord to some extent. While the independent-measures analysis of Table 7-7 shows no significant effect due to navigation mode, the repeated-measures distribution of Figure 7-25 shows that seven of the thirteen subjects clipped the spinal cord, and six of those seven irradiated more cord when using Orbital mode than when using Joystick mode.

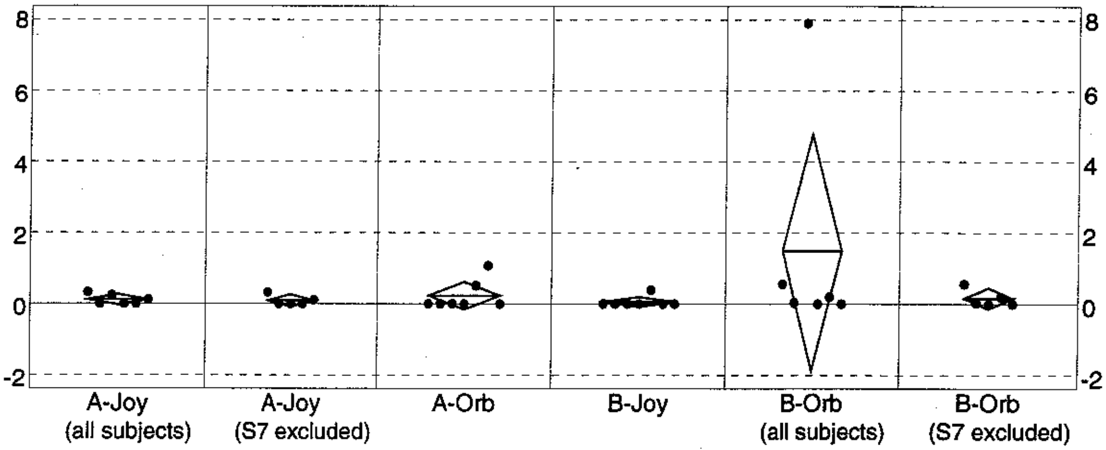


Figure 7-24. Distribution of spinal cord irradiation volume (cc) by sequence and navigation mode. A|B indicates sequence: A=1st trial, B=2nd trial. Joy|Orb indicates navigation mode used. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample I	vs. Sample II	*t*	*df*	Prob > |t| *p*	Est'd. Diff. of Means (I–II)	Lower 95% Conf. Limit	Upper 95% Conf. Limit
A-Joy (S7 excl.)	vs. A-Orb	0.711	10	0.494	-0.140±0.439	-0.579	0.299
B-Joy	vs. B-Orb (S7 excl.)	0.924	10	0.377	-0.102±0.246	-0.348	0.144
A-Joy (all)	vs. A-Orb	0.635	11	0.538	-0.114±0.396	-0.510	0.282
B-Joy	vs B-Orb (all)	1.170	11	0.267	-1.393±2.620	-4.014	1.227

Table 7-7. Comparison of spinal cord irradiation volumes using Student's t-test of null hypothesis $H_0{:}\mu_I - \mu_{II} = 0$. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

There are several reasons why cord hits occurred when every subject had it foremost in his mind to avoid the spinal cord. Probably the biggest reason is the poor display quality in the head-mounted display. The low resolution of the liquid crystal screens impedes the perception of fine detail in the image, and the overlap of treatment volume and spinal cord in the beam's-eye view may not have been discernible with the head-mounted display. Also, critical regions of the tumor-cord boundary can be occluded by other anatomical structures or by representations of beams already set. In his Orbital mode trial, Subject 7 set so many beams that he was unable to clearly see the silhouettes of the treatment volume and the spinal cord, and a relatively large amount of spinal cord was hit by treatment beams. Figures 7-24 and 7-25 clearly show how anomalous his performance was relative to the other subjects.

The above hindrances would apply equally to both navigation modes. Other problems apply solely to Orbital mode, and would therefore support the observation that Orbital mode users are more apt to hit the spinal cord than Joystick mode users. These problems include image lag, tracker error, and inherent imprecision of control using head

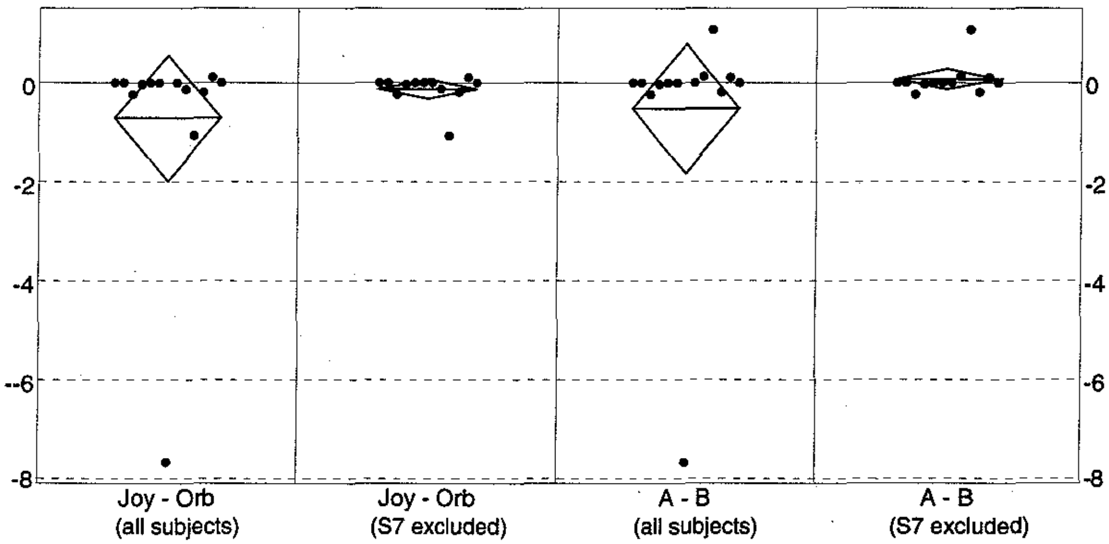


Figure 7-25. Repeated-measures analysis for spinal cord irradiation volume (cc). Joy-Orb shows difference between Joystick mode performance and Orbital mode performance for each subject. A-B shows performance difference between Trial A (1st trial) and Trial B (2nd trial) for each subject. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample	*t*	*df*	Prob > |t| *p*	Est'd. Mean	Lower 95% Conf. Limit	Upper 95% Conf. Limit
Joy-Orb (S7 excl.)	-1.435	11	0.179	-0.129±0.198	-0.328	0.069
A-B (S7 excl.)	-0.737	11	0.477	0.071±0.211	-0.140	0.282
Joy-Orb (all)	-1.211	12	0.249	-0.709±1.276	-1.985	0.567
A-B (all)	-0.872	12	0.400	-0.524±1.311	-1.835	0.786

Table 7-8. Student's t-test of null hypothesis, H_0: $\mu_D=0$ for each irradiated spinal cord volume repeated measures sample. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

and neck muscles. All three of these factors would complicate the precise positioning of treatment beams close to the spinal cord.

7.10.2. Crude dosimetry

The volumetric representation of the beam configuration can be thought of as a very crude dosimetry calculation, in which the dose received by a particular voxel is directly proportional to the number of beams that contain it. Voxels belonging to all the beams receive 100% of the target dose, and voxels that belong to only half the beams in the configuration receive 50% of the target dose. This interpretation assumes that all beams have been equally weighted and no beam modifiers are being used except for a perfectly shaped block. In addition, important radiation physics processes, such as scattering and attenuation, are assumed to be non-existent.

Conceivably this crude dosimetry could have been used to evaluate and compare beam configurations, but I felt this would have yielded unreliable results. Configuration judges could have viewed isodose contours representing this dosimetry, but basing judgements on an approximate dose distribution, for which the error would be difficult to characterize, would have been too dangerous. I put more faith in the judges' abilities to use their experience and to visualize dose distributions.

Dose-volume histograms also could have been easily generated from the volumetric grid by counting the number of voxels in each anatomical structure that belong to n beams, where n varies from 0 to the total number of beams used. They, too, would have been of minimal usefulness. First, dose-volume histograms are meaningful only when computed for the entire anatomical structure., and only two of the six cases used in this study contained complete lung models. Second, dose-volume histograms are typically used to supplement analysis of a dose distribution, but as mentioned above, I did not trust the crude dose distribution. Third, although methods have been developed for computing tissue complication probabilities from dose-volume histograms, which can then be combined to compute an overall score for the treatment plan, these methods require radiobiologic data and expertise which is beyond the scope of this research. In addition, because they are based on such crude dosimetry, such probabilities would be highly suspect.

Chapter 8

User Study: Results and Discussion —Critique

After all subjects had completed the targeting phase of the experiment, each beam configuration was critiqued by a panel of expert judges consisting of the targeting subjects and two other radiation oncology professionals. For each beam configuration, each judge assigned a score for each of the five criteria presented in the Criteria Survey (Chapter 6), and also assigned an overall quality score for the beam configuration. In addition, the importance responses given in each judge's Criteria Survey were used as relative weights in the computation of an objective overall score. Each judge reviewed all the beam configurations designed for a single case, which was specifically not a case for which the judge had targeted beams in the previous phase. Each judge was presented the six or seven configurations for his case in random order, and had the option of returning to any previously reviewed configuration at any point in the sequence. The intent was for each judge to become an expert on one particular case and to be able to score all the configurations for that case on an equal footing.

Several unanticipated complications arose in the reviewing process which reduce the validity of the critiques. The biggest complication, as each judge informed me, was that critiquing a beam configuration was VERY difficult without seeing the resulting dose distribution. The dose distribution represents the "bottom line" of a treatment plan, and without it a lot of guesswork is required in the evaluation. Having access to the dose distribution would resolve uncertainties in the judge's mind and reduce the inter-judge variability in the scores. I did not include dose calculations in this study because I wanted to concentrate on the geometric aspects of beam targeting, the part of treatment planning that would be affected by the introduction of a head-tracked navigation mode. Computing dose distributions would require specification of beam weights and beam modifiers, which would present a source of variability not related to navigation mode. Critiquing might be easier, but it would be difficult to determine if differences in dose distribution quality were due to differences in navigation mode used to target beams, or to differences in the quality of the specification of beam modifiers and beam weights.

Another source of variability was ambiguity in the scales used by the judges in scoring the beam configurations. For each criterion the judge assigned the configuration a score from 1 to 7, where in general 1 meant that the configuration poorly satisfied the criterion and 7 meant that it satisfied the criterion well. The ambiguity resulted from the question of whether the scale should be applied to the configuration in an absolute sense or in a relative sense. For example, the scenario used in this study—a lung tumor with the spinal cord already at tolerance—pretty much ensured that beam configurations would tend to be laterally oriented, and that treatment beams would pass through the width of the body. Consequently, on an absolute scale these beam configurations would rate quite low according to the Beam Length criteria, but what I intended was a relative scale in which 7 indicated "couldn't do any better for this case" and 1 indicated "couldn't do any worse." It was not until a judge asked me for clarification about this that I realized that some of the prior judges might have made the wrong assumption. Even with the correct assumption, however, judges indicated that it was difficult to know what was optimal for a particular case without first taking time to study the case and come up with their own solution. In light of this it may have been better to have each judge score a case on which he had worked.

As with the targeting of beams, the point was again raised that good critiquing required background knowledge of the patient and case. Such information as whether the therapy is palliative or curative and whether the patient can tolerate long treatment sessions is necessary any time an evaluation of a beam configuration or treatment plan must be made.

The tool used to study the beam configurations was *xvsim*, the virtual simulation program developed at the UNC Hospitals' Radiation Oncology Department. *xvsim* is in regular clinical use at UNC as a treatment planning tool, so all the judges were acquainted with its capabilities and familiar with interpreting its various displays. As with the targeting task, different judges brought different approaches to the task of critiquing beam configurations. The three main sources of information provided by *xvsim*—the shaded surface representation of the anatomy presented in beam's-eye view (BEV), the digitally reconstructed radiograph (DRR) continually updated as the beams-eye view changed, and the series of CT slice images which included beam outlines—were used to different degrees by different judges. Judge 1 used the DRR and the shaded surface BEV exclusively, and essentially ignored the CT slices. Judge 10 considered the shaded surface BEV good only for understanding the general, overall situation, and used the CT slices to answer specific questions of beam coverage. Judge 12 first used the shaded surface BEV to get the overall picture and then inspected CT slices to check spinal cord avoidance. Judge 0 examined the radiographs for each beam. "If I understand the radiograph, then that's all I need to look at. If I don't understand the radiograph, then I look [at the shaded-surface BEV]. The only reason I go to the CT scans is I want to be sure that the cord is really out of the field." Judge 3 studies all the CT slices to ensure tumor coverage and evaluate avoidance. Judge 4 felt most comfortable working with the CT slices, but used the shaded-surface rendering to get the overall picture. Judge 5 worked primarily with the CT slices, shrinking them down so that he could see many of them on the screen at once. Judge 13 relied mostly on the DRR and paid very little attention to the CT slices and the shaded-surface views.

The difficulties of using *xvsim* were pointed out by Judges 9 and 8, but were experienced to some extent by most of the judges. Judge 9 depended strongly on CT slices for his evaluation, and he was somewhat challenged to fully understand beams that were off the transverse plane. Comprehending such beams required integrating the information presented on successive CT slices, and Judge 9 suggested that non-transverse beams would be easier to understand if he could view a longitudinal cut through the CT data that would show the entire path of the beam in one image, just as a normal transverse CT slice would for a transverse beam. Judge 8 reported that even when using a state-of-the-art three-dimensional treatment planning tool such as *xvsim*, clinical practice still relies heavily on cardinal angle beams. This is a result of radiation oncology training that uses cardinal angle diagnostic images, and Judge 8 feels that there is not enough experience with oblique views to develop sufficient understanding. In its current state "*xvsim* is not good enough and not being used enough in the clinic to break the mold of cardinal angle treatments." For beam configurations comprising more than six beams, many subjects had difficulty isolating and studying one particular beam amidst the clutter of the display of all the beams. Much time was spent changing beam colors and anatomical representations, which required many button presses and menu selections. Additionally, *xvsim's* performance suffered from low update rates and large lags, which complicated exploration of the anatomy and beam configuration.

I preferred to let each judge operate *xvsim* himself to allow freer thinking and exploration. Two judges, however, did not feel competent enough in "driving" *xvsim* to do it themselves, and engaged me as their "chauffeur." This required extensive verbal communication to enable me to perform the desired operations that would allow them to study the beam configuration, and this verbalization most probably hindered the evaluation process. The other judges operated *xvsim* themselves, but with varying levels of ability. Ideally all judges would have been trained to a specified level of competence with the program, but limited judge time made that infeasible.

It would be a valid criticism to say that using old tools will constrain thinking and seeing to conventional modes, and will perhaps impede the appreciation of beam configurations designed with new methods. Would the improvements provided by intuitive navigation, if any, be apparent if old navigation methods are used to evaluate the results, or would they be masked by deficiencies in the old methods? Given the difficulties experienced by the judges, the danger of the latter result is clearly present. In designing this study, however, I felt that *xvsim* provided a known base from which we could extend our knowledge. All the judges were accustomed to evaluating and designing treatment plans with *xvsim* and making decisions based upon the information provided by *xvsim*, and I believed that the judgements they made on the beam configurations would have more validity than those made using some other presentation method. I still believe this to be true, despite the fact that some judges did experienced difficulties evaluating some of the configurations with *xvsim*.

A possible alternative scoring method would be to have a beam configuration evaluated by a judge who knew the relative importance values of the five criteria for the configuration's designer. With this information, the judge would be asked to answer the question, "How well does the configuration reflect the designer's values?" One would be looking for an effect in which one navigation mode enabled designers to better meet their goals than the other. This is a very interesting approach that would theoretically bypass confounding inter-subject factors such as the Practicality effect. I am skeptical, however, that the designer's point of view could be adequately communicated to the judge for accurate evaluation. I do not feel that the five relative importance values would be sufficient. One could consider the designers themselves to be the best judges of how well their solutions met their goals, in which case the fact that all subjects felt that they performed no differently between the two navigation modes (Section 7.3) strongly suggests equivalence between the two modes in this respect.

To analyze the critique responses, the two or three scores for each beam configuration were averaged together. Simple independent-measures t-tests were run on the these averaged values, as were repeated-measures analyses. The following sections discuss the results for each score category.

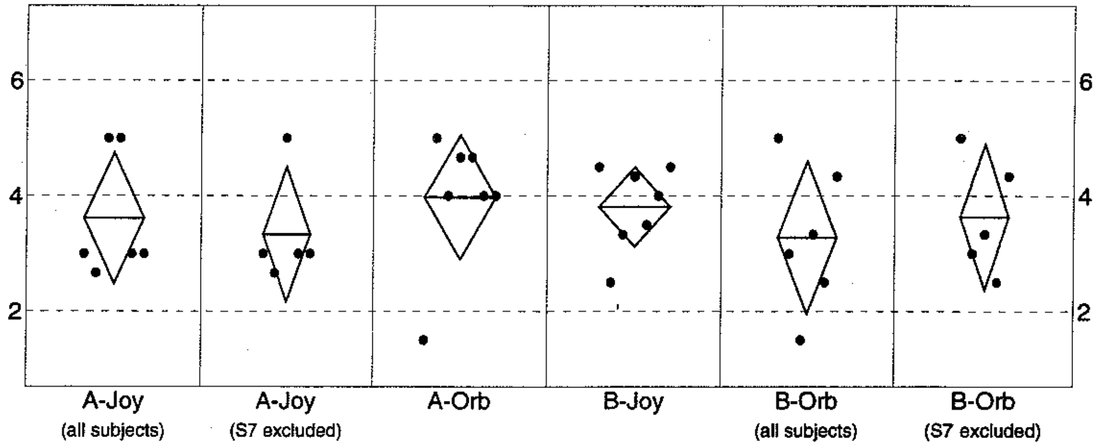


Figure 8-1. Distribution of Length scores by sequence and navigation mode. A|B indicates sequence: A=1st trial, B=2nd trial. Joy|Orb indicates navigation mode used. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample I	vs. Sample II	*t*	*df*	Prob > |t| *p*	Est'd. Diff. of Means (I–II)	Lower 95% Conf. Limit	Upper 95% Conf. Limit
A-Joy (S7 excl.)	vs. A-Orb	1.015	10	0.334	-0.643±1.411	-2.054	0.768
B-Joy	vs. B-Orb (S7 excl.)	0.349	10	0.734	0.177±1.125	-0.948	1.302
A-Joy (all)	vs. A-Orb	0.582	11	0.573	-0.365±1.382	-1.747	1.017
B-Joy	vs B-Orb (all)	0.946	11	0.364	0.532±1.237	-0.705	1.769

Table 8-1. Comparison of Length score samples using Student's t-test of null hypothesis $H_0:\mu_I-\mu_{II} = 0$. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

8.1. Beam length

The average critique scores for Beam Length are presented in Figure 8-1 and Table 8-1. While no significant navigation mode effect is seen here, there is a slight suggestion of a Practicality effect. The more practical A-Orb/B-Joy subjects averaged slightly higher scores than the A-Joy/B-Orb subjects. Practically-guided beams configurations will tend to be simple and not stray from the transverse plane. More complex configurations will usually have additional beams that are off the transverse plane, thereby making them longer than the transverse plane beam. This would yield lower Beam Length scores for the more complex, less practical beam configurations. This concurs with the analysis of the Simplicity scores, in which the A-Orb/B-Joy subjects also averaged higher scores than the A-Joy/B-Orb subjects. The repeated-measures analyses presented in Figure 8-2 and Table 8-2 show no significant effects, either due to navigation mode or to sequence.

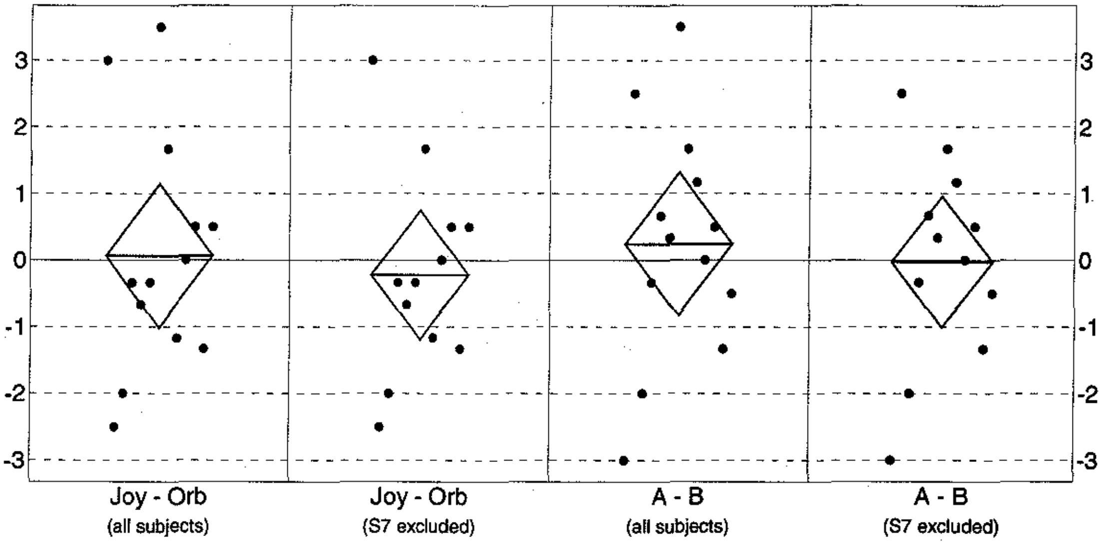


Figure 8-2. Repeated-measures analysis for Length score. Joy-Orb shows difference between Joystick mode performance and Orbital mode performance for each subject. A-B shows performance difference between Trial A (1st trial) and Trial B (2nd trial) for each subject. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample	*t*	*df*	Prob > |t| *p*	Est'd. Mean	Lower 95% Conf. Limit	Upper 95% Conf. Limit
Joy-Orb (S7 excl.)	-0.503	11	0.625	-0.222±0.972	-1.194	0.750
A-B (S7 excl.)	-0.062	11	0.952	-0.028±0.983	-1.011	0.955
Joy-Orb (all)	0.129	12	0.900	0.064±1.083	-1.019	1.147
A-B (all)	0.495	12	0.630	0.244±1.073	-0.829	1.316

Table 8-2. Student's t-test of null hypothesis, H_0: $\mu_D=0$ for eachLength repeated-measures sample. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

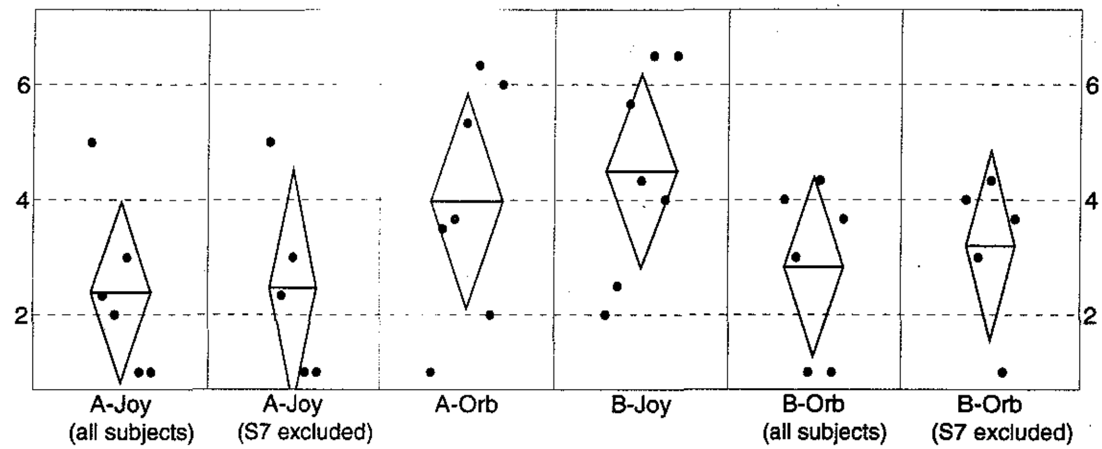


Figure 8-3. Distribution of Simplicity scores by sequence and navigation mode. AIB indicates sequence: A=1st trial, B=2nd trial. JoylOrb indicates navigation mode used. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample I	vs. Sample II	*t*	*df*	Prob > ltl *p*	Est'd. Diff. of Means (I–II)	Lower 95% Conf. Limit	Upper 95% Conf. Limit
A-Joy (S7 excl.)	vs. A-Orb	1.366	10	0.202	-1.509±2.462	-3.971	0.953
B-Joy	vs. B-Orb (S7 excl.)	1.353	10	0.206	1.300±2.140	-0.840	3.440
A-Joy (all)	vs. A-Orb	1.582	11	0.142	-1.587±2.209	-3.796	0.621
B-Joy	vs B-Orb (all)	1.786	11	0.102	1.667±2.054	-0.387	3.721

Table 8-3. Comparison of Simplicity score samples using Student's t-test of null hypothesis $H_0{:}\mu_I - \mu_{II} = 0$. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

8.2. Simplicity

The distribution of Simplicity scores is presented in Figure 8-3, with the accompanying statistical analyses presented by Table 8-3. Although the observed differences across navigation mode are not statistically significant, the Practicality effect is visible here, as the A-Orb/B-Joy group of subjects, which included more of the practically-minded subjects, tended to score higher than the more dosimetry-oriented A-Joy/B-Orb subjects. The repeated-measures analyses presented in Figure 8-4 and Table 8-4 show no effect due to navigation mode, but suggest a tendency for the B-trial to be slightly simpler than the A-trial. This may be the indicative of some type of learning effect that results from subjects becoming more familiar with the capabilities of the system with time.

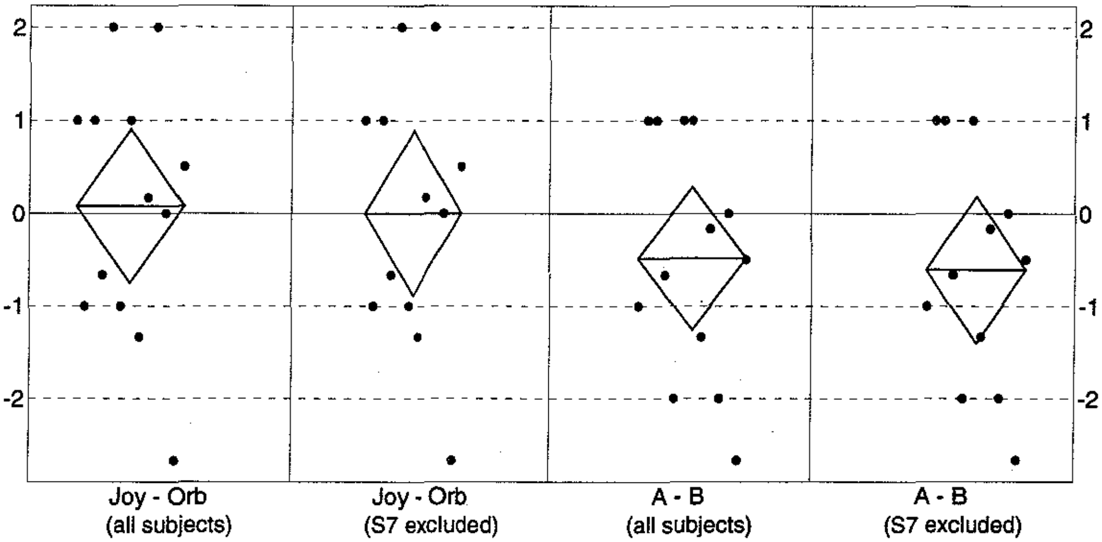


Figure 8-4. Repeated-measures analysis for Simplicity score. Joy-Orb shows difference between Joystick mode performance and Orbital mode performance for each subject. A-B shows performance difference between Trial A (1st trial) and Trial B (2nd trial) for each subject. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample	*t*	*df*	Prob > |t| *p*	Est'd. Mean	Lower 95% Conf. Limit	Upper 95% Conf. Limit
Joy-Orb (S7 excl.)	0.000	11	1.000	0.000±0.891	-0.891	0.891
A-B (S7 excl.)	-1.696	11	0.118	-0.611±0.793	-1.404	0.182
Joy-Orb (all)	0.202	12	0.843	0.077±0.828	-0.751	0.675
A-B (all)	-1.377	12	0.194	-0.487±0.771	-1.258	0.284

Table 8-4. Student's t-test of null hypothesis, H_0: $\mu_D=0$ for each Simplicity repeated-measures sample. Analyses with Subject 7 excluded and include are presented, and estimates of difference of means are also shown.

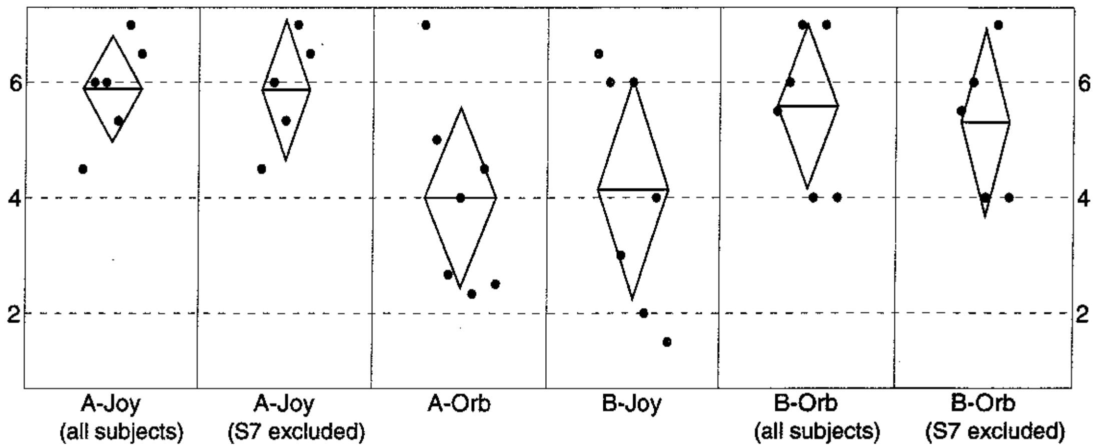


Figure 8-5. Distribution of Originality scores by sequence and navigation mode. A|B indicates sequence: A=1st trial, B=2nd trial. Joy|Orb indicates navigation mode used. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample I	vs. Sample II	t	df	Prob > |t| p	Est'd. Diff. of Means (I–II)	Lower 95% Conf. Limit	Upper 95% Conf. Limit
A-Joy (S7 excl.)	vs. A-Orb	2.205	10	0.052	1.867±1.887	-0.020	3.754
B-Joy	vs. B-Orb (S7 excl.)	1.102	10	0.296	-1.157±2.339	-3.496	1.182
A-Joy (all)	vs. A-Orb	2.462	11	0.032	1.889±1.689	-0.200	3.578
B-Joy	vs B-Orb (all)	1.461	11	0.172	-1.440±2.171	-3.611	0.731

Table 8-5. Comparison of Originality score samples using Student's t-test of null hypothesis $H_0: \mu_I - \mu_{II} = 0$. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

8.3. Originality

The analysis for Originality scores presented in Figure 8-5 and Table 8-5 show differences between the A-Joy samples and A-Orb sample that are statistically significant at the α=0.10 level. The differences between the B samples are not as large nor significant. Rather than conclude, however, that using Joystick mode on the first trial tends to make a subject more creative than using Orbital mode on the first trial, I believe that the Practicality effect is responsible for the differences. We have seen that the A-Orb/B-Joy subject group, which included most of the practically-minded subjects, have scored better on Simplicity and Reproducibility, and here they have scored lower on Originality than the A-Joy/B-Orb subjects. Simple, reproducible configurations do not provide much opportunity for exercising creativity. If a subject is planning to place only one opposed pair of beams, it is generally pretty clear where that opposed pair should go. On the other hand, if a subject is planning on using more beams, then there are more degrees of freedom and more opportunity to try beam directions that others would not think of. In general, judg-

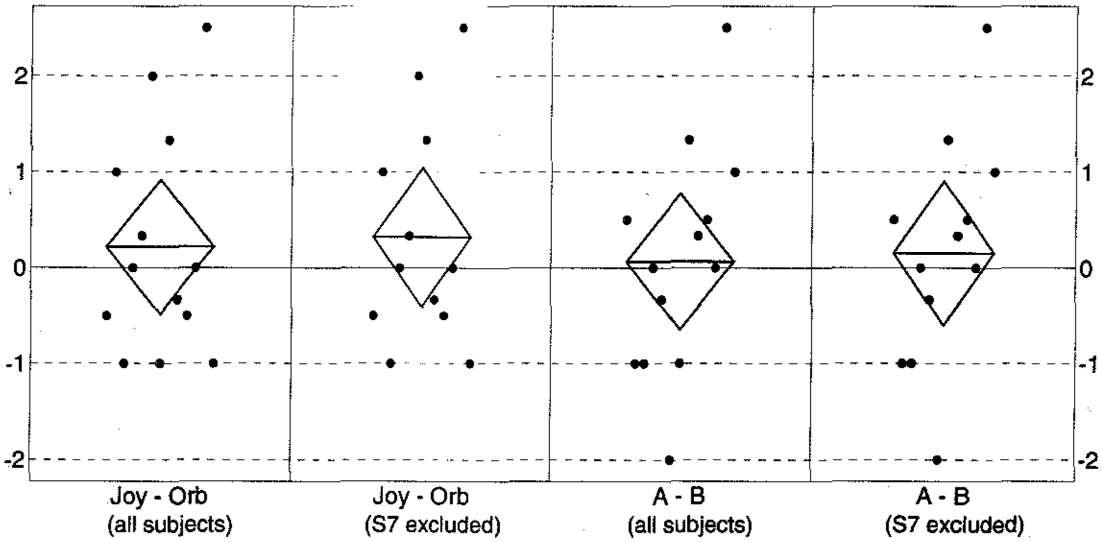


Figure 8-6. Repeated-measures analysis for Originality score. Joy-Orb shows difference between Joystick mode performance and Orbital mode performance for each subject. A-B shows performance difference between Trial A (1st trial) and Trial B (2nd trial) for each subject. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample	*t*	*df*	Prob > |t| *p*	Est'd. Mean	Lower 95% Conf. Limit	Upper 95% Conf. Limit
Joy-Orb (S7 excl.)	0.963	11	0.356	0.319±0.730	-0.411	1.050
A-B (S7 excl.)	0.446	11	0.664	0.153±0.753	-0.601	0.906
Joy-Orb (all)	0.678	12	0.511	0.218±0.701	-0.483	0.919
A-B (all)	0.196	12	0.848	0.064±0.713	-0.649	0.777

Table 8-6. Student's t-test of null hypothesis, H_0: μ_D=0 for each Originality repeated-measures sample. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

es scored beam configurations that used a lot of beams high in Originality. There were only a couple of instances in which a beam configuration scored high in Originality because of its quality. Judge 0 said of one configuration (designed with Orbital mode), "Clever. I like it. Well-done. Probably a very good beam. Not one I would have thought of." On a different configuration (designed with Joystick mode) Judge 11 commented, "This is extremely creative. You would never be able to come up with this in a regular simulation. What I like about this is the combination of creativity and simplicity in being able to achieve this objective of sparing that lung." For the most part, though, highly original beam configurations were not suitable for clinical practice. Figure 8-6 and Table 8-6 show no significant effects revealed by the repeated-measures analysis.

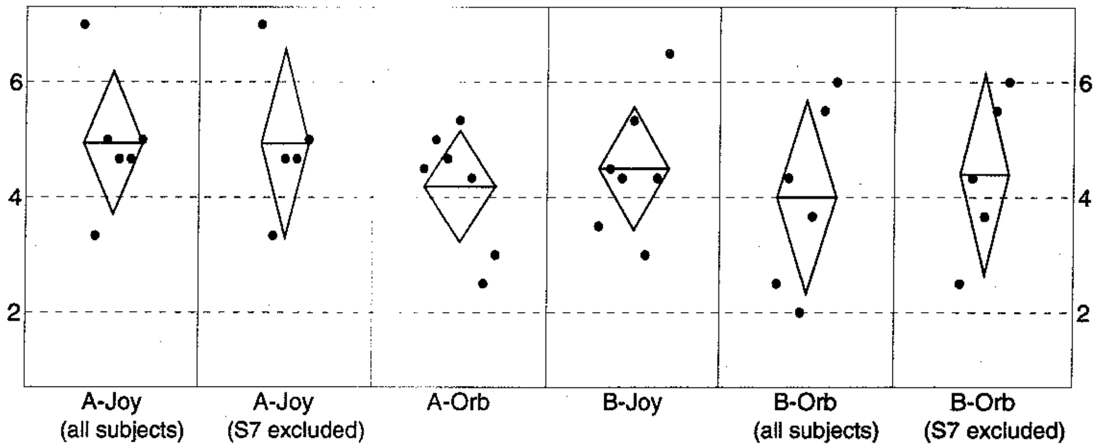


Figure 8-7. Distribution of Avoidance scores by sequence and navigation mode. A|B indicates sequence: A=1st trial, B=2nd trial. Joy|Orb indicates navigation mode used. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample I	vs. Sample II	*t*	*df*	Prob > |t| *p*	Est'd. Diff. of Means (I–II)	Lower 95% Conf. Limit	Upper 95% Conf. Limit
A-Joy (S7 excl.)	vs. A-Orb	1.090	10	0.302	0.743±1.519	-0.776	2.262
B-Joy	vs. B-Orb (S7 excl.)	0.135	10	0.895	0.100±1.646	-1.546	1.746
A-Joy (all)	vs. A-Orb	1.221	11	0.248	0.754±1.360	-0.606	2.114
B-Joy	vs B-Orb (all)	0.655	11	0.526	0.500±1.681	-1.181	2.181

Table 8-7. Comparison of Avoidance score samples using Student's t-test of null hypothesis H_0:$\mu_I - \mu_{II}$ = 0. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

8.4. Avoidance

The analyses for the average Avoidance scores are presented in Figure 8-7 and Table 8-7. No significant effect is seen here across navigation modes. Similarly, the repeated-measures analyses presented in Figure 8-8 and Table 8-8 show no significant effect across navigation mode, nor across sequence.

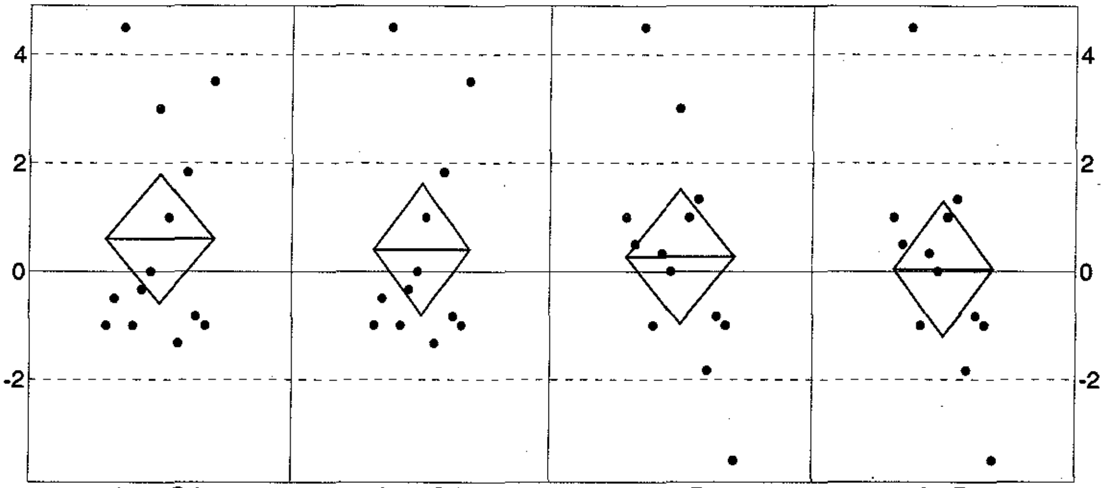


Figure 8-8. Repeated-measures analysis for Avoidance score. Joy-Orb shows difference between Joystick mode performance and Orbital mode performance for each subject. A-B shows performance difference between Trial A (1st trial) and Trial B (2nd trial) for each subject. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample	*t*	*df*	**Prob > |t|** *p*	**Est'd. Mean**	**Lower 95% Conf. Limit**	**Upper 95% Conf. Limit**
Joy-Orb (S7 excl.)	0.725	11	0.484	0.403±1.223	-0.820	1.626
A-B (S7 excl.)	0.073	11	0.943	0.042±1.251	-1.210	1.293
Joy-Orb (all)	1.098	12	0.294	0.603±1.196	-0.751	1.798
A-B (all)	0.472	12	0.645	0.269±1.243	-0.974	1.512

Table 8-8. Student's t-test of null hypothesis, H_0: μ_D=0 for each Avoidance repeated-measures sample. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

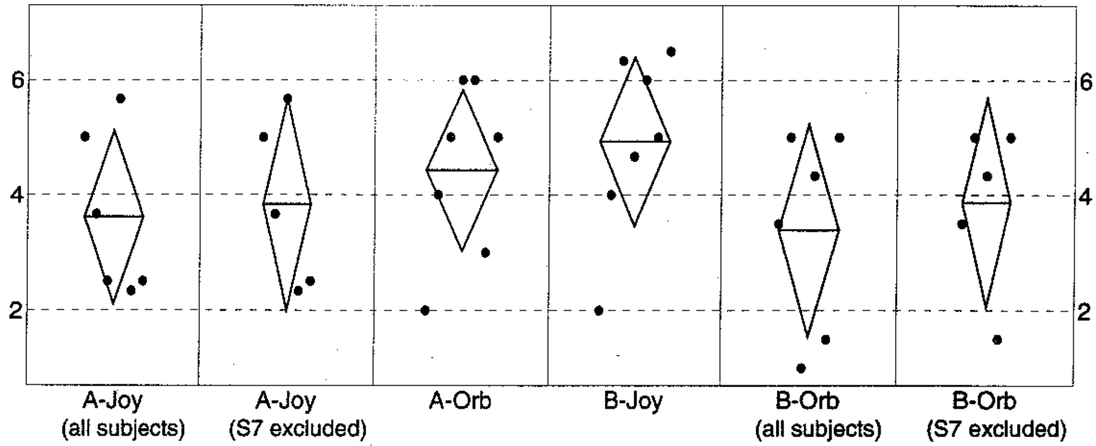


Figure 8-9. Distribution of Reproducibility scores by sequence and navigation mode. A|B indicates sequence: A=1st trial, B=2nd trial. Joy|Orb indicates navigation mode used. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample I	vs. Sample II	*t*	*df*	Prob > |t| *p*	Est'd. Diff. of Means (I–II)	Lower 95% Conf. Limit	Upper 95% Conf. Limit
A-Joy (S7 excl.)	vs. A-Orb	0.678	10	0.513	-0.596±1.957	-2.553	1.361
B-Joy	vs. B-Orb (S7 excl.)	1.180	10	0.265	1.062±1.005	-0.943	3.067
A-Joy (all)	vs. A-Orb	0.995	11	0.341	-0.818±1.808	-2.625	0.990
B-Joy	vs B-Orb (all)	1.663	11	0.125	1.540±2.038	-0.498	3.578

Table 8-9. Comparison of Reproducibility score samples using Student's t-test of null hypothesis $H_0{:}\mu_I - \mu_{II} = 0$. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

8.5. Reproducibility

Figure 8-9 and Table 8-9 show the analysis for Reproducibility. Again, the Practicality effect is apparent, for the A-Orb/B-Joy subjects, who scored better in Simplicity and Beam Length, are seen to also score better in Reproducibility. The differences are not statistically significant, but in light of this pattern appearing in other criteria scores, I believe them to be important. The repeated-measures analyses, shown in Figure 8-10 and Table 8-10 show no significant effects from navigation mode nor from sequence.

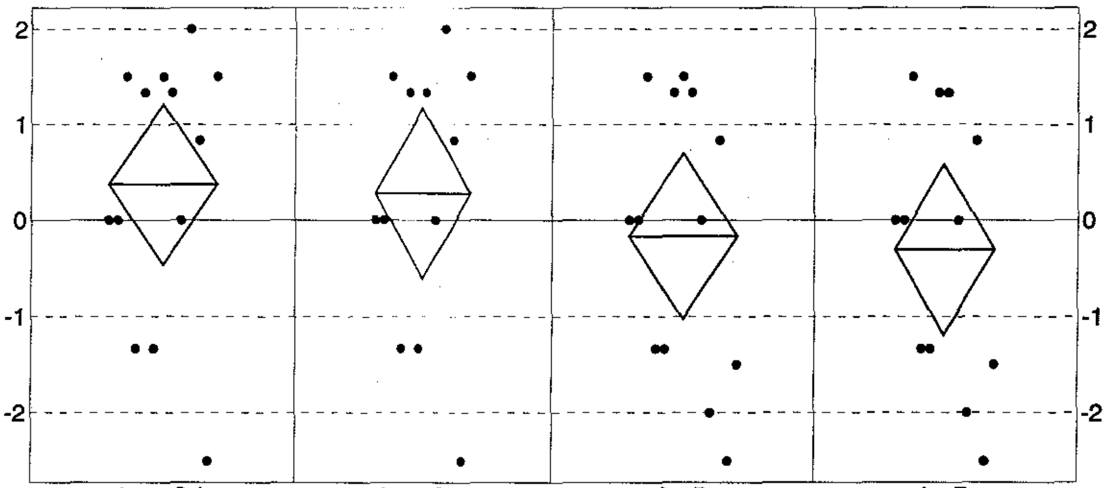


Figure 8-10. Repeated-measures analysis for Reproducibility score. Joy-Orb shows difference between Joystick mode performance and Orbital mode performance for each subject. A-B shows performance difference between Trial A (1st trial) and Trial B (2nd trial) for each subject. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample	*t*	*df*	Prob > |t| *p*	Est'd. Mean	Lower 95% Conf. Limit	Upper 95% Conf. Limit
Joy-Orb (S7 excl.)	0.686	11	0.507	0.278±0.891	-0.613	1.169
A-B (S7 excl.)	-0.758	11	0.464	-0.306±0.887	-1.192	0.270
Joy-Orb (all)	0.968	12	0.352	0.372±0.837	-0.465	1.209
A-B (all)	-0.421	12	0.681	-0.167±0.862	-1.019	0.696

Table 8-10. Student's t-test of null hypothesis, H_0: μ_D=0 for each Reproducibility repeated-measures sample. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

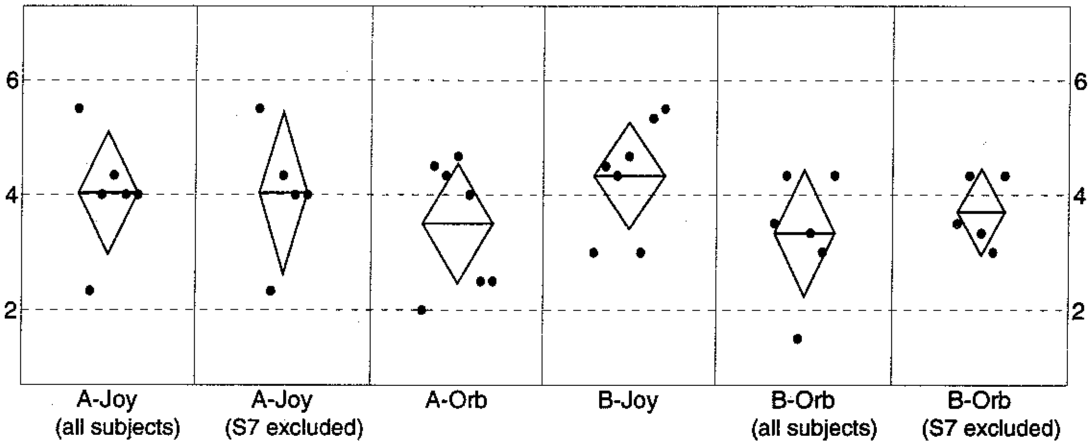


Figure 8-11. Distribution of Overall scores by sequence and navigation mode. AIB indicates sequence: A=1st trial, B=2nd trial. JoyIOrb indicates navigation mode used. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample I	vs. Sample II	*t*	*df*	Prob > ItI *p*	Est'd. Diff. of Means (I–II)	Lower 95% Conf. Limit	Upper 95% Conf. Limit
A-Joy (S7 excl.)	vs. A-Orb	0.809	10	0.438	0.533±1.470	-0.937	2.002
B-Joy	vs. B-Orb (S7 excl.)	1.247	10	0.241	0.633±1.132	-0.499	1.765
A-Joy (all)	vs. A-Orb	0.883	11	0.396	0.528±1.315	-0.787	1.843
B-Joy	vs B-Orb (all)	1.754	11	0.107	1.000±1.255	-0.255	2.255

Table 8-11. Comparison of Overall score samples using Student's t-test of null hypothesis $H_0{:}\mu_I-\mu_{II} = 0$. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

8.6. Overall

Figure 8-11 and Table 8-11 present the independent-measures analysis of the Overall scores assigned by the judges to the beam configurations. Although not statistically significant, there is a suggestion of a navigation mode effect, for in both A-trials and B-trials the Joystick mode configurations averaged higher scores than the Orbital mode configurations. This suggestion is reinforced by the repeated-measures analyses presented in Figure 8-12 and Table 8-12, which show a tendency for Joystick mode configurations to score higher than Orbital mode. I believe this difference to be related to the Joystick-Orbital differences observed for Avoidance and Reproducibility (Figures 8-8 and 8-10) which were not statistically significant but still showed an advantage for Joystick mode.

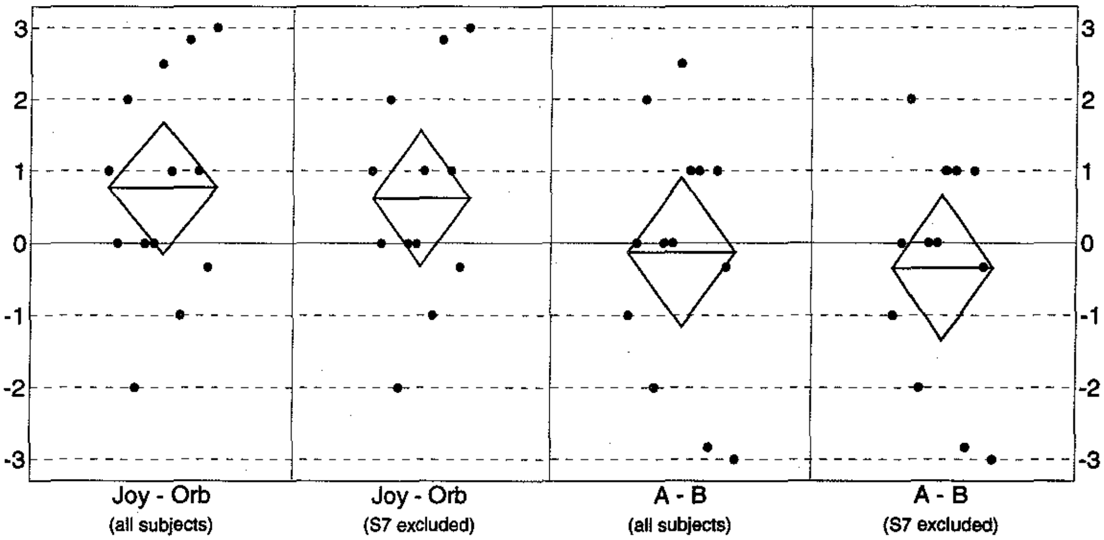


Figure 8-12. Repeated-measures analysis for Overall score. Joy-Orb shows difference between Joystick mode performance and Orbital mode performance for each subject. A-B shows performance difference between Trial A (1st trial) and Trial B (2nd trial) for each subject. Analyses with Subject 7 excluded and included are shown. Diamonds indicate sample mean with 95% confidence interval for estimate of population mean.

Sample	*t*	*df*	Prob > |t| *p*	Est'd. Mean	Lower 95% Conf. Limit	Upper 95% Conf. Limit
Joy-Orb (S7 excl.)	1.452	11	0.174	0.625±0.947	-0.322	1.572
A-B (S7 excl.)	-0.758	11	0.464	-0.347±1.008	-1.355	0.661
Joy-Orb (all)	1.826	12	0.093	0.769±0.918	-0.149	1.687
A-B (all)	-0.270	12	0.792	-0.128±1.035	-1.163	0.906

Table 8-12. Student's t-test of null hypothesis, H_0: μ_D=0 for each Overall repeated-measures sample. Analyses with Subject 7 excluded and included are presented, and estimates of difference of means are also shown.

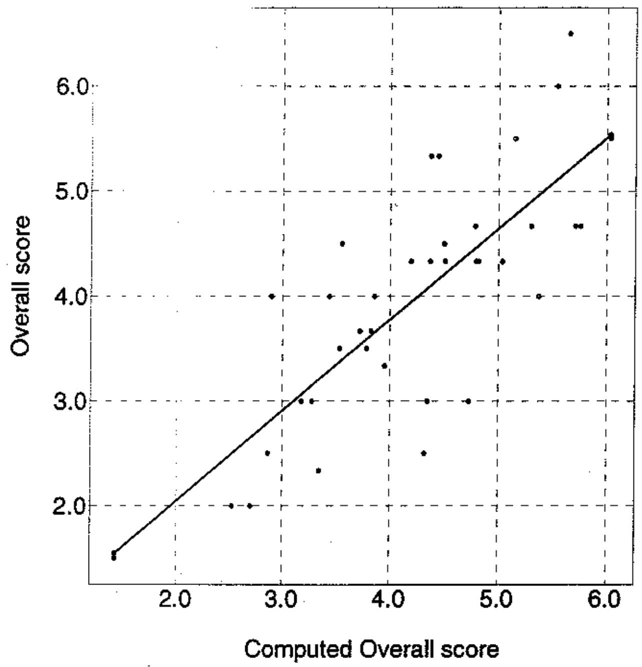


Figure 8-13. Scatter plot of Overall score and Computed Overall score showing linear regression fit.

8.7. Computed overall

Recall from Chapter 6 that each judge completed a Criteria Survey, the responses from which were used to compute each judge's profile of relative weights for the five criteria. For each beam configuration that a judge graded a Computed Overall score was calculated by combining the individual criterion scores assigned by the judge in accordance with the judge's relative weights. Comparing the Computed Overall scores with the Overall scores gives an idea of how well each judge understands his own evaluation process. A linear regression of Overall score onto Computed Overall score produced the regression equation:

$$\text{Overall} = 0.866 * \text{Computed Overall} + 0.313$$

with a coefficient of determination r^2=0.609 and *root mean square error*=0.731. Figure 8-13 shows a scatter plot of the two quantities, showing that individual criterion scores served well as predictors for the Overall score.

Chapter 9

User Study: Summary

9.1. Navigation modes

The information collected in this study and discussed above paints pictures of Orbital and Joystick modes as being complementary, in that each provided advantages that were essentially opposite sides of the same coin.

Joystick mode provided constrained movement, in that the mechanical action of the joystick encouraged users to deflect the joystick only along the cardinal axes. This reduced the joystick to a two-dimensional controller that was used to control only one dimension at a time, and controlling only one dimension at a time gave the users more precise control. Subjects who preferred Joystick mode over Orbital mode (roughly half of the subjects) appreciated this precision, but felt that in order to make the best use of the joystick, its sensitivity and gain[1] must be adjusted to the user's liking. Because subjects were able to sit down, Joystick mode provided more comfort and stability while targeting treatment beams. Regarding task performance, Joystick mode was significantly slower than Orbital mode, averaging a difference of 3 cases/hr. out of a typical performance rate of 20 cases/hr. Judging from the Critique Overall scores, the beam configurations produced by Joystick mode were slightly better than those produced with Orbital mode. Joystick mode also tended to produce fewer cord hits of less volume than Orbital mode.

---

1. Sensitivity is inversely related to the size of the "dead zone" of the joystick, within which deflection of the joystick will have no effect. Once the joystick is deflected beyond the dead zone, the relationship between magnitude of deflection and magnitude of rotation rate determines the gain.

These trends are interesting in light of the fact that almost all of the subjects stated that they felt they performed just as well with either navigation mode.

Orbital mode was preferred by its proponents (roughly half of the subjects) because it provided freer movement through more natural and more intuitive control. Subjects credited this freedom with enabling them to better understand the three dimensional structure of the patient's anatomy. Because it involved head and body movement, however, Orbital mode was more susceptible to the ergonomic problems of the head-mounted display. The front-heaviness of the unit made it more difficult to look up than to look down. Inferior hemisphere coverage was less extensive than with Joystick mode, and some subjects never visited the inferior hemisphere when using Orbital mode. In addition, the bundle of cables carrying power, audio, and video signals to the head-mounted display hindered circumnavigation of the patient anatomy, and somewhat restricted the subject's exploratory freedom. A major drawback of Orbital mode is that subjects often felt unstable and unsure of their balance, and yet they were required to move their feet and their bodies to explore the patient's anatomy. In addition to this instability making the task more difficult, I would expect Orbital mode users to be more susceptible to motion sickness than Joystick mode users, because there is a greater opportunity for vection (it is more likely for the user to feel as if he is orbiting the virtual object) and the weight, balance, and distortion of the head-mounted display may produce sensorimotor rearrangements. As mentioned above, users completed their tasks more quickly with Orbital mode than with Joystick mode, although Orbital mode beam configurations scored lower in Overall quality and tended to hit the spinal cord more.

The two navigation modes have different strengths and weaknesses. Orbital mode is good for the medium- and large-scale movements used to explore the gross structure of a virtual scene, and as such, it is very well suited for the god's-eye views used by many applications to provide overall views of some particular area of interest. The imprecision of the control provided by head and neck muscles, however, makes it relatively unsuited for fine adjustment and precision alignment tasks. When precision is required, Joystick mode provides better results by making use of the excellent fine motor control of human fingers and hands. This suggests that some hybrid or combination of the two modes would be most suitable for use in a clinical treatment planning tool. Most joysticks have a neutral position to which they automatically return when released by the user. Thus no effort is required to maintain a constant view, which may be desirable when evaluating how well a given beam avoids critical structures. In Orbital mode, one must contend with low bandwidth muscular control, a heavy head-mounted display, noisy tracker data, and image lag, all of which contribute to the difficulty of maintaining a constant view. Subjects in this study deflected the joystick almost exclusively along the principal axes. This was probably due to the construction of the joystick, but I believe that it was used to advantage by the subjects by reducing the number of degrees of freedom that had to be controlled. It would be interesting to see how performance changes when using an unbiased joystick.

9.2. Task domain

In this study it has been difficult to find statistically significant effects of navigation mode on task performance. There are two possible explanations. Either no effect exists to be found, or the experiment was inadequate to find the effect. At this point it is impossible to say which is the case, but it is possible to discuss the possible inadequacies of the experiment. Chapter 3 discusses technological problems that might hinder the exposure of a navigation mode effect. The course of this study has shed light on many unanticipated aspects of beam targeting and radiation treatment planning, which may have masked the effect (if it existed) by providing sources of variability that reduced the power and sensitivity of the experiment.

9.2.1. Subjects

I found that subjects vary greatly in their approaches to beam targeting, and this variation, which I called the *practicality effect*, confounded the data greatly. Some subjects aimed at designing beam configurations that were dosimetrically optimal, i.e. that would yield an optimal dose distribution. Others aimed at designing clinically optimal configurations, i.e. configurations that minimized the patient's time obligation and the opportunity for error by using a minimal number of beams. I suspect that navigation mode effects are more likely to be apparent when designing dosimetrically optimal configurations than when designing clinically optimal configurations. While some, but not much, exploration of the patient's anatomy is needed to find the optimal orientation for the single pair of opposed beams to be used in a clinically optimal configuration, extensive exploration and evaluation is required to design the dosimetrically optimal configuration, which will normally require a larger number of beams. In essence they are two different tasks, and unless they are recognized as such and handled appropriately, the results of the study will not be trustworthy.

9.2.2. Cases

Because the spatial arrangement of tumor and normal anatomy differs from case to case, the use of six different lung tumor cases in this study rather than just two assuredly introduced additional variability into the results. In designing the study, however, I felt that the greater generalizability of the results from a study involving six cases outweighed the reduced sensitivity of the study. Despite the mostly inconclusive results of the critique, I still feel the same way, for inconclusive results from a study that uses a larger, more representative sample of the stimulus domain are more meaningful than inconclusive results from a study that used a small, narrow sample.

In addition to the number of cases, the completeness of the case information can affect variability in results. When working in the clinic on a real patient, the treatment planner is familiar with the entire case history, including the patient's physical condition and the goals of the treatment. These factors can deeply affect a treatment plan. If the treatment is to be palliative only, then heroic measures do not need to be taken to design and deliver the perfect treatment plan. If the patient has emphysema, then special attention must be paid to the amount of lung volume that will get irradiated. In the absence of such information, different treatment planners will make different assumptions regarding the patient, and the resulting beam configurations will reflect this variability.

9.2.3. Judges

The critiquing process also provided opportunities for variability to enter the data. The largest source of variability was the lack of dosimetry information. The dose distribution resulting from a given treatment plan is the "bottom line" by which that plan is normally evaluated. The judges in this study evaluated beam configurations without the benefit of such information, and were thus required to estimate and guess about the dose distribution that might by produced by the given configuration. As with the case histories in treatment planning, different judges will make different estimates and the scores they give to the beam configuration will reflect that variability.

There were other sources of judge-related variability in this study. Just as with beam targeting, different judges had different ideas about what makes a good treatment plan, and these ideas affected the scores they gave the beam configurations. In addition, the judges were of varying levels of competence in the operation of *xvsim*, the program used to evaluate the beam configurations. This could have affected their impressions of the beam configurations, as did possibly the different emphases they placed on the three types of information provided by *xvsim* (shaded surface beam's-eye view, DRR beam's-eye view, CT slices).

Chapter 10

Conclusion and Future Directions

10.1. Summary

The goal of this research was to demonstrate that the intuitive navigation provided by a tracked head-mounted display would enable radiation treatment planners to target radiation beams more efficiently, which is to say that equally good beam configurations would be able to be designed more quickly or better configurations in the same amount of time. I believed that a navigation mode that effectively used the natural spatial skills possessed by most human adults would offer advantages over a non-intuitive mode, which did not provide the kinesthetic and vestibular cues required by these spatial skills.

A preliminary study was conducted to evaluate several possible intuitive navigation modes and to determine which would be best suited to the task of targeting radiation treatment beams. Of the four head-tracked navigation modes considered in this study (Walkaround, Walk/Rotate, Immersion, Orbital) the best prospect turned out to be Orbital mode, a new navigation mode that offered many advantages as a user interface for beam targeting. Orbital mode was very easy to learn and master, which was surprising in light of the fact that it had no real-world metaphor to help users understand its operation. It did benefit, however, from a much better fit to the task's two degrees-of-freedom than that of Walkaround and Walk/Rotate modes. Orbital mode also provided the user with a beam's-eye view of the target, an essential tool in radiation treatment planning, and left the user's hands free for other tasks. Three non-head-tracked, non-intuitive navigation modes were also examined in this study, and the velocity-control joystick proved itself superior to both the isometric Spaceball and the six degree-of-freedom mouse.

This preliminary study involved non-expert subjects targeting a beam in an abstract anatomical model, in which tumor and healthy organs were all represented as spheres. Building on the results of this experiment, a full user study was conducted to compare Orbital mode and Joystick mode as used by radiation oncology professionals in designing beam configurations for real-life lung tumor cases. The results of this study were interesting and somewhat unexpected.

Perhaps the most interesting result arising from this experiment was that Orbital mode proved itself to be significantly faster than Joystick mode, enabling its users to average 20 cases/hr. compared to Joystick mode's 17 cases/hr. For busy radiation oncology clinicians, who are often faced with too many patients and too little time, the ability to perform part of their job significantly faster is very helpful and appealing. In light of this, it is somewhat dismaying that the Critique Overall scores suggested Joystick mode to be a better performer than Orbital mode. The statistical analyses showed the conclusion of Orbital mode being faster to be stronger than the conclusion of Joystick mode producing better results.

The power of the radiation oncology study was reduced by sample variances that were too large to yield tight confidence intervals for the population mean differences. The variability in the beam configuration scores derived from numerous sources, because the task of targeting radiation treatment beams is very complex and difficult to define precisely enough for a controlled experiment such as this. The major source was the confounding influence of what I called the Practicality effect—the variability in subjects' approaches to beam targeting. Other contributors included variability in the judges, inadequate patient history, lack of dose distribution, and variability in the spatial and motor skills of the subjects. These factors were not anticipated in the design of this study, but should be properly accounted for in future research.

The observational data gathered during this study proved to be quite valuable in providing insight into the two navigation modes. Analysis of videotapes recorded during the experiment sessions and of the log files recorded during actual beam targeting shed much light on the types of movements used by subjects to explore the patient's anatomy.

Orbital mode movement was seen to be somewhat freer and more direct than movement in Joystick mode. This was most likely due to the mechanical operation of the joystick, which exhibited a preference for deflection along the principal axes. In terms of subject preferences, Orbital mode was not the hands-down winner I expected. The subjects were split roughly 50-50 in their preferences, but all subjects agreed that the navigation mode had very little effect on their performance of the task. The complaints against Orbital mode were mostly rooted in ergonomic and human factors problems. Chief among these was the lack of stability experienced in Orbital mode resulting mainly from image lag, but I believe this may be adequately remedied by having the users sit in a swiveling chair. The imbalance of the head-mounted display plagued subjects with smaller heads, and there was evidence that exploration of the solution space was compromised as a result. Imbalance and excessive weight decreased the precision of control. Subjects also complained about discomfort from heat and the unit's tight headband. Poor display quality may have contributed to targeting errors.

For these reasons head-mounted display technology in its current state cannot be considered suitable for clinical use. The technological problems mentioned above outweigh the modest performance speedup. I believe, however, that time will see these problems reduced or eliminated, and once this occurs the advantages provided by Orbital mode—ease of movement, intuitive control—may very well bring intuitive navigation into the mainstream of clinical use.

10.2. Future directions

10.2.1. Observational follow-up

Given the inconclusive statistical results of the user study, there is great temptation to repeat the experiment with its known design deficiencies corrected. At this point, however, it may be more fruitful and more beneficial to the user interface community at large to pursue the questions raised by the observational data. This will enable us to build a more precise characterization of this new navigation mode. It will also provide us with a set of metrics that can be used to evaluate the performance improvements that result from new developments in head-mounted display technology. I believe that this direction will profitably use the time spent waiting for technology developments by developing an intellectual framework in which to measure the developments.

The user study revealed a set of constituent skills that are called upon at various times in the targeting of treatment beams. These skills may be affected by the navigation mode and may lend themselves more readily to objective measurement than did the entire beam targeting task. I suggest that future research attempt to characterize the accuracy and speed with which judgements pertinent to beam targeting can be made. Specifically, the ability to judge the *volume of intersection* between a beam and an irregularly-shaped object such as a lung could be studied. In this, accurate relative judgement of volume (beam A intersects more lung than beam B) would be more relevant than absolute judgement. Similar to volume judgement is the determination of *beam length*—which beam has a longer traversal through the body? Estimation of the *angle* between a pair of treatment beams is another important skill that may be affected by navigation mode, because some modes may build more accurate understanding of the situation or more easily facilitate movement to a normal direction from which the angle can be accurately seen. For some subjects in the user study it was important to know where the patient's *cardinal directions* were, and navigation mode may have an effect on the ability to find the cardinal directions.

The observational data also revealed more basic performance parameters that are relevant to other domains outside of beam targeting. These parameters could not be accurately and completely described with the data collected here, but more focused investigation should yield interesting conclusions. One of these parameters is the *coverage* of the solution space. If the subject were somehow required to explore all portions of the solution sphere, would the subject's movement show preference for certain regions over others? Would this preference vary among navigation modes? Another issue is *movement direction*. Is the movement omnidirectional, or are there preferred directions? Are these preferences dictated by the navigation mode or are they rooted in human perception and motoric ability? It would also be interesting to understand the *precision, speed, and accuracy* with which movements can be made using different navigation modes. A Fitts' law type of study investigating the *index of difficulty* of movement and the *bandwidth* (MacKenzie 1992) provided by various navigation modes would be appropriate here, involving large- and small-scale movement. Such a study may produce insight into the similarity between the speed histograms of the two modes presented in Figure 7-20. In addition, the bandwidth of the navigation mode is a convenient metric by which developments in technology can be evaluated. It would also be helpful to understand how well subjects can stay still with a given navigation mode, for there has been evidence that this is somewhat difficult with Orbital mode. The accuracy of *muscle memory* is also of interest, because one of the claims of Orbital mode is that enables users to take advantage of proprioception to maintain their bearings. This raises the question of how accurately can one turn his head to a direction that he had visited before. Does the accuracy decrease with time?

Although the preliminary study showed velocity-control Joystick mode to be best suited for beam targeting of the three non-head-tracked modes tested, there are other devices that may be even better suited and these should be investigated. I feel the foremost candidate for further study is the positional joystick, which would provide much the same benefit as Orbital mode. A positional joystick would provide one-to-one correspondence between its state and the beam's-eye view seen by the user, thereby simplifying navigation as Orbital mode did. Users of the positional joystick, however, would not suffer from the instability experienced by Orbital mode users.

10.2.2. Beam targeting study

When head-mounted display technology has improved sufficiently, it may be fruitful to repeat the beam targeting study. The following modifications to the experimental method may help reduce the variability in the data and increase the power of the experiment. Complete, detailed case histories should be provided to the subjects so that they know as much about the test case as they do their real patients. This should eliminate the need to make assumptions regarding the case that might affect the treatment plan. To avoid the confounding influence of the Practicality effect, the experimental task should be clearly stated as the design of the beam configuration that would produce the optimal dose distribution for the case. Clinical considerations of patient time and error risk should be disregarded. The case history can be constructed to support this goal by stressing the need for heroic curative therapy, regardless of cost. (Perhaps the patient could be the President of the United States.) To equalize conditions for evaluation of the beam configurations, judges should be trained to some appropriate level of competence on the evaluation tool. Also, dosimetry should be incorporated into the beam configuration to give the judges as much information as possible. The best way to do this may be to have the judges add beam modifiers to each configuration they critique and then to compute the dose distribution for that setup. Then, the judges will have dosimetric information with which to make decisions, and the dosimetry will reflect only differences in the beam configurations. It may also be useful to choose a more challenging tumor site for the experiment. The more complex situations surrounding prostate tumors might be more sensitive to navigation modes than the lung tumors used in here.

Observation of the subjects has suggested to me that the effectiveness of new tools and user interfaces may be hindered by strategies and techniques that were developed based upon the old tools. To take full advantage of the benefits afforded by new tools, or even to evaluate what those benefits might be, it may be necessary to retrain and re-educate the users. This, of course, would require an enormous effort; before the radiation oncology community can be persuaded to change its approach, there must be solid evidence that that such an effort would be worthwhile. That evidence has yet proven too difficult to be gathered, but I believe the constant advance of head-mounted display technology will eventually make it possible.

10.3. Contribution

At present, the corpus of scientific findings and observations regarding human performance in virtual environments is small, but growing. Below I list and discuss what I feel are the significant contributions to this field produced by this research.

Invention of Orbital mode — a new, intuitive navigation mode for movement about a spherical surface. Orbital mode lacked a handy metaphor, but provided advantages of intuitive control, beam's-eye view, and hands-free operation.

Effectiveness demonstration of Orbital mode. A basic human performance study was conducted comparing with Orbital mode with three other intuitive, head-tracked navigation modes (Walkaround, Walk/Rotate, Immersion) and with three less intuitive, non-head-tracked modes (Joystick, Spaceball, Mouse). This study yielded the following contributions:

- *Navigation modes exhibit statistically significant differences in performance.*

- *Navigation modes exhibit qualitative differences.* Such differences were evident in subject rankings of navigation modes according to preference and ease-of-use, and in subjective comments.

Radiation Oncology study. The results of the preliminary study were used in the design of a radiation oncology user study comparing Orbital mode (head-tracked) with Joystick mode (non-head-tracked) in the targeting of radiation treatment beams. The contributions of this study included:

- *Orbital mode is faster than Joystick mode.* Orbital mode users averaged a beam-targeting rate of roughly 20 cases/hr., whereas Joystick mode users averaged 17 cases/hr. This result is especially important because of the demanding schedules found in radiation oncology clinics today.
- *Joystick mode produces better results.* As measured by individual criterion scores, no significant differences in performance quality were measured. The Overall scores, however, revealed Joystick mode to be a better performer than Orbital mode.
- *The navigation modes have many qualitative differences.* Observational data was analyzed for differences and similarities between subject performance and behavior in the two navigation modes. Differences found between navigation modes include profiles of movement direction, accuracy and precision of control, and ergonomic considerations. The performance observations can serve as the foundation for more rigorous scientific investigation. The current state of head-mounted display technology is inadequate for clinical use, but future developments may change that.
- *Unforeseen aspects of task have strong effects on results.* The tasks of targeting and evaluating radiation treatment beam configurations were analyzed. Several unforeseen aspects were identified (e.g. Practicality effect) and their confounding influences on the experimental results were discussed.

References

Badcock, P.C. (1984). The value of computed tomography in planning radiotherapy of intra-thoracic tumors. *Strahlenther. Onkol.*, 160:26.

Bentel, G.C., C.E. Nelson and K.T. Noell (1989). Treatment Planning and Dose Calculation in Radiation Oncology (4th ed.). Pergamon Press: New York, 350 pp.

Bourne, L.E., Jr., R.L. Dominowski, E.F. Loftus and A.F. Healy (1986). Cognitive Processes (2nd ed.). Prentice-Hall: Englewood Cliffs, New Jersey, 404 pp.

Brooks, F.P., Jr. (1986). Walkthrough—A dynamic graphics system for simulating virtual buildings. In Proceedings of *1986 Workshop on Interactive 3D Graphics*, F. Crow and S.M. Pizer (Ed.), 23-24 Oct '86, Chapel Hill, North Carolina, ACM: New York, 9-21.

Brooks, F.P., Jr. (1988). Grasping reality through illusion—Interactive graphics serving science. In Proceedings of *CHI'88*, E. Soloway, D. Frye and S.B. Sheppard (Ed.), 15-19 May '88, Washington, D.C., ACM: New York, 1-11.

Chin, L.M., R.L. Siddon, G.K. Svensson and C. Rose (1985). Progress in 3-D Treatment Planning for Photon Beam Therapy. *International Journal of Radiation Oncology, Biology, and Physics*, 11:2011-2020.

Chung, J.C. (1992). A Comparison of Head-tracked and Non-head-tracked Steering Modes in the Targeting of Radiotherapy Treatment Beams. In Proceedings of *1992 Symposium on Interactive 3D Graphics*, April 1992, Cambridge, Mass., ACM: New York, 193-196.

Clark, J.H. (1976). Designing Surfaces in 3-D. *Communications of the ACM*, 19(8):454-460.

DiZio, P. and J.R. Lackner (1992). Spatial Orientation, Adaptation, and Motion Sickness in Real and Virtual Environments. *Presence*, 1(3):319-328.

Dodd, M.J. (1987). Managing Side Effects of Chemotherapy and Radiation Therapy. Appleton and Lange: Norwalk, Conn., 156 pp.

Ebenholtz, S.M. (1992). Motion Sickness and Oculomotor Systems in Virtual Environments. *Presence*, 1(3):302-305.

Ellis, S.R. and A.J. Grunwald (1989). Head-mounted spatial instruments II: Synthetic reality or impossible dream. In Proceedings of *NASA Conference on Space Telerobotics*, 31 Jan - 2 Feb '89, JPL, 1-16.

Ellis, S.R., S. Smith, A. Grunwald and M.W. McGreevy (1991). Direction judgement error in computer generated displays and actual scenes. In *Pictorial communication in virtual and real environments*, S.R. Ellis (Ed.), Taylor & Francis Ltd.: London, 504-526.

Eyles, J., J. Austin, H. Fuchs, T. Greer, and J. Poulton (1988). Pixel-planes 4: A Summary. In *Advances in Computer Graphics Hardware*, Eurographics Seminars, 1988, 183-208.

Fuchs, H., J. Poulton, J. Eyles, T. Greer, J. Goldfeather, D. Ellsworth, S. Molnar, G. Turk, B. Tebbs and L. Israel (1989). Pixel-Planes 5: A Heterogeneous Multiprocessor Graphics System Using Processor-Enhanced Memories. *Computer Graphics*, 23:79-88.

Gerbi, B.J. (1992). Use of the simulator in treatment planning and determination and definition of treatment volume. In *Levitt and Tapley's Technological Basis of Radiation Therapy: Practical Clinical Applications*, S.H. Levitt, F.M. Khan and R.A. Potish (Ed.), Lea & Feiberger: Philadelphia, 27-37.

Goitein, M. (1979a). Computed tomography in planning radiation therapy. *International Journal of Radiation Oncology, Biology, and Physics*, 5:445.

Goitein, M. (1979b). The utility of computed tomography in radiation therapy: An estimate of outcome. *International Journal of Radiation Oncology, Biology, and Physics*, 5:1799.

Goitein, M. (1980). Benefits and costs of computerized tomography in radiation therapy. *Journal of the American Medical Association*, 244:1347.

Goitein, M. (1982). Limitations of two-dimensional treatment planning systems. *Medical Physics*, 3:580.

Goitein, M. and M. Abrams (1983). Multi-dimensional Treatment Planning: I. Delineation of Anatomy. *International Journal of Radiation Oncology, Biology, and Physics*, 9:777-787.

Goitein, M., M. Abrams, D. Rowell, H. Pollari and J. Wiles (1983). Multi-dimensional Treatment Planning: II. Beam's-eye View, Back Projection, and Projection Through CT Sections. *International Journal of Radiation Oncology, Biology, and Physics*, 9:789-797.

Goitein, M. and T.E. Schultheiss (1985). Strategies for treating possible tumor extension: Some theoretical considerations. *International Journal of Radiation Oncology, Biology, and Physics*, 11:1519-1528.

Gregory, R.L. (1991). Seeing by exploring. In *Pictorial communication in virtual and real environments*, S.R. Ellis (Ed.), Taylor & Francis Ltd.: London, 328-337.

Hall, E.J. and J.D. Cox (1989). Physical and biologic basis of radiation therapy. In *Radiation Oncology: Rationale, Technique, Results*, W.T. Moss and J.D. Cox (Ed.), C.V. Mosby Co.: St. Louis, 1-57.

Henry, D. and T. Furness (1993). Spatial perception in Virtual Environments: Evaluating an Architectural Application. In Proceedings of *VRAIS '93*, 18-22 Sep '93, Seattle, Wash., IEEE: Piscataway, New Jersey, 33-40.

Hettinger, L.J. and G.E. Riccio (1992). Visually Induced Motion Sickness in Virtual Environments. *Presence*, 1(3):306-310.

Holloway, R.L. (1993). A Systems-Engineering Study of the Registration Errors in a Virtual-Environment System for Cranio-Facial Surgery Planning. Ph.D. dissertation research proposal, Department of Computer Science, University of North Carolina at Chapel Hill.

Holloway, R.L. and A. Lastra (1993). Virtual Environments: A Survey of the Technology (TR93-033). Dept. of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, 55 pp.

Howlett, E. (1983). Wide Angle Color Photography Method and System. U.S. Patent.

Jagacinski, R.J. and D.L.. Monk (1985). Fitts' Law in Two Dimensions with Hand and Head Movements. *Journal of Motor Behavior*, 17(1):77-95.

Kaiser, M.K. (1991). Knowing. In *Pictorial communication in virtual and real environments*, S.R. Ellis (Ed.), Taylor & Francis Ltd.: London, 43-46.

Kennedy, R.S., N.E. Lane, M.G. Lilienthal, K.S. Berbaum and L.J. Hettinger (1992). Profile Analysis of Simulator Sickness Symptoms: Application to Virtual Environment Systems. *Presence*, 1(3):295-301.

Kilpatrick, P.J.. (1976). The Use of a Kinesthetic Supplement in an Interactive Graphics System. Ph.D. dissertation, Department of Computer Science, University of North Carolina at Chapel Hill, 175 pp.

Kutcher, G.J. and C. Burman (1989). Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method. *International Journal of Radiation Oncology, Biology, and Physics*, 16:1623-1630.

Ling, C.C., C.C. Rogers and R.J. Morton (1983). Computed Tomography in Radiation Therapy. Raven Press: New York, 268 pp.

Linksz, A. (1952). Physiology of the Eye, Vol. 2: Vision. Grune & Stratton: New York.

Lyman, J.T. (1985). Complication probability as assessed from dose-volume histograms. *Radiation Research*, 104:S13-S19.

MacKenzie, I.S. (1992). Fitts' Law as a Research and Design Tool in Human-Computer Interaction. *Human-Computer Interaction*, 7:91-139.

McCauley, M.E. and T.J. Sharkey (1992). Cybersickness: Perception of Self-Motion in Virtual Environments. *Presence*, 1(3):311-318.

McGee, M.G. (1979). Human Spatial Abilities. Praeger: New York, 146 pp.

McKinnon, G.M. and R.V. Kruk (1991). Multi-axis control in telemanipulation and vehicle guidance. In *Pictorial Communication in Virtual and Real Environments*, S.R. Ellis (Ed.), Taylor & Francis Ltd.: London, 247-264.

Mine, M.R. (1993). Characterization of End-to-End Delays in Head-Mounted Display Systems (TR93-001). Dept. of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, 11 pp.

Mosher, C.E., G.W. Sherouse, E.L. Chaney and J.G. Rosenman (1988). 3D Displays and User Interface Design for a Radiation Therapy Treatment Planning CAD Tool. In Proceedings of *Three-Dimensional Imaging and Remote Sensing Imaging*, W.E. Robbins (Ed.), 14-15 January 1988, Los Angeles, SPIE: Bellingham, WA, 902:64-71.

Norman, D.A. (1988). The Psychology of Everyday Things . Basic Books: New York, 257 pp.

Pausch, R., M.A. Shackelford and D. Proffitt (1993) A User Study Comparing Head-Mounted and Stationary Displays. In Proceedings of *IEEE Symposium on Research Frontiers in Virtual Reality*, 25-26 October 1993, San Jose, Calif.

Pausch, R., J.C. Chung, R. Eggleston and T. Piantanida (1993). Evaluation of VR Systems: Experience and Advice. In Proceedings of *IEEE Symposium on Research Frontiers in Virtual Reality*, 25-26 October 1993, San Jose, Calif.

Perez, C.A. and J.A. Purdy (1992). Rationale for treatment planning in radiation therapy. In *Levitt and Tapley's Technological Basis of Radiation Therapy: Practical Clinical Applications*, S.H. Levitt, F.M. Khan and R.A. Potish (Ed.), Lea & Feiberger: Philadelphia, PA, 14-26.

Perrone, J.A. and P. Wenderoth (1991). Visual slant underestimation. In *Pictorial communication in virtual and real environments*, S.R. Ellis (Ed.), Taylor & Francis Ltd.: London, 496-503.

Proffitt, D.R. and M.K. Kaiser (1991). Perceiving environmental properties from motion information: minimal conditions. In *Pictorial communication in virtual and real environments*, S.R. Ellis (Ed.), Taylor & Francis Ltd.: London, 47-60.

Purdy, J.A. and B. Emami (1992). Computed Tomography and Three-Dimensional Approaches to Radiation Therapy Treatment Planning. In *Levitt and Tapley's Technological Basis of Radiation Therapy: Practical Clinical Applications*, S.H. Levitt, F.M. Khan and R.A. Potish (Ed.), Lea & Feiberger: Philadelphia, 56-66.

Purdy, J.A., J.W. Wong, W.B. Harms, B. Emami and J.W. Matthews (1987). State of the art of high energy photon treatment planning. In *Treatment Planning in the Radiation Therapy of Cancer*, J.M. Vaeth and J. Meyer (Ed.), Karger: Basel, Switzerland, 4-24.

Raven, R.W. (1990). The Theory and Practice of Oncology . Parthenon Publishing Group: Carnforth, 366 pp.

Robinett, W. and J.P. Rolland (1992). A computational model for the stereoscopic optics of a head-mounted display. *Presence*, 1(1):45-62.

Rosenman, J.G. (1991). 3D Imaging in Radiotherapy Treatment Planning. In *3-D Imaging in Medicine*, J. Udupa and G. Herman (Ed.), Lewis Publishers: Chelsea, Michigan.

Rosenman, J.G., E.L. Chaney, S. Sailer, G.W. Sherouse and J.E. Tepper (1991). Recent advances in radiotherapy treatment planning. *Cancer Investigation*, 9:465-481.

Rosenman, J.G., G.W. Sherouse, H. Fuchs, S.M. Pizer, A.L. Skinner, C.E. Mosher Jr., K. Novins and E.L. Chaney (1989). Three-dimensional display techniques in radiation therapy treatment planning. *International Journal of Radiation Oncology, Biology, and Physics*, 16:263-269.

Rottenberg, D.A. (1991). Acute and chronic effects of radiation therapy on the nervous system. In *Neurological Complications of Cancer Treatment*, D.A. Rottenberg (Ed.), Butterworth-Heinemann: Boston, 228.

Rushton, S., and J. Wann (1993). Problems in Perception and Action in Virtual Worlds. In Proceedings of *Virtual Reality 1993: 3rd Annual Conference on Virtual Reality*, London, Antony Rowe: Wilts, 43-55.

Sailer, S.L., E.L. Chaney, J.G. Rosenman, G.W. Sherouse and J.E. Tepper (1992). Treatment planning at the University of North Carolina at Chapel Hill. *Seminars in Radiation Oncology*, 2(4):267-273.

Sutherland, I.E. (1965). The Ultimate Display. In Proceedings of *IFIP Congress 65*, 506-508.

Sutherland, I.E. (1968). A head-mounted three dimensional display. In Proceedings of *1968 Fall Joint Computer Conference*, 33:757-764, AFIPS Press: Montvale, New Jersey.

Tremewan, R.N. (1988). Computed tomography in radiotherapy treatment planning. *Australas Radiol*, 32:50.

Tharp, G., A. Liu, L.French, S. Lai and L. Stark (1992). Timing Considerations of Helmet Mounted Display Performance. In Proceedings of *Human Vision, Visual Processing and Digital Display*, February 1992, San Jose, Calif., SPIE: Bellingham, WA.

Ware, C. and S. Osborne (1990). Exploration and virtual camera control in virtual three dimensional environment. In Proceedings of *1990 Symposium on Interactive 3D Graphics*, 25-28 Mar '90, Snowbird, Utah, ACM: New York, 175-183.

Ware, C. and L. Slipp (1991). Exploring virtual environments using velocity control: A comparison of three devices. In Proceedings of *35th Annual Meeting of the Human Factors Society*, Sep'91, San Francisco, 300-304.

Woolf, H.B. (1975). Webster's New Collegiate Dictionary G. & C. Merriam Company: Springfield, Mass., 1536 pp.

Zeltzer, D. (1992). Autonomy, interaction, and presence. *Presence*, 1(1):127-132.

Appendix: Source Code