FACULTY OF SCIENCE & HEALTH
SCHOOL OF HEALTH & CARE PROFESSIONS

The role of an immersive 3-D virtual reality environment in the development of the spatial visualisation skill of pre-registration therapeutic radiography students

ANDREW JAMES WILLIAMS

The thesis is submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of the University of Portsmouth

May 2020

Abstract

Background

The treatment of cancer with radiation uses advanced techniques such as intensity modulated and stereotactic radiotherapy. These modalities can provide sub millimetre accuracy, delivering high radiation dose to the tumour and reduced dose to the surrounding normal anatomy. Building a mental model of the size, shape and position of the tumour in relation to the surrounding anatomy and proposed radiation beam direction requires radiotherapy radiographers to have well developed three- spatial visualisation skills. The introduction of the Virtual Environment for Radiotherapy Training (Vertual Ltd, Hull, United Kingdom), an immersive 3-D visualisation platform in 2007, offered the potential for developing new ways of supporting the development of these visualisation skills in pre-registration radiotherapy learners in a simulated environment.

Aims

This programme of research aimed to measure the baseline three dimensional spatial visualisation skills of new radiotherapy learners, to compare their performance with new learners in diagnostic imaging, to determine if 3-D visualisation skills could develop over time and to identify those learners most likely to benefit from learning in the virtual environment for radiotherapy training.

Methods

This programme of research employed a QUANT + qual mixed model approach with purposive convenience sampling of year one diagnostic imaging and radiotherapy students to develop an online, three-dimensional spatial visualisation test using objects from two traditional paper based spatial visualisation tests for mental rotation and cross sectional visual perception (the pilot phase). The experimental phase employed an online test platform in a controlled, single subject design, longitudinal study using a second cohort of students to determine their 3-D spatial visualisation skill at the start of their pre-registration radiography programme and to track any change over three additional time points during an 18-month period between the start of year one and the end of year two. For the radiotherapy cohort, the relationship between baseline spatial visualisation

skill and patient positioning performance was investigated using a simulated treatment delivery task conducted within the Virtual Environment for Radiotherapy Training.

Results

The pilot phase comparison of performance scores for the paper based and online versions of the three-dimensional spatial visualisation test did not produce statistically significant differences, suggesting that a move to an online test platform would not disadvantage any participant. Results from the experimental phase (study four) identified that the baseline 3-D spatial visualisation skill of 54 pre-registration learners in radiotherapy (n = 15) and diagnostic imaging (n = 39) could be measured and performance classified as being low, intermediate or high at the start of their radiography education. Across both pathways, 13 participants (24%) were identified as having low skill, 36 (67%) were intermediate and 5 (9%). Performance gains were observed in the growth trajectory for mean spatial visualisation test score over the 18 month time period for both pathways. Analysis of performance in the mental rotation and cross section subcomponents indicated that one third of all participants might benefit from additional support in mental rotation or perception of cross sections. For the radiotherapy positioning task, correlations between task performance metrics for task completion time, number of equipment adjustments and set up accuracy and baseline demonstrated a weak positive relationship meaning that the results were inconclusive.

Conclusions and contributions to knowledge

The measurement of 3-D spatial visualisation subcomponent performance as a surrogate for accurate patient positioning and beam alignment has provided an enhanced understanding of baseline visualisation skill. Analysis of these subcomponents, with an emphasis on patterns of incorrect answers, in addition to overall performance score provides a method for identifying those individuals with less well-developed skills. These are the learners who may have difficulty with mental model and relationship building and would benefit from the additional support of focused tutorials with personalised spatial visualisation syllabus activities. This enhanced understanding will provide opportunities for the development of the spatial visualisation syllabus beyond the often opportunistic, and ad-hoc structure of clinical practice and a one-size fits all approach to campus based simulation and visualisation activities.

Contents

Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

Word count: 83,770

List of Tables

Table of Figures

Abbreviations

3-D CRT	3-dimensional conformal radiotherapy
3-D SVT	3-dimensional spatial visualisation test
CBCT	Cone beam computed (or computerised) tomography
CT	Computed tomography
CTV	Clinical target volume
GTV	Gross tumour volume
Gy	Gray
HEI	Higher Education Institute
IGRT	Image guided radiotherapy
IMRT	Intensity modulated radiotherapy
MRT	Mental rotation test
OEM	Original equipment manufacturer
PET	Positron emission tomography
PTV	Planning target volume
SBRT	Stereotactic body radiotherapy
SBST	Santa Barbara solids test
STEM	Science, technology engineering and maths
SVS	3-dimensional spatial visualisation skill
VMAT	Volumetric arc therapy

Glossary

Dose rate:

The amount of radiation dose absorbed per unit of time

Gray:

The amount of energy transferred from the beam of radiation to human cells and tissues (1Gray = 1 joule of energy delivered to 1 kilogram of tissue)

Half-life:

The time required for a radioactive substance to decay to one half of its original strength

Image guided radiotherapy:

The process of verifying correct anatomical and tumour position prior to delivery of treatment

Intensity modulated radiotherapy:

The delivery of two or more different sized radiotherapy beams, of different shape from a single gantry position to deliver dose of varying intensity

Percentage depth dose:

The proportion (expressed as a percentage) of the absorbed dose at a specified depth (d) compared to the absorbed dose at a fixed reference depth (d 0) along the central axis of the beam

Positron emission tomography:

A diagnostic imaging technique to measure physiological function at a molecular level

Specific activity:

The amount of radioactivity or decay rate of a radioactive substance per unit

Stereotactic body radiotherapy:

The delivery of radiotherapy from a number of small fields and high dose with sub millimetre accuracy

Tumour target volume:

The tumour target volume consists of the visible or clinically detectable tumour - the gross tumour volume, plus a margin to accommodate undetectable microscopic spread – the clinical target volume plus a margin to account for movement because of breathing – the planning target

Volumetric arc therapy:

A type of intensity-modulated radiotherapy but with simultaneous movement of the gantry

Acknowledgements

Many people have supported me through this journey of learning, research and discovery. Much of what I have achieved since becoming a full time academic in 2006 would not have been possible without one man. When I completed my teaching in higher education qualification, Derek Adrian- Harris asked me “what next?” and planted the PhD seed, thank you for seeing something in me that I could not see at the time.

My first Director of Studies, Dr Sally Kilburn, set me off on this journey and continually challenged me to think about what I needed to achieve. When Sally moved on to pastures new, Dr Jason Oakley took over as Director and brought a new focus and impetus to my work, for which I will always be thankful.

Along the way, I have always been able to count on the support of my second supervisors; to Dr Ann Dewey I say a very big thank you for persevering, for being forthright and telling me the things that I did not always want to hear. Dr Mick Harper, thank you, not just for the continuous feedback, but also for our wide-ranging discussions relating to simulation and education in general. To Drs Penny Delf and Rebecca Stores, your insight, comments and recommendations along the way have been invaluable. While my clinical life was spent manipulating data for treatment planning and descriptive analysis of audit results, I am indebted to Dr Ngianga II Kandala (Kandala) for his sage advice on the wider aspects of statistical analysis.

I would also like to express my gratitude to all the members of the Radiography course team, past and present, who have offered me help and support from the outset. However, it would not have been possible to complete this research had it not been for the skills of Sarah Cooper who was instrumental in building the online survey tools and test instruments. To all the study participants from the 2010 and 2013 Diagnostic Imaging and Radiotherapy cohorts who willingly gave up their time to have their spatial visualisation skill measured – thank you. You are our future, go and make it yours!

And finally, to four very special people, thank you Dr Tom Williams for providing an external focus on early drafts and for finding more corrections in my thesis than I was able to find in yours and to Matthew, for just being around. To Ryan, thank you for providing me with an excuse to play with cars and trains on Saturday mornings when I really should have been more focussed on writing and to Lindsey – you have always

believed that this was possible. Thank you for giving me the space, time and encouragement to continue, especially when I was ready to give up and for not reminding me too many times as the list of DIY jobs around the house got longer.

List of dissemination outputs from this research

Peer reviewed publications:

Bridge, P., Giles, E., Williams, A., Bøjen, A., Appleyard, R., Kirby, M. (2017). International Audit of VERT™ Academic Practice, *Journal of Radiotherapy in Practice*, 6(4), 375-382.

Conference presentations:

Williams, A.J. (2011, March). Learning in the virtual world: what do we need to know? Paper presented at the 8th Annual Toronto Rti3 Radiation Therapy Conference

Williams, A.J. (2012, June). Can learning in virtual reality environments enhance visualisation skills? Paper presented at the 17th International Society of Radiographers and Radiological Technologists Conference.

Harper, M. & Williams, A.J. (2012, June). Clinical learning & assessment in simulated & virtual worlds. Paper presented at the 17th International Society of Radiographers and Radiological Technologists Conference.

Williams, A.J. (2016, July). Measuring the spatial visualisation skill of student radiographers. Paper presented at the University of Portsmouth Science Together Post-Graduate Research Conference.

Invited conference presentations:

Williams, A.J. (2015, September). The theory and practice of using VERT™ as an assessment tool. Paper delivered at the International VERT™ User Group Meeting, Portsmouth, UK.

Williams, A.J. (2017, June). The Virtual Environment for Radiotherapy Training (VERT™) & Spatial Visualisation Skill. Paper delivered at the International VERT™ User Group Meeting, Haarlem, Netherlands.

Poster presentations

July 2010: University of Portsmouth Science Faculty Research Forum

“The virtual world of clinical assessment in radiotherapy”

November 2010: University of Portsmouth Health Development Forum

“Using a 3D virtual environment for the acquisition & assessment of clinical skills in radiotherapy”

July 2011: University of Portsmouth Science Faculty Research Forum

Determining Spatial Visualisation Ability & Learning Style: A Radiographic Perspective

Chapter 1

Background and introduction to this programme of research

1.1 Introduction to chapter one

In order to set the contextual background for this programme of research, the chapter will begin by tracing the history of cancer management strategies for England over the past two decades. It will then identify the physiological and pathological processes involved in the development of cancer before reporting the current position and future predictions for cancer incidence in England. It will continue by outlining the recommendations that have been made for the delivery of a world class service for cancer patients with an emphasis on those relating to radiotherapy. In doing so, it will discuss the implications of these recommendations for the radiotherapy workforce and in particular the impact on radiotherapy radiographer pre-registration education. It will conclude by providing an overview of the aims and objectives for this programme of research.

1.2 Background to this programme of research

The publication of the National Health Service (NHS) Cancer Plan in 2000 heralded the first coordinated 10-year strategy for the improvement of cancer services in England. The stated aims of the plan were: to save more lives by reducing inequalities in health, to invest in the cancer workforce and research and to ensure that patients would get the best treatment and the right professional care and support (National Health Service Executive, 2000, p. 5). A progress report by the House of Commons Public Accounts Committee (2006, p. 3) noted that, while encouraging progress had been made, further review was required. Aspects of that review included the National Radiotherapy Advisory Group (NRAG) report to Ministers on the development of world class radiotherapy services in England. This report identified that the projected need for radiotherapy had been underestimated during the previous two decades. If radiotherapy were to be given to all patients who might benefit from it the National Radiotherapy Advisory Group (2007, p. 3) reported that the access rate was 63% lower than the optimal requirement for treatment To address this shortfall, the report identified that the radiotherapy service in England needed to increase the level of access to radiotherapy treatment. Increasing access would require investment in additional equipment and an associated increase in the radiotherapy radiographer workforce by up to 50% (NRAG, 2007, p. 6). To support these increases, the report identified the need to provide a quality training experience for pre-registration learners while at the same time improving retention and reducing the training burden on clinical departments. The report recommended the introduction of hybrid virtual environment skills training facilities across the 10 educational providers and 52 clinical radiotherapy departments in England from 2007 (ibid, p. 25). This recommendation resulted in the Department of Health for England allocating five million pounds of capital funding for virtual reality radiotherapy training platforms (Department of Health, 2007, p. 61). The virtual environment for radiotherapy training (VERT™), as it became known, employs 2-D and 3-D projection to replicate real radiotherapy linear accelerators virtually. The virtual linear accelerator can replicate the full range of mechanical movements found on in-service machines. This is achieved through user interaction with original equipment manufacturer (OEM) hand pendants to control the virtual linear accelerator models available in VERT™. Although the platform does not produce X-ray beams, their position and path can be visualised on a range of virtual treatment delivery plans which are based on real but anonymised patient data (referred to as the virtual patient). Following the roll out of funding and the installation of the platform during 2008, the Department of Health (England) commissioned an 18 month evaluation project led by the Society and College of Radiographers, to manage the implementation of the VERT™ technology and evaluate its impact on training. The aims of that project as identified by Appleyard & Coleman (2010, p. 2) were to assess the potential use of VERT™ and its impact on the student learning experience and in the development of practical skills, confidence and knowledge. The findings of one of the studies conducted by Appleyard and Coleman (2010, p. 33) as part of the project are discussed in more detail in chapter 2.8, p. 51.

The NRAG report (2007, p. 29) also identified the need to increase the introduction of new technology to support the delivery of intensity modulated radiotherapy (IMRT) and 4-D adaptive techniques which allow the treatment position and dose delivered to be verified and changed as necessary during a course of treatment. The delivery of these advanced and complex radiotherapy techniques, discussed in chapter 2.3, pp. 29 – 32, have increasingly encroached on clinical resources as suggested by Bridge et al., (2007, p. 482) and have added to the time pressures on departmental workload and workflows. While there is limited data reporting of radiotherapy workload and capacity issues in the literature, a small study of 324 randomly selected radiotherapy treatment sessions conducted by Van de Werf, Lievens, Verstraete, Pauwels and Van den Bogaert (2009, p.138) reported that the mean room occupancy time (the time from the patient entering the room to the time the patient leaves the room) was reported as 11.6 minutes (SD = 5.9) for conventional 3-D conformal radiotherapy (3-D CRT) delivery. For more complex treatment delivery, such as a seven field IMRT technique, this time increased to 13.6 minutes (SD = 5.4) and with the addition of IGRT the mean was 17.3 minutes (SD = 6.8). The need to increase access to radiotherapy as a result of increasing incidence and the longer time required to deliver more complex treatments is likely to add to perceived time pressures in clinical departments. It is also possible that these pressures will mean that students may often feel rushed and unable to maximise their learning of these more demanding techniques. This likely to have an adverse impact on clinical experience and may lead to some students not completing their studies. These issues of demand, access and staffing remain pertinent today, with the Health Education England Cancer Workforce Plan (Health Education England, 2017, p. 37) reporting that NHS Trusts in England had 2632 funded radiotherapy radiographer posts in 2016 and would require in the region of 300 additional posts by 2021. Associated with this requirement, but outside the scope of this programme of research, is a reported drop of 27.5% in UCAS applications to Subjects Allied to Medicine which includes radiotherapy programmes with 81,720 fewer applicants from England during the period 2014-2018 (Health Education England, 2018, p. 9). The Reducing Pre-registration Attrition and Improving Retention project (RePAIR) report also identified that between 2013 and 2015, 33% of students did not complete their programme of radiotherapy education (Health Education England 2018, p. 26). Supporting research for the reasons for poor retention in radiotherapy is limited, however, evidence from nursing indicates that they are multi-faceted (Jeffreys, 2015, p. 426) and identified environmental and professional integration factors and the interaction between psychological and academic outcomes. In addition, the RePAIR project (Health Education England, 2018, p. 12) suggested a number of reasons for non-completion including personal reasons, poor placement learning experience and assessment failure. Given the complexity of modern radiotherapy techniques and workload time pressures, it is possible that some students may experience difficulty with the visualisation and processing of complex information relating to the application of radiotherapy in cancer management. If this is the case, it may contribute to a poorer clinical experience and performance. The use of additional visualisation opportunities such as those provided by VERT™ may assist the development of 3-D spatial visualisation skills in these learners.

1.3 The physiological and pathological development of cancer

The division and reproduction of cells in normal tissues is tightly regulated and results in the accurate duplication of the human genome into two daughter cells (Shen, 2011, p. 1). Under these conditions the number of new cells produced is equal to the number of mature differentiated cells that are lost through damage and death. If the control mechanism for regulated cell division, growth and death are disrupted then uncontrolled cell division occurs and may result in the development of a benign (non-invasive) or malignant (invasive cancerous) primary tumour within a specific organ.

Invasive cancers develop because of the accumulation of genomic alterations in cell structure combined with deoxyribose nucleic acid (DNA) instability over time (Orth, et al., 2014, p. 5) and have the potential to invade adjacent tissues in the same organ and spread to adjacent organs. These invading cells can also migrate into lymphatic vessels and lymph nodes and have the capacity to survive in the circulating blood supply which in turn can give rise to their survival in distant organs (Massagué & Obenauf, 2016, p. 298). This is referred to as metastatic or secondary disease and occurs as a result of the downregulation of the damage surveillance mechanisms of the immune system. This downregulation occurs as a result of increased genetic instability of cells as described by Jeggo, Pearl and Carr, (2016, p. 35). Curative treatment management interventions are designed to remove the tumour completely or inhibit cell division while the tumour is still localised within its organ of origin. Patients with metastatic tumour deposits distributed across several anatomical sites are unlikely to be cured of their cancer, but with appropriate treatment interventions could survive symptom free for relatively long periods of time between each episode of tumour growth and spread.

1.4 Cancer incidence in England

The data for cancer registrations collected by the National Cancer Registration and Analysis Service, a division of Public Health England shows that the incidence has continued to rise over recent decades. The latest incidence data for cancer registrations in England indicate that in 2017 there were 305,683 new cancer diagnoses, excluding non-melanoma skin cancer (Office for National Statistics, 2019, p. 3). Recent estimates by Smittenaar, Petersen, Stewart and Moitt (2016, p.1149) predicted that the incidence of new cancer cases is likely to rise by 42.5% by 2035. While primary cancers can arise in any cell type, in any organ, Bray et al., (2018, p. 395) identified 36 major cancer types arising worldwide. In England, figures from the Office for National Statistics (2018, p. 3) demonstrate that the four most common cancers in males and females, make up 60.5% of cases in males and 63.3% of cases in females. The percentage breakdown of each of these sites in comparison to all the other sites combined is summarised in table 1.1.

Table 1.1: The number of cancer diagnoses registered in England in 2017 by most frequently occurring anatomical site

Male		Female	
Prostate	39.5%	Breast	30.7%
Lung	13%	Lung	12.4%
Colorectal	12.4%	Colorectal	10.3%
Non-Hodgkin's Lymphoma	4.3%	Uterus	5.3%
Other	39.5%	Other	36.7%

Recent demand modelling has indicated that radiotherapy can be of benefit in the region of 45 - 50% of patients (Round et al., 2013, p. 529; Independent Cancer Taskforce, 2015, p. 35). If the incidence of new cancer cases increases as predicted, this will result in approximately 514,000 new cases diagnosed in England by 2035. This would result in an increase in the demand for radiotherapy in the region of 21,150 additional cases per year. In addition, estimates from the European Society of Radiation Oncology Health Economics in Radiation Oncology project reported by Borras et al., (2015, p. 5) also indicated that there would be an increase in the demand for radiotherapy to the breast, prostate, lung, colon and rectum of an average of 21.9% across the United Kingdom by 2025.

A review by the Independent Cancer Task Force (2015, p. 13) predicted that 50% of those patients diagnosed with cancer now will survive at least 10 years. They identified that this would be due, in part, to advances in the treatment management of the disease. It is important, therefore, that these survivors can live disease free without suffering any long-term side effects from the treatment that they have received.

1.5 Cancer treatment options and outcomes

1.5.1 Surgery

Surgery has been reported by Watson, Barrett, Spence and Twelves, (2006, p. 34) as the modality most likely to cure patients with primary cancers confined to the organ of origin and with no clinical evidence of metastatic spread. But they also acknowledge that success is dependent on the complete removal of the tumour together with a margin of normal tissue to take account of any undetectable microscopic spread beyond the visible extent of the tumour. The authors also acknowledged that not all patients will be eligible for surgery because of comorbidities which may lead to an anaesthesia risk, the tumour being inaccessible because of its proximity to other vital organs or those who have declined surgery because of reported adverse impact on quality of life. Examples include patients with comorbidities such as chronic obstructive airways disease and limited lung capacity who may be at risk from anaesthesia. Other examples of when patients may decline surgery can be seen in the case of primary tumours arising in the prostate gland. Localised tumours confined within the capsule of the gland can be managed by surgical removal of the gland, part of the adjacent bladder neck, the seminal vesicles and the vas deferens, as described by Neal and Hoskin (2012, p. 168). But the surgical procedure is complex and associated with significant long term morbidity including erectile dysfunction and urinary incontinence in approximately 25% of patients.

1.5.2 Radiotherapy

Radiotherapy has been defined by Orth et al., (2014, p. 1) as the clinical application of ionising radiation in the form of X or gamma rays to restrict the growth of a tumour. In cases where patients are not fit enough for, or decline surgery, Erridge, Toy and Campbell (2012, p. 112) suggest that advanced techniques such as IMRT and stereotactic radiotherapy (SBRT) may be offered as an alternative. These techniques are discussed in more detail in chapter 2.3, p. 29. High energy x-ray beams are produced when electrons, accelerated in a linear accelerator waveguide to energies approaching the speed of sound, interact with a tungsten target. Conversely, gamma rays are produced as a result of the decay process of unstable radioactive substances such as Cobalt - 60. Whatever the production process, both X and gamma rays have high energy which, when transferred to the electrons orbiting the atoms of individual cells, cause the electrons to be ejected from the atoms leaving them in an ionised state (Sibtain, Morgan, & MacDougal, 2012, p. 2). This process results in the creation of free radical ions which are highly reactive at the molecular level and the source of radiation damage. The delivery of ionising radiation to the tumour can be achieved in one of two ways. Radiation delivered via external sources such as high energy X-ray and proton generators is a process known as external beam radiotherapy. External beam radiotherapy processes and workflows are covered in more detail in chapter 2.4, p. 35 as they form the background to this programme of research.

Radiation delivered by radioactive sources implanted directly into the tumour, or in close proximity to it, is referred to as brachytherapy (Delwiche, 2013, p. 120). The brachytherapy workflow involves the insertion of catheters or applicators under image guided control, following which; radioactive sources in the form of wires or seeds are loaded into them in a process known as afterloading. A range of different dose and fractionation regimes are available and are determined by the organ to be treated, the dose rate and specific activity of the sources being used (Hoskin & Coyle, 2013, p. 2) and whether brachytherapy is being used as a monotherapy or as an adjuvant to surgery and / or radiotherapy.

While brachytherapy is highly customisable, it is also highly dependent on the accurate placement of the applicators, catheters and sources. For example, a 1mm uncertainty in the position of the source may result in a 20% dose difference at a point 1cm from the centre of the source (Cunha, et al., 2019, p. 94). This is due to the rapid dose fall off so small geometric inaccuracies can translate into relatively large dose uncertainties (ibid, p. 102). Pre-registration learners in radiotherapy are unlikely to take an active role in the administration of brachytherapy during their training as it is viewed as an area of post graduate advanced clinical practice. However, they do need to have an appreciation of the fundamental principles of the process, which in turn requires mental visualisation of source placement in relation to the tumour and surrounding normal tissue.

1.5.3 Chemotherapy and biological therapies

The term generally refers to a group of agents which are used in the systemic management of cancer (Watson et al., 2006, p. 38). In a review of development timelines, DeVita and Chu (2008, p. 8643) reported that, until the 1960`s, while surgery radiotherapy dominated cancer management, cure rates remained at approximately 30%. This was attributed to the presence of undetected micro-metastases beyond the site of the primary tumour. Following the development of drugs to treat leukaemia and lymphoma (ibid p. 8648), chemotherapy became an accepted adjuvant treatment for breast and colorectal cancer in patients identified as having a high risk of developing metastatic disease. The anti-cancer action of chemotherapy agents is the targeting of the DNA of highly proliferating cells by using a combination of drugs to target the different phases of the cell cycle. Treatment is usually delivered in cycles repeated every two to four weeks over a course of administration lasting four to six months (Watson et al., 2006, p. 41). In the adjuvant setting, chemotherapy is given following local, primary site treatment, with surgery, radiotherapy, or a combination of both, to reduce the risk of the patient developing metastatic disease at a later date. Examples include cancer of the breast, lung and colon, where it can also be administered in the neoadjuvant setting (preceding definitive local treatment) for large tumours. The aim is to reduce the size of the tumour which will make it more accessible with surgery, or in the case of radiotherapy a reduction in the size of the tumour will reduce the size of the tumour target volume required to encompass the tumour.

In recent years chemotherapy has transitioned into an era of more targeted agents as identified by DeVita and Chu (2008, p. 8651). These agents offer a new dimension in cancer management through the selection of treatment regimens based on molecular profiling of cancer cell characteristics (Abrams, et al, 2014, p. 74). Collectively they are known as biological therapies and have been developed to target specific cancer cell molecular abnormalities and tumour microenvironment functions. Examples include monoclonal antibodies which can attach to cell surface proteins to inhibit cell growth and small molecule signalling and angiogenesis inhibitors which bind to growth receptors to switch signalling pathways off.

1.5.4 Summary of treatment options

The three modalities of surgery, radiotherapy and chemotherapy administered singly, as a monotherapy, or in combination play a key role in the management of cancer. As Bristow et al., (2018, p. e240) suggested, the targeting of tumours early in their natural history of development through the combination of the modalities of precision radiotherapy and molecular agents can augment local control and have a role in the ablation of micro-metastatic disease.

1.5.5 Cancer treatment outcomes

When determining the most appropriate treatment management intervention, the effectiveness of the proposed treatment option(s) may be determined by the impact that the interventions, either singly or in combination, may have on tumour growth and patient quality of life. From an individual patient perspective, the effectiveness of treatment has been identified by Neal and Hoskin (2009, pp. 29-30) as having five components which need to be considered. The first is disease free survival and relates to the length of time from completion of treatment that the patient will survive without signs and symptoms of the disease. It is linked to survival rate, which provides an indication of the percentage of patients within a specific diagnosis or treatment group who are still alive at a given time point after diagnosis. The second is response rate and while it can be linked to disease free survival, it also gives an indication of the amount of measurable disease present at a stated point in time post treatment. A complete response would indicate that there is no disease present whereas a partial response would signify a reduction in the size of the measurable tumour. Toxicity refers to the impact that the treatment may have on the patient in terms of acute (short term) reactions or the chronic (longer term) late effects. The risk of short-term impairment of normal function of an organ must be balanced against the risk of longer-term damage to the organ caused by the treatment, and from which the patient may not recover. Toxicity is closely linked to quality of life which has been identified by the World Health Organisation (2018) as a broad ranging concept covering an individual's perception of their physical health and psychological state. The nature of radiation damage is such that late toxicity in normal tissues may not become evident until many years post radiotherapy. Therefore it is important that the dose is delivered predominantly to the tumour while limiting the dose to surrounding normal tissue as much as possible. This appreciation of the relationship between dose, tumour and normal anatomy requires radiotherapy radiographers to have a high degree of mental modelling and spatial visualisation skills. While this is the case for both brachytherapy and external beam radiotherapy, the focus and emphasis for this programme of research will centre on the development and application of 3-D spatial visualisation in external beam radiotherapy.

1.6 The pre-registration radiotherapy curriculum

Radiotherapy pre-registration education typically incorporates a combination of academic study within a higher education institution and periods of experiential or work based learning components delivered within a hospital based radiotherapy department. Experiential learning has been defined by Kolb and Kolb (2009, p. 44) as the process through which knowledge is created via the transformation of experience. Learning therefore occurs as a result of observations and hands on practice in the workplace.

The curriculum structure in the UK has remained relatively unchanged over the last three decades despite the move from largely hospital based professional diploma level training to Framework for Higher Education Qualifications Level 6 education (Paterson, 2012, p. 48). Curriculum and programme specification documents tend to use the language of outcomes based learning with an emphasis on the development and application of knowledge and skills in order to achieve threshold standards of proficiency in line with the Quality Assurance Agency`s subject benchmark statements for radiography (Pratt & Adams, 2003, p. 319), statutory regulatory body standards of proficiency (Health & Care Professions Council, 2013) and National Occupational Standards (Skills for Health, 2011, n.p). Of the 15 standards of proficiency for radiographers published by the Health & Care Professions Council (HCPC), one refers specifically to the importance of what they refer to as spatial awareness. Standard 14.19 states that radiographers should:

> *"Be able to demonstrate spatial awareness, visual precision and manual dexterity in the precise and safe manipulation of treatment units or imaging equipment and related accessory equipment"* (HCPC 2013, p. 9).

This statement implies the need for a combination of spatial visualisation skills and coordinated psychomotor skills, but with no mandate for how this can be measured and achieved through pre-registration training. In addition, the professional body for radiographers in the UK, the College of Radiographers, have published an Education and Career Framework for the Radiography Workforce policy document (Society & College of Radiographers, 2013, p. 1) which provides an indicative curriculum for radiotherapy through a framework of learning outcomes for all levels of practice. The framework identifies 33 outcome statements which are mapped to the HCPC Standards of Education and Training for entry level, practitioner practice (HCPC, 2017) so that pre-registration learners can achieve the HCPC Standards of Proficiency for Radiographers (Society & College of Radiographers, 2013, pp. 12-16) by the time they complete their programme of education. Five of these statements (shown in table 1.2) specifically cover areas of practice which are considered to require highly developed spatial visualisation skills.

Table 1.2: Education and career framework statements of practice requiring spatial visualisation skill

Statement N°	Area of practice / Spatial visualisation skill
15	Identify, evaluate and interpret normal and abnormal anatomy and pathophysiology relevant to clinical practice
17	Employ effective patient positioning and immobilisation, customising devices as appropriate
19	Monitor and assess the adequacy of images
20	Interpret results and, where necessary, carry out additional image manipulation, imaging or adaptation of treatment delivery
24	Generate an optimal treatment plan and interpret radiotherapy prescriptions accurately, modifying these during treatment when necessary

While overall success on a programme of study involves the integration of information from multiple sources, gaining an understanding of new entrants' baseline spatial visualisation skill and any changes which may occur over time could contribute to the identification of those who may have less well-developed skills at the start of training. This understanding may offer the opportunity for academic and clinical educators to provide learners with additional focussed workshops and tutorials to develop these skills during their programme of study.

While the emphasis of this programme of research will focus on the importance and development of 3-D spatial visualisation in external beam radiotherapy, it is envisaged that the findings can be transferable to the development of the 3-D spatial visualisation skills required by other members of the cancer multidisciplinary team. For example, in minimally invasive colorectal surgery, novice surgeons will need to translate the information from a series of cross section CT images into a mental model of tumour position in order to gain optimal positioning of the patient and robotic surgical probes. In the delivery of chemotherapy, mental and spatial relationships between subcutaneous tissues, veins and the position of drug delivery cannula and catheters also need to be visualised in 3-D.

1.7 Supporting deliberate clinical practice with simulated and virtual environments

The development of clinical skills through observation and practice in the workplace is commonly referred to as the see one-do one apprenticeship model of teaching and learning (Mantovani, Castelnuovo, Gaggioli & Riva, 2003, p. 389). In this model, Woolley and Jarvis (2006, p. 75) identified that a novice (the new learner) may observe a task being performed by an expert (or master) and then attempt to carry out the same task while being observed. An illustration from radiotherapy practice would relate to the performance of a combined psycho-motor and spatial visualisation task for the accurate positioning of a patient prior to treatment delivery. Radiotherapy students undertake what McPake (2019, p. 222) refers to as a type of apprenticeship, in that they begin as novices and journey to expertise. However, while the accepted view is that in this arrangement there is a one-one partnership; radiotherapy clinical educators tend to work in teams of two or three. Each team will operate within a specified work area and do not generally have individually assigned patients. This has led McPake to argue that radiotherapy education is not strictly an apprenticeship model. That said, while a single student may be supervised by a single member of the team, they will not treat patients in the absence of radiotherapy clinical supervisors and educators.

Being the only student in a specific work area is important and may govern their perception of the success of their placement learning. For example, working in a linear accelerator treatment room all day enables more hands-on experience and one-to-one support from their clinical educators without competition from another student, therefore meeting the requirements for threshold skills. However, the need to increase placement capacity in order to increase training throughput has led many departments to adopt a paired student collaborative or peer assisted learning model. While this model offers the opportunity for reflection and sharing experiences and gaining a more holistic experience of the role of a radiotherapy radiographer, pairs of students do not collaborate in the treatment of individual patients because each student works with a different pair of clinical educators and is likely to treat different patients. This results in each student treating approximately 50% of patients on a daily appointment schedule. However, as Palma (2017, p. 510) observed in a recent editorial, high patient numbers do not necessarily guarantee high levels of proficiency. Repeating an incorrect practice again and again will not lead to skill development and improvement. A further disadvantage of this model is that while the daily radiotherapy treatment workload can be predicted, the case mix of new patients entering the schedules cannot. This unpredictability has been identified by Bridge & Carmichael (2014, p. 45) as having the potential for opportunistic, rather than systematic learning in that the same knowledge and skill development cannot be guaranteed for all learners and may lead to dissatisfaction with clinical placement experience.

During the last two decades there has been a tendency to move away from this apprenticeship style and sometimes ad-hoc teaching in the workplace to learning based on theoretical knowledge as reported by Monaghan (2015, p. 1). This is the case especially where consequences of error may be high (refer to chapter 2.6, p. 46 for more detail relating to radiotherapy errors). The need for improved clinical skill preparation in a safe environment, which is removed from the clinical environment, has seen a rise in the implementation of simulation platforms and simulated exercises in the UK. This has been coupled with changes in patient pathways and the increasing complexity of care delivery which have resulted in fewer opportunities for students to learn from a range of real patients (Ker & Bradley, 2010, p. 165). One of the major drivers for the use of simulation is to develop safe healthcare practitioners (ibid, p. 164). Proposals for increasing the use of simulation for high risk activities have also arisen from a number of patient safety reports and public enquiries, notably the Bristol Royal Infirmary report, which placed an emphasis on improving the safety and standard of care (Kennedy, 2001, p. 450) and reducing the number of adverse harmful incidents (Donaldson, Panesar & Darzi, 2014, p. 3). The use of simulated and virtual learning environments provides systematic learning of high-risk interventions which can be practised and repeated many times in a safe and controlled environment. Novices can learn from their mistakes (error consequence) with no risk to patients. In relation to radiography in general, Kong, Hodgson and Druva (2015, p. 31) identified the risk of delivering inappropriate doses of ionising radiation and proposed that the risk-free practise of clinical skills without compromising patients' safety can be achieved in simulation.

Simulation as a technique, has been referred to by Gaba (2004, p. 2) as a method which could amplify or replace real experiences. This is a view that is supported by Ker and Bradley (2010, p. 164), who report that simulation may involve a range of techniques applicable to learners at all levels from novice to expert. An alternative view is offered by Tavakol, Mohagheghi and Dennick (2008, p. 77) who referred to simulation as a technology rather than a technique in that situations and conditions can be created artificially in order to experience something that would exist in reality. Both viewpoints can be considered as valid, since the simulation technology that is available can support teaching techniques that support learning through deliberate practice to improve student performance in a safe environment (Kong et al., 2015, p. 30).

The concept of deliberate practice, as proposed by Anders-Ericsson and Lehmann (1996, p. 279), is that of individualized training activities specially designed by a teacher or facilitator to improve specific aspects of an individual's performance through rehearsal, repetition and successive refinement. These activities also support opportunities for the provision of immediate contingent feedback to support critical reflection on the part of the learner to identify any limitations and knowledge gaps (Okuda et al., 2009, p. 333). The deliberate practice of clinical tasks can be supported by using a range of simulated environments or simulation platforms which reproduce aspects of clinical work ranging from replication of a single task, or part thereof, to recreation of an entire environment (Maran & Glavin, 2003, p. 24). In radiotherapy specifically, Chamunyonga, Edwards, Caldwell, Rutledge and Burbery, (2018, p. 241) defined deliberate practice as:

> *"Purposeful skill augmentation through a strategic repletion of an area of practice guided by feedback from a mentor throughout the process".*

They go on to propose that radiotherapy educators need to ensure that new graduates have met the threshold values and skills, firstly to meet regulatory body standards of proficiency and secondly, to support improvements in the quality and safety of care delivery. They also identify that while the evidence base supports the role of simulation based learning combined with deliberate practice, the evidence base supporting approaches to radiotherapy specific deliberate practice is limited.

During the course of this programme of research there has been a tendency to move away from apprenticeship style training in the clinical environment to an emphasis on safety and rehearsal in simulation based education. It is interesting to note therefore, that today there are a range of flexible, learning at work opportunities within the NHS, which range from intermediate level apprenticeships at level 2 (GCSE grades 4 – 9 equivalent) of the Framework for Higher Education Qualifications (FHEQ) to levels 6 and 7 degree courses. These programmes will have different delivery structures compared to more traditional degree programmes with more time being spent in the clinical environment. The Universities and Colleges Admissions Service have indicated that 80% of a working week would be spent in the place of employment and 20% at a place of study.

1.8 Types of simulation platforms

The literature describes a range of simulation platforms with Okuda et al., (2009, pp. 331-332) dividing them in four distinct classifications based on complexity: standardised patient actors; desktop personal computer based systems; partial task trainers and high fidelity manikins and virtual reality environments. For the purpose of setting the context for this programme of research, attention will focus on the final classification which is where the virtual environment for radiotherapy training (VERT™) platform is positioned.

Virtual reality refers to a collection of technologies that provide realistic experiences that allow individuals to interact with 3-D computerised databases in real time by using their natural senses and skills (Tavakol, Mohagheghi, & Dennick 2008, p. 81). In a typology of medical simulation tools, Alinier (2007, p. 243) identified interactive virtual reality patient simulators as being just below real life. Virtual reality has been defined by Maran and Glavin (2003, p. 24) as the ultimate computer based technology while Dalgarno and Lee (2010, p. 19) have reported that 3-D virtual learning environments deliver the opportunity for the experiential learning of tasks that are impractical or impossible to undertake in the real world because the concept or domain is abstract.

1.9 High fidelity virtual reality environments

A high-fidelity virtual environment can be considered as a recreation of a working environment which can combine sophisticated whole or part body manikins and computer modelling of treatment interventions, equipment or physiological processes (Maran & Glavin, 2003, p. 25). One such platform is the Virtual Environment for Radiotherapy Training (VERT™). It was initially developed to model and visualise complex radiation treatment delivery subcomponent processes to support the training of medical physicists and treatment planners, also known as dosimetrists (Phillips, Ward & Beavis 2005, p. 392). The nature of the immersive 3-D visualisation capabilities of the VERT™ platform places it in the domain of high fidelity virtual reality. The platform has three operational modes: 2-D visualisation; 3-D stereoscopic visualisation and 3-D stereoscopic visualisation with incorporated user tracking. Visualisation in 3-D is achieved via stereoscopic, forward or back projection with users wearing liquid crystal active shutter goggles or polarising glasses. Full 3-D immersion is achieved via active user position tracking. The different operating modes can support individual and peer assisted learning of fundamental radiotherapy first principles and equipment terminology through visualisation. This is combined with user interaction with an original equipment manufacturer (OEM) hand pendant to develop and practice the psychomotor skills required to control the virtual linear accelerator.

Visualisation has been reported to be helpful in the early stages of radiotherapy education when a number of complex and unfamiliar concepts need to be integrated into the accurate positioning of patients prior to treatment (Green & Appleyard, 2011, p. 178). As the understanding of the first principles of radiotherapy progress and technical skills develop, the platform can be used to visualise and practice more complex techniques such as IMRT and volumetric modulated arc therapy (VMAT). The VERT™ platform can also be used to model correct and incorrect patient positioning. This particular function can assist in the visualisation of correct and incorrect tumour target volume position and the impact on other organ position in relation to X-ray beam placement (Thoirs, Giles & Barber, 2011, p. 9).

As Montgomerie, Kane, Leong and Mudie (2016, p. 204) have also identified, VERT™ can provide an alternative method to support the transition between 2-D representations and visualisation of organs in text books and CT datasets to a real world understanding of internal anatomy. They also reported that, in their experience, the integration of VERT™ across the curriculum in conjunction with traditional teaching methods has enhanced the learning of the underpinning conceptual knowledge of radiotherapy. A detailed examination of the research base for the use of VERT™ can be found in chapter 2.8 (pp. 50 - 73). Following the installation of the VERT™ platform across England during 2008 the platform has been introduced across the rest of the world. Currently there are over 140 systems operating across 130 academic, clinical and industry sites in 30 countries worldwide (T. Swayne, personal communication, 2018; www.vertual.co.uk 2020).

1.10 The integration of simulation into the University of Portsmouth radiotherapy curriculum

The Centre for Simulation in Healthcare, located within the School of Health and Care Professions, comprises a number of physical spaces including ward bays and, open access, multi-functional clinical observation areas. The equipment inventory covers a range of part task trainers, high fidelity human patient simulators (which can support initial clinical assessment, postoperative care and advanced cardiac life support) and an Anatomage 3-D virtual human dissection table to support anatomy learning and teaching. In addition to the above, there are also a number of fully equipped laboratories including an operating theatre, diagnostic imaging and haptic ultrasound suites and the virtual environment for radiotherapy training (VERT™).

The VERT™ platform was installed and commissioned at the University of Portsmouth in June 2008. Employed for the first time in September 2008, the platform has been embedded into the curriculum at FHEQ levels four and five (years one and two) of the radiotherapy pre-registration programme to support the structured understanding of the basic terminology and first principles of the science of radiotherapy and pre-clinical orientation. The learning outcomes based practical workshop lesson plans (appendix 1) have been designed to encourage peer to peer learning through exposure to the concepts of and practice with radiotherapy treatment delivery methods.

The platform also supports the transition from understanding the first principles of radiotherapy to the understanding the more advanced techniques of IMRT and VMAT as identified above. The development of clinical and technical skills in a safe, campus based, environment is supported by the provision of a series of formative rehearsal and feedback sessions using individual and team scenarios prior to experiential learning in radiotherapy clinical departments. Between 2009 and 2017[1], it was also used with first year students for the summative assessment of the understanding and application of the processes required for daily linear accelerator quality assurance checks. During their first year of study, radiotherapy students will have a minimum of 15 hours scheduled small group practical workshop and tutorial time prior to the commencement of the first clinical practice placement in November, eight to nine weeks after the start of the programme. In addition, the campus timetable supports a minimum of six hours post placement time for student led peer to peer problem solving activities. During year two, students have the opportunity, over a minimum of eight hours of timetabled practical workshops, to practice more advanced positioning tasks. These are usually conducted in teams of three to replicate the team working patterns usually encountered in radiotherapy departments, with each student taking the role of team leader in rotation. First and second year students can also access the platform on an ad-hoc basis, depending on room availability, for unfacilitated peer to peer problem solving. Following software upgrades in 2018, year two and year three students receive a minimum of nine hours dedicated to 3-D cone beam CT image acquisition, pattern recognition and verification image approval practice. The acquisition of a laptop, 3-D data projector and a software licence for a patient education and learning module (PEARL) means that the VERT™ platform can be used to support the teaching of the concepts of radiotherapy to other health science students and for careers and educational outreach events across the University and wider community.

[1] The VERT™ assessment was removed in September 2017 as a result of programme revalidation and realignment of assessment methods

1.11 The rationale, aims and objectives for this programme of research

This section will summarise the themes identified and discussed throughout this chapter which together underpin the rationale and the justification for conducting this programme of research. The rising incidence of cancer will increase the demand for radiotherapy, which in turn will result in the need for additional equipment and an increase in the therapeutic radiography (radiotherapy) workforce. In addition, radiotherapy treatment planning and delivery techniques have evolved from relatively simple 2-D to complex 3-D techniques such as IMRT, VMAT and SBRT. These techniques are now being delivered as standard and require new graduates to have specialist skills which have, until recently, been viewed as being part of the advanced practitioner (post graduate) skill set and will be discussed further in chapter 2.5, p. 45. Combined, these factors have the potential for conflict, in that, increasing demand and complexity need to be balanced against achieving a high quality clinical placement experience for learners. The National Radiotherapy Advisory Group (NRAG, 2007, p.25) identified that increasing capacity demands and time pressures on clinical radiotherapy had led to a poor experience for some students and a resultant high attrition rate. However, the reasons for poor retention are multi-faceted and in the researcher's experience, students often cite "personal reasons" for leaving a radiotherapy education programme. What is not clear from this and remains relatively under researched, is whether poor experience, personal reasons, or a combination of both is due solely to time pressures. Could these time pressures be having an impact on those learners who have difficulty with the processing of complex visual information, if so, what additional support might assist them in the development of these skills?

The increasing role for simulation platforms and environments in clinical education and preparation for clinical practice has been identified. Simulation, as an educational technique, has the potential to increase time on task, thereby maximising the impact of hands on learning in complex and potentially high risk tasks in a safe environment. However simulation on its own may not contribute to an increase in staffing levels unless more students can be retained by improving their clinical experiences. The challenge is how to identify an effective use of simulation resources which moves beyond a one size fits all approach. While the VERT™ platform has several advantages, particularly in the visualisation of complex radiotherapy principles, its role in supporting the development of and improvement in 3-D spatial visualisation skill remains under researched. This research set out to investigate if a more focused use of the platform was possible. This required the identification of learners who may be challenged by 3-D concepts and therefore at risk of not being able to process and act on complex 3-D visual information. If this could be done, would it lead to the identification of more focussed VERT™ activities, matched to individual learners 3-D spatial visualisation skill development needs? These themes provided the drivers for the formulation of the research aims and objectives which set out in table 1.3. It is recognised that, while the focus for this programme of research centred on VERT™, the findings could be transferable to other areas where simulation is employed.

Table 1.3: Summary of research aims and objectives

Research Aims	
1	To gain an understanding of the spatial visualisation skill of pre-registration learners in radiotherapy in one United Kingdom Higher Education Institute;
2	To determine whether it was possible to stratify pre-registration radiotherapy learners in terms of their baseline spatial visualisation skill;
3	To determine the longer term potential of VERT™ in relation to the development of 3-D spatial visualisation skill.
Research Objectives	
1	Conduct a systematic review of the literature to identify and define the components of spatial visualisation skill required for radiotherapy practice;
2	To identify valid and reliable 3-D spatial visualisation skill measurement tools via a critical evaluation of the spatial visualisation testing literature;
3	Develop an appropriate test instrument for use in radiotherapy;
4	Compare performance in paper and online versions of the test instrument
5	To determine if the baseline spatial visualisation skill of pre-registration learners in radiotherapy could be measured;
6	To determine if spatial visualisation changes over time;
7	To determine if a relationship exists between baseline spatial visualisation skill and performance in a complex radiotherapy positioning task;
8	To determine if a relationship between baseline spatial visualisation skill and previous spatial visualisation experience exists;
9	To make recommendations for future educational practice.

1.12 Thesis layout, content and structure

The thesis continues in:

Chapter 2 which will provide a technical narrative review of the external beam radiotherapy treatment planning and delivery pathway;

Chapter 3 will review the literature relating to general intelligence and spatial visualisation skill and the factors which may influence the development of spatial visualisation skills. It will then provide a critical evaluation of the literature reporting the measurement of 3-D spatial visualisation skill and the test instruments used. It will conclude with a summary of the research conducted with VERT™ to date;

Chapter 4 will discuss the underpinning epistemology and the supporting ontological perspectives for this programme of research before examining the research design and methodological considerations;

Chapter 5 details the design and findings of two quantitative studies and one qualitative study which were conducted as the pilot phase of this programme of research:

- Study one was designed to determine if the spatial visualisation skill of a cohort of radiography learners could be measured using traditional paper-based measurement instruments and an alternative online version;
- Study two sought to determine if the test platforms used in study one could detect any change in SVS over time;
- Study three, the qualitative study, was employed to ascertain the utility and participant acceptability of the online version (utility being defined as useful and acceptability as suitable or agreeable by the New Oxford Dictionary of English).

Chapter 6 will continue by reporting the experimental phase which consisted of three studies:

- Study four was designed to measure the baseline spatial visualisation skill of novice diagnostic imaging students (the control group) and radiotherapy students (the experimental group) and to determine if spatial visualisation remains stable or changes over time;

- Study five sought to determine if there was any relationship between baseline skill and performance in a complex simulated clinical task;
- Study six investigated the relationship between baseline spatial visualisation skill and previous spatial activities.

Finally, chapter 7 will provide the overall discussion of this programme of research in the context of the spatial visualisation literature and the findings of each of the six studies. This will be followed by an outline of the contribution of this research to the understanding of 3-D spatial visualisation skill in radiotherapy, whilst acknowledging the scope and limitations. Finally the chapter will draw conclusions identifying the wider implications of the findings and make appropriate recommendations for education, practice and further research.

1.13 Chapter summary

The chapter has identified the current position and predictions for the incidence of cancer in the UK and the treatment methods which are aimed at achieving long term disease free survival. The most successful, non-surgical, treatment is radiotherapy; however, its success depends on the ability to deliver a high dose of radiation to the tumour while minimising the dose to surrounding normal tissues and organs. Advanced planning and delivery techniques can achieve this but require a high degree of accuracy. Three-dimensional spatial visualisation skills can support this accuracy but the challenge for new learners is to develop these skills in the often time pressured clinic environment. The importance of safety in the clinical setting in general has been identified and the impact and role of simulated environments in the development of clinical skills has been discussed. The introduction of the VERT™ platform may offer educators in radiotherapy the opportunity to develop new ways of teaching the underpinning principles and visualisation of the radiotherapy process. However the nature of spatial visualisation in relation to the radiotherapy process and the impact of the platform on the development of spatial visualisation skill remain under researched. Therefore the research aims and objectives have reflected the need to understand the field of spatial visualisation, its application to the radiotherapy process and how it may be measured. These themes will provide the focus of the literature review presented in chapters two and three.

Chapter 2:

The evolution of radiotherapy and the radiotherapy pathway

2.1 Introduction to chapter two

This chapter will begin by exploring the evolution and development of radiotherapy treatment planning and delivery techniques before providing an appraisal of current radiotherapy practice. It will do this by adopting a technical narrative review approach as advocated by Greenhalgh, Thorne and Malterud (2018, p. 12933). The technical narrative review draws on published literature, critical reflection and the researchers experience in the field to provide evidence informed interpretations of the current state of play (Kane 2018, p. 131). It will continue with a discussion focussing on the importance of 3-D spatial visualisation in the safe and accurate planning and delivery of radiotherapy and the consequence of radiotherapy errors. The chapter will then conclude with an examination of the research base reporting the role of VERT™.

2.2 The evolution and development of radiotherapy practice

The primary goal for both external beam radiotherapy and brachytherapy, as outlined by Hand, Kim and Waldow (2004, p. 77), is to eradicate the tumour by controlling tumour growth while avoiding un-repairable damage to normal tissue. The probability of controlling tumour growth with radiotherapy is proportional to the dose of radiation that can be delivered (Verellen et al., 2007, p. 949). The aim of radiotherapy, therefore, is to deliver as close as possible to 100% of the radiation dose to the tumour target volume and as close to no dose at all to the normal tissues and organs surrounding the tumour target volume (Thariat, Hannoun-Levi, Myint, Vuong & Gérard, 2013, p. 52). In practice this is not achievable, but steep dose differentials between the tumour target volume and surrounding tissues can be achieved with the use of advanced treatment delivery techniques (discussed in chapter 2.3, pp.28 - 35).

In the early part of the 20th century the X-ray tubes used for radiotherapy operated at energies of between 50 and 200 kilo-volts (kV). This relatively low generating energy resulted in an inability to deliver adequate dose to deep seated tumours as identified by Thariat et al., (2013, p. 52). This lack of penetrating power resulted in the cure of cancer with radiotherapy being limited to small tumours on, or just below, the skin surface. Delivering radiation with enough energy to reach deep seated tumours requires equipment capable of generating radiation beams with high energy in the megavoltage (MV) range (Thwaites & Tuohy, 2006, p. 347). The early 1950`s saw the development of equipment capable of generating these high energies following the invention of microwave power sources for radar during the 1940`s. The first medical linear accelerators, capable of generating beams with energies between 6 and 20 MV were introduced into clinical practice in 1953 (Thwaites & Tuohy, 2006, p. 343). While the basic structure has remained relatively unchanged, the range of functions and capabilities has continued to evolve into the 21st Century. The components of a modern linear accelerator are depicted in figure 2.1.

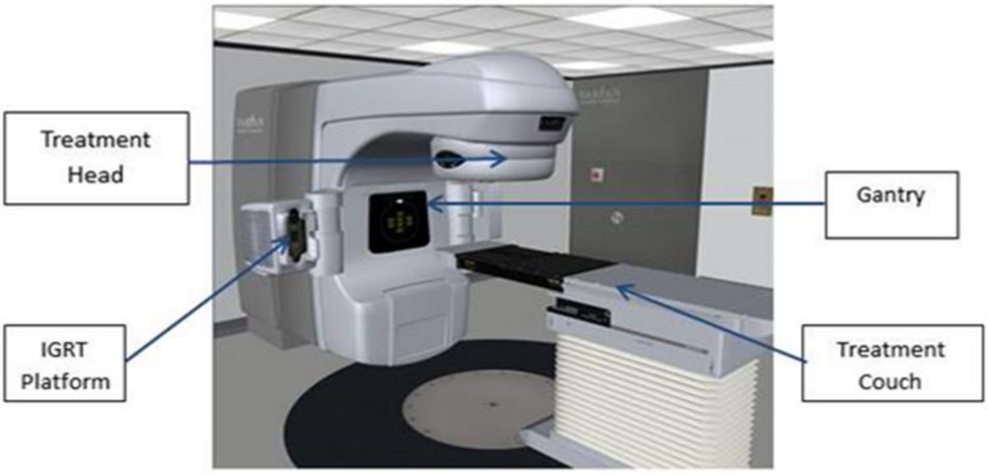


Figure 2.1: A Varian Clinac iX linear accelerator as modelled in VERT™ v 3.0. (Screenshot from UoP VERT™ platform with permission of Vertual Ltd, 2016)

The low energy machines generating X-ray beams with energy of 100kV deposit 100% of the dose on the surface of the skin, but only 13% at 10cms deep. Whereas linear accelerators generating X-ray beams with energies of 6MV can deliver 100% of the dose at 1.5cms deep and 67% of the dose at 10cms deep. In contrast, electron beams, also generated by a linear accelerator, can be utilised for the treatment of superficial or subcutaneous disease. Unlike X-ray dose distribution, a mid- range electron beam energy of 10MeV will deposit 90% of the prescribed dose at 3.1cms deep, falling to 50% at 3.9cms and a practical range of 4.8cms, beyond which no dose is received. This provides an alternative option in the treatment of superficial tumours in close proximity to organs at risk (OAR) as identified by Strydom, Parker and Olivares (2005, p. 286). When using electrons for radiotherapy treatment, the beam coverage is determined by an applicator of appropriate size attached to the linear accelerator treatment head. Trimmers can be attached to the applicator to conform the shape of the beam to the treatment area if it has an irregular shape. To achieve dose homogeneity within the treated area, the technique usually employs a single direct field with the central axis of the beam entering the skin surface perpendicular to it. This technique is known as a skin apposition technique and will be explained in more detail in the introduction to study six in chapter 6.6.1 p.227.

A more recent introduction to radiotherapy treatment in the UK is that of proton beam therapy. As Barker, Lowe and Radhakrishna (2019, p. 575) have reported, proton beams penetrate tissue to a limited depth and deposit their energy as their velocity decreases at the end of their path. This results in the radiation dose building up to a maximum (known as the Bragg peak) and then falling off sharply with no dose deposited beyond the finite range. This range is governed by the generating energy of the beam. Typical beam energies range from 60 to 230 MeV, the higher the generating energy, the greater the penetration and hence the depth of the Bragg peak. At lower energies, for example 80 MeV, the Bragg peak will occur at a depth of 5 cm, while the mid-range energies of 140 and 180 MeV will produce a Bragg peak at 14cms and 21cms deep and at 230MeV the peak will occur at 30cms (Almhagen, Boersma, Nyström & Ahnesjö, 2018, p. 31). The Bragg peak is narrow compared to the size of the target volume that needs to be covered, so in order to achieve optimal dose coverage of the volume in depth, a spread out Bragg peak is created by using different proton energies. As Barker et al., (2019, p. 576) continue, using a technique known as pencil beam scanning, the tumour is treated in layers where the whole target volume at a particular depth is covered with a specific proton beam energy whose Bragg peak is matched to that depth. The next layer is then treated with a different beam energy and repeated until the delivered dose encompasses the longitudinal extent of the tumour. However proton beams are particularly sensitive to changes in patient position and therefore tissue composition along the beam path because of their finite range and sharp distal fall-off. These can occur due to daily variation in patient set-up for treatment, because of organ motion with breathing, or with changes in patient anatomy such as weight loss during treatment.

2.3 The development of advanced radiotherapy treatment techniques

The developments in equipment design and the associated advances in radiotherapy techniques have influenced both the position of radiotherapy in cancer management and the role of radiotherapy radiographers (White & Kane, 2007, p. 298). The introduction of 3-D imaging for tumour location has changed the emphasis from a broader organ based tumour position to a more focussed optimisation of target volume dose delivery as discussed in the following section.

2.3.1 Three-dimensional conformal radiotherapy

Three dimensional conformal radiotherapy techniques developed during the 1990`s in order to achieve a high-dose volume shaped to conform to the tumour target volume in three dimensions. The techniques were designed to irradiate tumour sites with radiation beams whose apertures were shaped using customised dense alloy blocks or by multileaf collimators (part of the beam defining system) housed in the treatment head of a linear. Multileaf collimators consist of 40 - 60 pairs of tungsten bars 0.5 to 1 cm wide which can be adjusted in length to conform the X-ray beam to the irregular shape of the 3-D tumour target volume based on 3-D reconstructions of target and anatomical information available from cross sectional computed tomography imaging (CT). This 3-D planning information also delivered the capability to graphically reconstruct the relationships between the tumour target volume shape and its position in relation to other organs. Clinical teams were able to view computer generated anatomical data as if it was being viewed along the axis of the radiation beam (known as a beam`s eye view) as described by Pradu, Starkshall and Mohan (2007, p. 124). The ability to conform the shape of radiation beams to the 3-D shape of the tumour target volume also led to the ability to increase the dose delivered to the tumour target volume, while at the same time, reducing the volume of normal tissue irradiated by up to 50% in a technique known as dose escalation as reported by Senan et al., (1999, p. 247). This ability to deliver higher doses to the tumour also led to improved rates of local control of tumour growth compared to the earlier techniques based on 2-D planning as identified by Rosenzweig et al., (2005, p. 2124) since it was possible to reduce the volume of normal tissue within the tumour target volume.

2.3.2 Intensity modulated radiotherapy

Intensity modulated radiotherapy is an enhanced application of 3-D CRT which can conform dose delivery to the shape of the tumour target volume by constantly changing the shape and position of multi-leaf collimators as described by Mackay, Staffurth, Poynter and Routsis (2010, p. 629) and Høyer et al., (2011, p. 149). X-ray beams delivered from a number of discrete linear accelerator gantry angles, are composed of a number of small beamlets of different shapes which vary the intensity of dose which is delivered across each beam. This varying intensity is designed to achieve the predetermined specification of dose requirements to the tumour target volume and dose constraints to normal tissue surrounding the target volume. The following screenshots from the University of Portsmouth VERT™ platform (figures 2.2 and 2.3) demonstrate the different beamlet shapes from one static gantry angle and the combination of several gantry angles which converge at the centre of the tumour target volume.

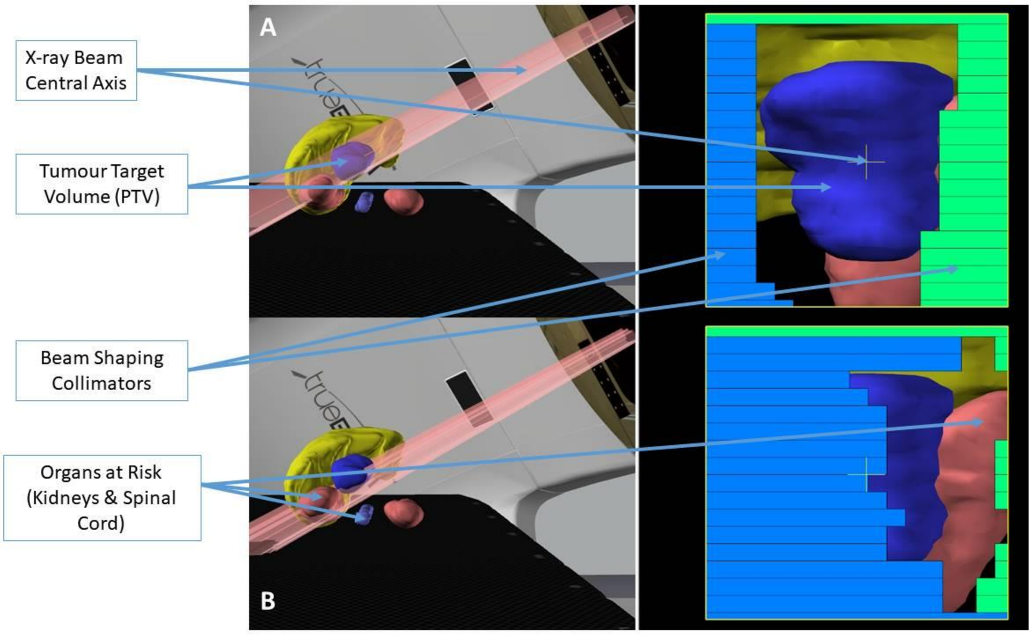


Figure 2.2: Beamlet A will deliver dose to the tumour target volume, while beamlet B will deliver dose to part of the tumour target volume while avoiding the underlying spinal cord

The close tumour target volume conformity and varying intensity serves to reduce the volume of high dose outside the PTV and therefore reduces the dose to normal tissue within the treated volume. This technique has been shown to be particularly beneficial for tumours arising in the head and neck region. This is due to the reduction in dose delivered to the parotid gland which reduces the incidence of dry mouth post radiotherapy. For example, a matched case control study conducted by Agee (2017, p. 349) compared IMRT and 3-D conformal radiotherapy in 207 patients with head and neck tumours. The results showed that at the 12-month post treatment follow-up point, significant differences related to problems with dry mouth favouring the IMRT group (72.1 vs 62.8; p = .018) were observed in patient self-report quality of life questionnaires.

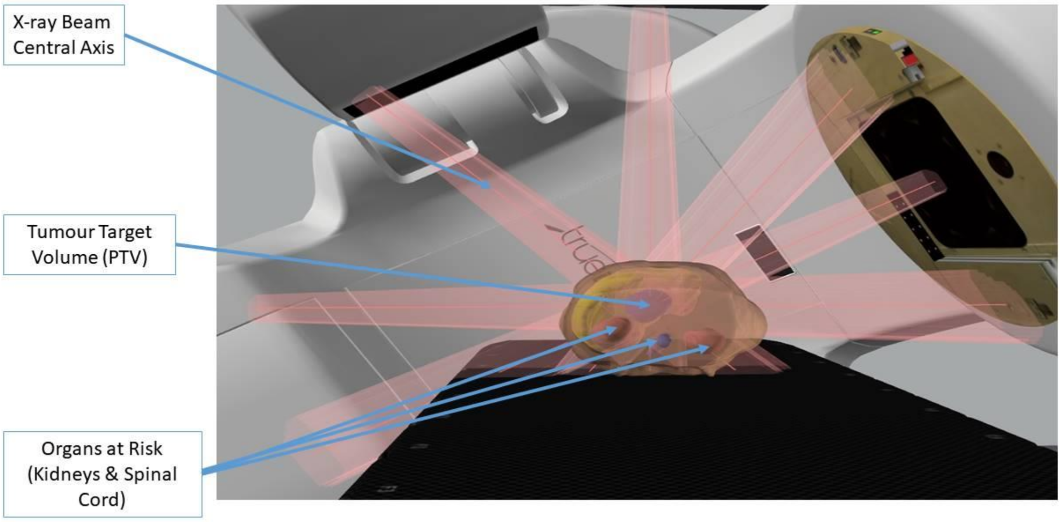


Figure 2.3: Demonstrating the convergence of five IMRT X-ray beams at the centre of the PTV

2.3.3 Volumetric modulated arc therapy

Volumetric modulated arc therapy was introduced in 2007 as an advance on static gantry IMRT which added rotation of the linear accelerator gantry, variation of its speed of movement and rate of dose delivery simultaneously as described by Teoh, Clark, Wood, Whitaker and Nisbet (2011, p. 968). Delivery of radiation from a continuous 360^{o} rotation of the gantry is more efficient than IMRT. A single arc VMAT treatment session has a typical beam on time less than two minutes compared with up to 10 minutes for a five or seven field fixed gantry IMRT treatment session (ibid, p. 969). Decreasing the overall treatment delivery time reduces the risk of organ movement during each session. Minimising this risk of movement can be of particular importance in the treatment of prostate tumours where significant changes in rectal and bladder volumes due to organ refilling during IMRT treatment delivery could compromise tumour target volume dose coverage and reduce tumour local control. Figure 2.4 shows a sequence of three beamlets of differing sizes, shapes and positions delivered from VMAT different angles.

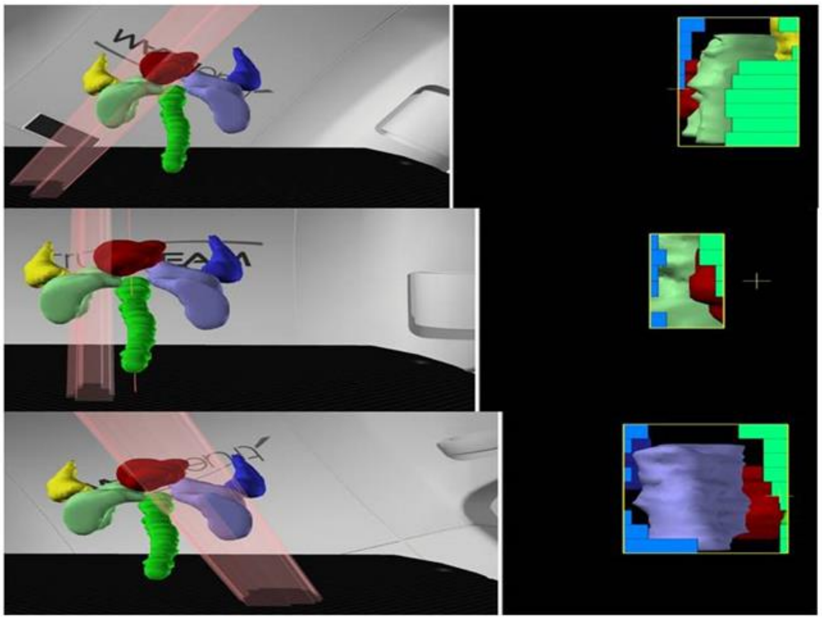


Figure 2.4: Three VMAT beams directed to the tumour target volume (shown in red)

2.3.4 Stereotactic radiotherapy

Stereotactic radiotherapy techniques were originally developed to treat tumours within the brain but over recent years, their use has been extended to other anatomical sites in other regions of the body including lung, prostate, liver and pancreas. Stereotactic body radiotherapy is considered the primary alternative to surgical lobectomy in patients with early stage lung tumours who are unfit for surgery (Høyer et al., 2011, p. 149). Overall target volume sizes are typically less than five centimetres in diameter with treatment being delivered using higher daily doses over fewer treatment sessions compared to IMRT and VMAT. Doses range from 22 – 55 Gy delivered over three sessions (7.3 – 18.3 Gy per session) compared to 60 - 66 Gy in 30 - 33 sessions with IMRT (2 Gy per session) as reported by Franks, Jain and Snee (2015, p. 286).

2.3.4 Image guided radiotherapy

High precision in the delineation of the tumour target volume and dose delivery requires a reduction in the uncertainty relating to the position of the tumour target volume and the relational organs during and between treatment sessions. This has become increasingly important as tumour target volume sizes decrease and daily radiation doses increase. The development of linear accelerator-based CT scanning has resulted in the ability to acquire in the region of 200 planar images in one rotation of the gantry prior to the delivery of each session of radiotherapy (Srinivasan, Mohammadi & Shepherd, 2014, p. 184). The soft tissue spatial detail available from CT facilitates online tumour target volume alignment in the X (lateral), Y (longitudinal) and Z (vertical) planes and the three rotations around these planes (known as roll, pitch and yaw). Any discrepancy between the intended (planned) and actual (daily) position of the target in these six directions and the magnitude of that discrepancy is calculated using computer algorithms. Following comparison of the images acquired daily prior to treatment delivery and the original localisation scan, data corrections to the position of the target volume and radiation beam path can be applied. This process is referred to as image guided radiotherapy (IGRT) and can detect and correct for the random (patient related) and systematic (equipment and process related) positional errors which may occur in the daily delivery of radiotherapy. While image guided radiotherapy (IGRT) is not a treatment delivery technique as such, it is embedded in the advanced treatment delivery workflows to support accurate patient position and dose delivery. Correction of positional errors minimises the risk of under dosing the target volume and overdosing of normal tissue with Høyer et al., (2011, p. 150) reporting that the use of 3-D CBCT in radiotherapy to the lung reduces the median set up error from 6 mm to 2 mm. Therefore daily 3-D position verification delivers target volume-oriented positioning rather than patient-oriented position offered by 2-D imaging when the presumed position of the tumour was determined by its proximity to bony anatomy. As Verellen et al., (2007, p. 949) have identified, the introduction of IGRT has enabled visualisation of the exact position of the tumour prior to each treatment delivery and has decreased PTV margins from centimetres to millimetres. For example, Bhide and Nutting (2010, p. 440) have reported that the safety margin for a spherical tumour of 5cms diameter can be reduced from 2cms to 5mm with a resultant decrease in the irradiated volume of the surrounding organs from 316 cm^3 to 48 cm^3. Until recently, image review and approval for IGRT has been considered to be outside the radiographer skill set, as identified by (Gillan, Li & Harnett, 2013, p. 242). However as radiotherapy radiographers roles change, discussed further in section 2.5, p.43, below, then the need for the development of visualisation of 3-D structures increases. For pre-registration education programmes the VERT™ platform can integrate academic theory and clinical application for the development of conceptual understanding of the complex radiotherapy treatment planning and delivery techniques of IMRT and VMAT and the principles and application of IGRT. Qualified radiotherapy radiographers, on the other hand, may have developed their 3-D skills spatial visualisation skills along the same lines as the researcher (refer to chapter 4.14, p. 150); however, the evidence is largely anecdotal and lacks a research base. For more detail on the role of VERT™ in supporting staff in the development of their CT skills refer to the study by Shah and Williams (2010) in section 2.8, p. 54, below.

2.3.5 Adaptive radiotherapy

The use of 3-D CBCT can identify, quantify and track target volume movement and change in its shape and position over time as result of tumour shrinkage due to radiotherapy. It therefore follows that this ability can lead to modifications to the treatment delivery plan being made during a course of radiotherapy. This process is known as 4-D adaptive radiotherapy and ensures correct dose delivery to the tumour target volume (Høyer et al., 2011, p. 151). One method of achieving this employs a plan of the day strategy that facilitates the selection of the most appropriate plan to achieve optimal dose coverage of tumour target volume. Proposed adaptive workflows suggest that the radiotherapy radiographers delivering the treatment will take the decisions relating to the most appropriate plan of the day.

During this programme of research, the technology for the delivery of the radiotherapy techniques referred to above have continued to evolve and develop. One example is the introduction of the Halcyon 2™ treatment delivery platform in 2017 (Varian Medical Systems, Palo Alto, CA, USA). The platform provides a fully automated nine step process for patient positioning and verification followed by a two minute beam-on time with remote patient unload at the end of the procedure to minimise the length of time a patient has to remain on the treatment couch. While this platform is likely to streamline the treatment delivery workflow, it may also reduce opportunities for learners to synthesise and assimilate fundamental principles into clinical practice and foundational academic learning further. In addition, the automation of patient positioning and treatment delivery will change the way that learners gain their visual information and will rely increasingly on digital sources such as computer monitors, rather than traditional paper-based methods such as hard copy dosimetry plans and treatment delivery charts.

2.4 The radiotherapy workflow

Following a clinical decision to treat a patient with radiotherapy, a number of steps need to be completed to ensure the safe and effective delivery of dose to the correct tumour target volume. This process has been referred to as the radiotherapy chain of operations by Vieira, Hans, Van Vliet-Vroegindeweij, Van de Kamer, and Van Harten, (2017, p. 130). The chain is characterised by two distinct phases which have been identified by Joustra, Kolfin, van Dijk, Koning and Bakker (2012, p. 451) as the pre-treatment preparation phase (comprised of three stages: patient immobilisation, CT scanning and dosimetry) and the treatment delivery phase (comprised of two stages; verification and delivery). These phases and stages as summarised in Figure 2.5 below.

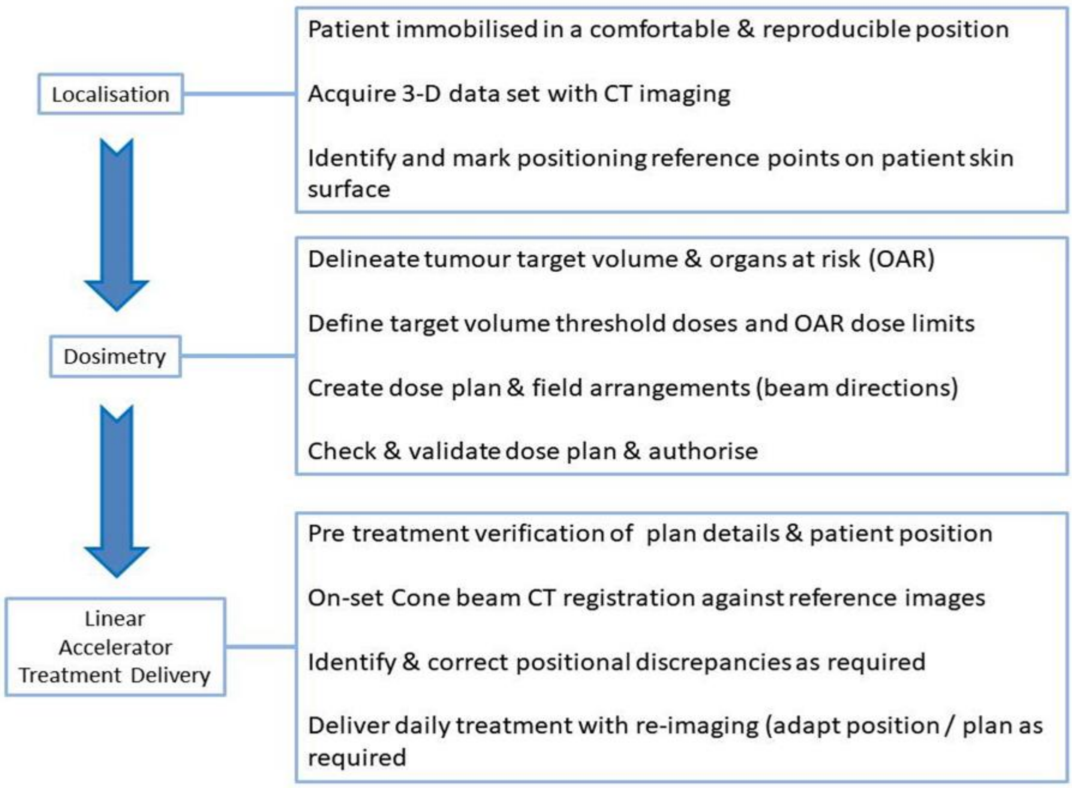


Figure 2.5: The relationships between the processes & activities in the radiotherapy workflow phases

2.4.1 The pre-treatment preparation phase

The pre-treatment phase begins with the identification of a stable and reproducible position for the patient which can be maintained throughout the course of radiotherapy treatment delivery. Known as immobilisation, the accurate positioning of the patient facilitates accurate field placement and dose delivery to the tumour target volume. Deciding on the most appropriate position requires mental visualisation of the position of the tumour and the relationships between this position within its organ of origin and the surrounding organs as summarised in figure 2.6 below. This is followed by CT localisation which involves the acquisition of 2-D cross-sectional CT X-ray data from which the size, shape and position of the tumour target volume and surrounding normal anatomy can be identified and reconstructed in 3-D. Following the delineation of the relative positions of the tumour target volume and the surrounding normal organs, optimal beam path directions, the size and shape of these beams and the radiation dose to be delivered by these beams is calculated in the dosimetry phase.

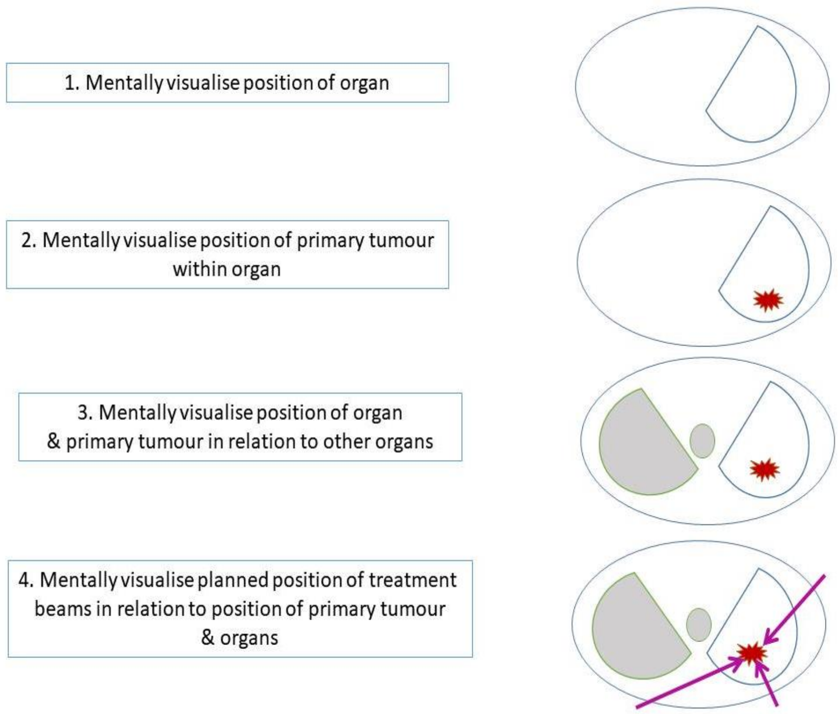


Figure 2.6: Schematic diagram to demonstrate the mental visualisation phases of the radiotherapy planning process

2.4.2 Determining the position of the tumour target volume

In order to deliver radiation treatment dose to these deep-seated tumours, their size, shape and position need to be identified. Prior to the invention of X-ray computed tomography (CT), the identification of the tumour target volume was based on 2-D X-ray images acquired in orthogonal planes, for example anterior and lateral images.

From these images, the position of the tumour was determined by its presumed relationships to skeletal landmarks as described by Thariat et al., (2013, p. 53). CT can demonstrate small differences in tissue contrast based on the density, atomic number and the number of electrons per gram of that tissue and can differentiate between disease processes and normal tissue. This tissue density information, known as absorption value, is reconstructed to form a 2-D image matrix which can represent the body in the transverse, sagittal and coronal planes. In radiotherapy it is most common to view images in the transverse plane and from the direction of the patient`s feet upwards (Bridge & Tipper, 2011, pp. 5-6). This is an important factor for new learners in radiotherapy to recognise in relation to, what Auer et al., (2008, p. 428) referred to as left – right discrimination (LRD). Two types of LRD have been described by Constant and Mellet (2018, p.1), the first is egocentric LRD and relates to the ability to discriminate between left and right from one's own perspective and secondly, allocentric or extra-egocentric LRD (Auer et al, 2008, p. 435) which relates to identifying features which are independent of and external to the viewer's position. Extra-egocentric LRD incorporates elements of egocentric LRD and mental rotation since most individuals, when considering features or objects external to themselves, will mentally rotate those objects in order to compare them with their own body parts. Viewing images from the perspective outlined above is a fundamental concept for learners to assimilate in relation to their own egocentric frame of reference and that of patient position and therefore the position and relationships of internal anatomy.

Following the increasing use of 2-D CT data for diagnosis in the late 1970`s, CT scanners were introduced into radiotherapy departments from the 1980`s onwards. As Thariat et al., (2013, p. 52) have observed, the addition of reconstruction algorithms to radiotherapy treatment planning computers led to the ability to visualise the soft tissue boundaries of the tumour within its organ of origin in 3-D. In combination with computer-based radiotherapy treatment planning algorithms it became possible to deliver dose to complex 3-D target volumes while further limiting the delivery of dose to normal tissue. Normal tissues which have sensitivity to radiation dose and therefore a limit to the amount of radiation that they can safely receive (known as a tolerance dose) are termed organs at risk (Berthelsen et al., 2007,

p. 109). Computed X-ray tomography has therefore become the standard imaging modality for radiotherapy planning due to its spatial resolution (the ability to differentiate between two tissue structures with different densities) and the availability of electron density data for each type of tissue. This is information which facilitates accurate dose calculation as identified by Høyer et al., (2011, p. 147). It is recognised that CT delivers a higher dose of radiation to the patient then 2-D imaging. Dose reference levels, based on surveys of median doses representing typical practice, are a quality assurance and improvement tool for controlling and optimising radiation imaging dose. The entrance surface dose for a conventional chest X-ray is 0.15 mGy / cm^2, while that quoted for a 4-D CT scan for planning radiotherapy treatment to the lung is has a dose length product of 1750 mGy/cm over a scan length of 34 cms (Public Health England, 2018). Nevertheless, the benefit of improved visualisation of the tumour and its relationship with normal anatomy outweighs the risk of any long term detrimental effect. In addition, the dose received from a CT planning scan is also much less than the treatment dose that a patient with a tumour in the lung would receive. For example, the radiotherapy treatment dose to the lung using a standard IMRT technique would be in the region of 50-55Gy[2], so it is important to acknowledge and be aware of this additional (concomitant) dose and the relatively small risk of subsequent development of a second cancer.

However, as Parodi (2017, p. 72) has identified, spatial resolution may not be sufficient for the definition of some tumour target volumes, with Bhide and Nutting (2010, p. 2) indicating that for tumours arising in the head and neck region, CT cannot detect microscopic extension of tumours with the same accuracy as magnetic resonance imaging (MR). This is related to the ability of MR to provide functional detail of metabolic processes which tend to be higher in regions of active cell growth (Høyer et al., 2011, p. 148). They also reported that an alternative to MR for the detection of cell proliferation and tumour metabolic activity is positron emission tomography (PET) which is also generally more accurate than CT for the delineation of the clinical extent of the tumour volume. The safe delivery of high radiotherapy doses with a steep dose gradient outside the tumour target volume is therefore predicated

[2] The dose prescribed for radiotherapy treatment is an expression of the amount of energy from the beam which is absorbed by the organ or region of interest (usually the PTV). 1 Gray = 1 Joule of energy / 1Kg of tissue

by the need for precise target volume visualisation and localisation using 3-D multi imaging modalities. There is, therefore, an associated need for radiotherapy radiographers to be familiar with the visualisation and pattern recognition of organs across a range of imaging modalities.

To support the delineation of the tumour target volume the International Commission on Radiation Units and Measurements (ICRU) have published recommendations for the identification and definition of the tumour target volume (ICRU, 1993, p. 5; ICRU, 2010, p. 46). The visible or palpable extent of the tumour is referred to as the gross tumour volume (GTV) and has been defined by Høyer et al., (2011, p. 147) as the extent of the tumour which is palpable by clinical examination or visible via any imaging modality. This volume will usually constitute the region of the tumour where the maximum concentration of cancer cells will be found. A further margin to include direct or local sub-clinical (microscopic) spread is added to the GTV. The GTV and this additional margin are known as the clinical target volume (CTV) which is determined by anatomical, topographical and biological factors relating to the stage of the tumour. For radiotherapy treatment planning purposes Hamilton and Ebert (2005, p. 456) indicate that a further margin of 4 – 10 mm is required to account for the variation in the size and position of the tumour related to patient movement due, for example, to respiration. The ICRU (1993, p. 16) refers to this margin as the planning target volume (PTV). For this programme of research, from this point forward, any reference to the tumour target volume will be synonymous with the PTV. Surrounding the PTV is a region referred to as the treated volume and is the volume that will receive a radiation dose that is appropriate to the intended outcome of treatment and usually identified as the volume receiving 95% of the prescribed dose. Beyond the treated volume is the irradiated volume which has been defined as the tissue volume which will receive a dose that is significant in relation to normal tissue. The ICRU (1999, p. 13) also reports that the size of the irradiated volume may increase as the number of beams increases. Shaping the radiation beam in 3-D CRT techniques to conform closely to the shape of the PTV with multi leaf collimators will reduce the size of both the treated and irradiated volumes. While the relative age of these publications is acknowledged, the fundamental concepts of target definition contained within them remain pertinent for current and emerging radiotherapy techniques (discussed above in section2.3).The relationships between these volumes in 2-D is shown in figure 2.7.

Associated with the identification of the target volumes identified above, there is also a need to identify surrounding anatomical structures on individual 2-D transverse CT slices in a process known as contouring or outlining, figure 2.8 shows an example of a single 2-D CT slice and the position of the prostate gland PTV. Outlining the PTV structure on a number of slices facilitates the reconstruction of these outlines into a 3-D model. By adding contours for additional anatomical structures (figure 2.9) This model becomes the platform for the development and calculation of the treatment delivery plan, as described by Bridge, Fielding, Pullar and Rowntree (2016, p. 38).

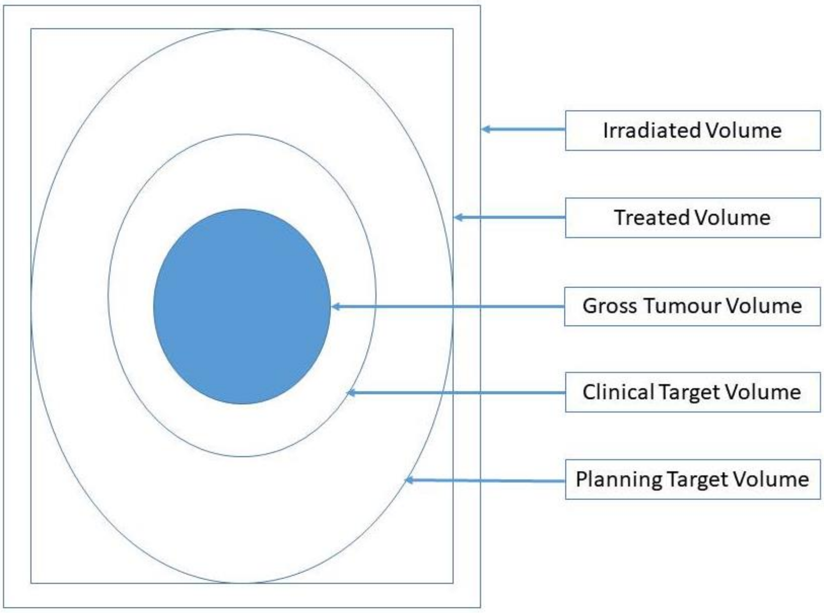


Figure 2.7: The 2-D relationships between ICRU tumour target volumes and the treated volume

The ability to mentally visualise and translate the appearance of 2-D figures or diagrams into a 3-D geometrical representation demands considerable conceptualisation as identified by Pittalis and Christou (2010, p. 193). The evidence base for the use of VERT™ in the visualisation of these concepts and its potential in supporting the development of 3-D mental model building skills is discussed in more detail in chapter 2.8, pp 50 - 73.

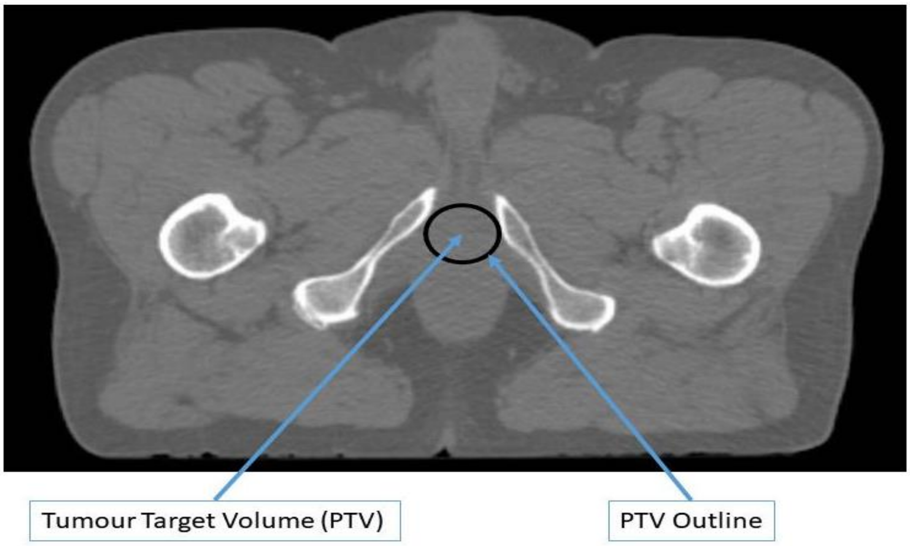


Figure 2.8: The planning target volume enclosed by the circular PTV contour

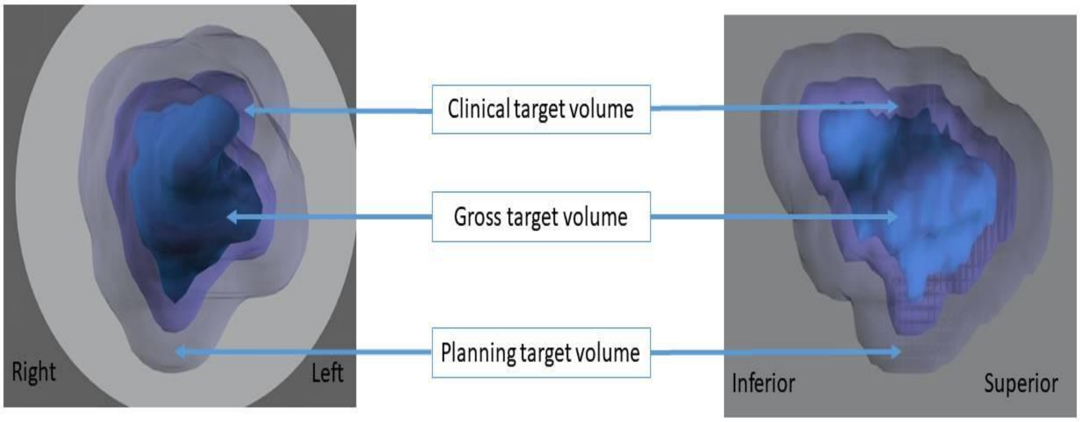


Figure 2.9: The 3-D relationships between target volumes and their margins

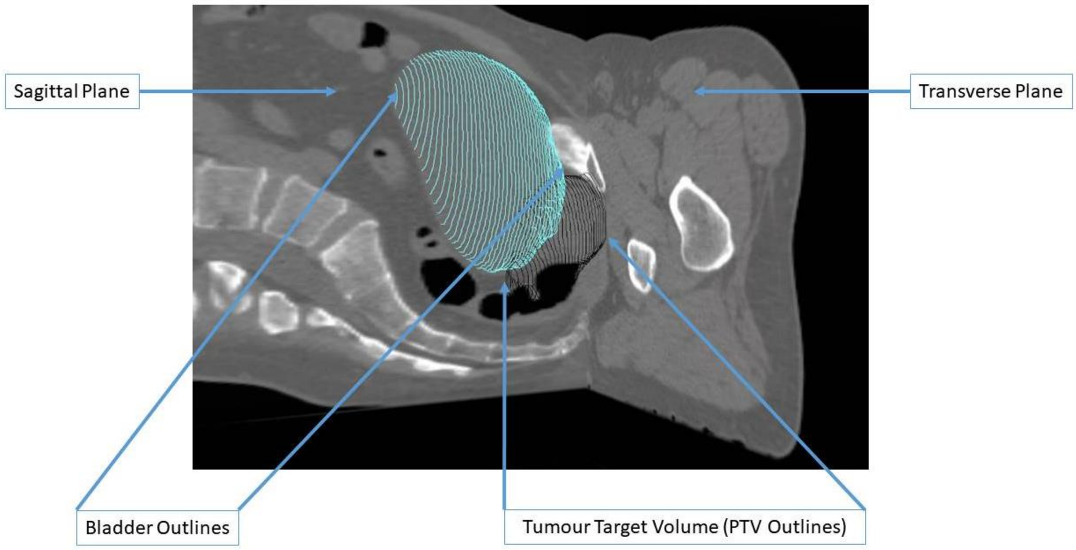


Figure 2.10: The relationships between prostate gland PTV and bladder outlines

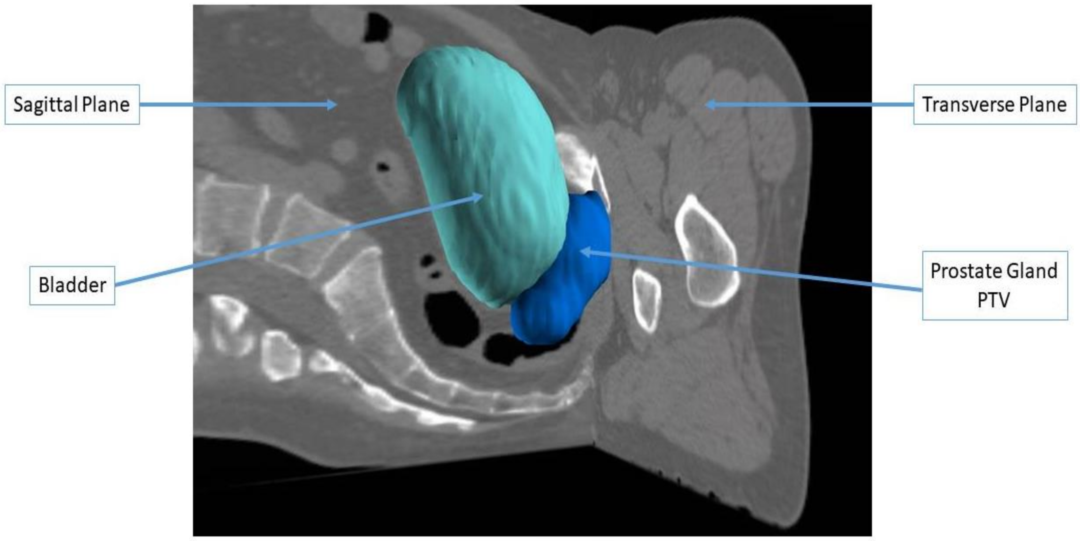


Figure 2.11: The outlines from figure 2.8 reconstructed as 3-D structures

2.4.3 The treatment verification and delivery phase

Prior to each treatment delivery session patients will be positioned on the linear accelerator treatment couch by utilising positioning and immobilisation instructions developed at the treatment planning stage. Following this, the treatment couch needs to be positioned so that the centre of the tumour target volume within the patient coincides with a point in space known as the isocentre. This is the point around which the linear accelerator gantry, the treatment head (field defining system) and the treatment couch will rotate and is located at a distance of 100 cms from the X-ray source as shown in a screen shot from the University of Portsmouth VERT™ platform (figure 2.12). The process of alignment has been identified by Sibtain et al., (2012, p. 161) as having two constituent parts. The first uses is a system of room lasers which are independent of the linear accelerator and project light in the X, Y and Z planes with the isocentre at the intersection of these planes as demonstrated in figure 2.12. The second set of visual indicators, shown in figure 2.13, is aligned to the linear accelerator isocentre. The beam definition system light field provides a visual indication of the size and shape of the radiation beam, its central axis coincides with the centre of the radiation beam and the optical distance indicator, calibrated to read 100 cms at the isocentre, provides a double check of the distance between the patient's skin surface and the radiation target.

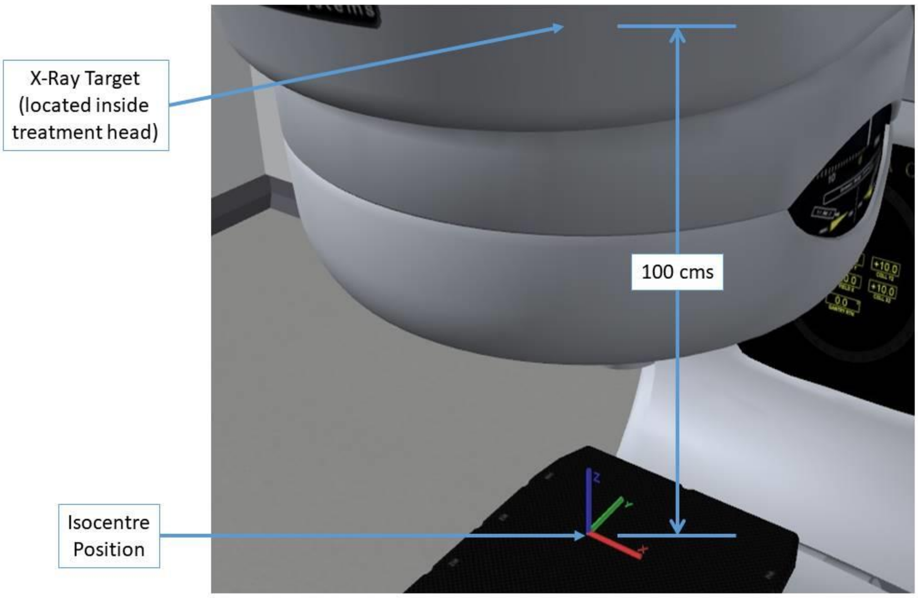


Figure 2.12: Screenshot from VERT™ demonstrating position of linear accelerator isocentre

A number of positioning coordinates on the patients' skin surface act as a surrogate for the centre of the tumour target volume within the patient. These will have been determined during localisation and target volume delineation and are aligned to the lasers prior to each treatment delivery. This facilitates correct tumour target volume alignment with the linear accelerator isocentre as described by MacDougal, Nalder and Morgan (2012, p. 115).

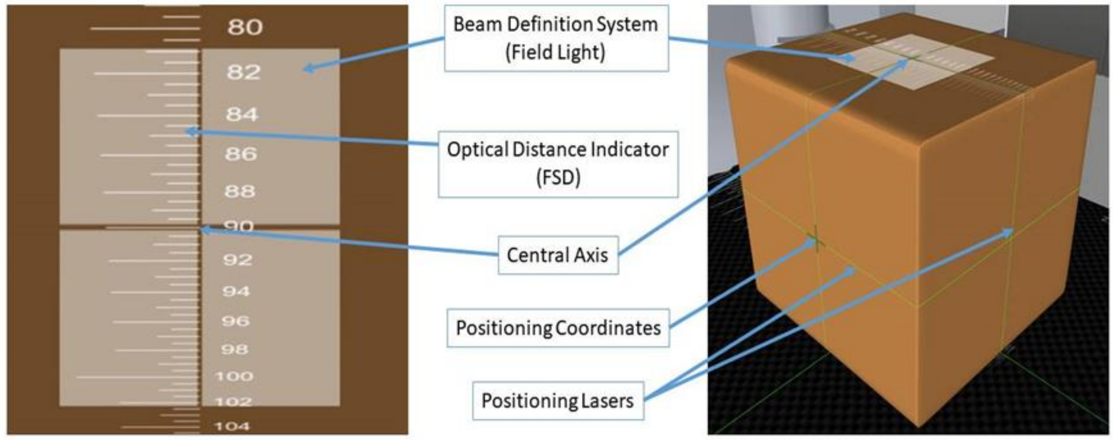


Figure 2.13: Screenshot from VERT™ showing the optical indicators which may be used to position the tumour target volume at the linear accelerator isocentre

Any deviation from the correct position will result in the potential for that volume to receive a dose that is lower than planned. This deviation in planned dose delivery to the tumour target volume is also associated with an increased risk of tumour regrowth due to geographic misses (set-up errors) which may occur as a result of incorrect patient position or incorrect isocentre position, beam size, shape and orientation. The impact of these set up errors may result in the likelihood of tumour under dosing (Royal College of Radiologists, Institute of Physics & Engineering in Medicine, Society & College of Radiographers, 2008, p. 11). At the same time, any reduction in the dose delivered to the tumour target volume will result in a corresponding higher than anticipated dose being delivered to the surrounding organs. This unplanned dose differential may result in the patient experiencing an increase in short term side effects or a longer-term risk of treatment related irreversible effects including the development of a second primary cancer.

Through mapping the sub tasks in the radiotherapy workflow which directly involved a patient or the manipulation of patient information such as CT images, Ford et al., (2009, p. 852) identified 269 activities. They referred to these activities as nodes, with each node identifying where actions are taken, decisions made, where data is manipulated and information is processed. This led the authors to conclude that radiotherapy is one of the most complex processes in healthcare. While the primary objective of the study was to identify the roles, responsibilities and inputs of the different staff groups within the radiotherapy multidisciplinary team, the findings would support the premise that clinical learning in radiotherapy takes place in an information rich environment.

2.5 The importance of spatial visualisation skill in radiotherapy

The effective delivery of curative radiotherapy is predicated by the successful integration of the physical principles of radiotherapy, the development of complex psychomotor skills and well-developed spatial visualisation skills. All of these factors are fundamental to safe and accurate decision making, patient care and treatment delivery. As treatment techniques for delivering higher doses to the tumour target volume have advanced, the underpinning concepts and fundamental principles supporting them need to be delivered from an early stage of pre-registration learning. As van der Merwe et al., (2017, p. 5) recently observed, while computer algorithms can support decision making, operators still need to have a high degree of pattern recognition skills so that they can differentiate between normal and abnormal appearances. This requires the front loading of visual and spatial relationship training from an early stage in their practice during pre-registration education programmes.

The impact of the 3 and 4-D processes required for IMRT, VMAT, SBRT and adaptive radiotherapy have changed the nature of practice and has resulted in a shift of professional role and responsibilities of radiotherapy radiographers as reported by White and Kane (2007, p. 298). This is particularly so in the case of IGRT image review and approval for geometric accuracy of the radiation beam direction, size and shape. Over recent years the responsibility for image review and approval has largely been transferred from the physician (who still holds responsibility for prescribing and approving the treatment course) to the radiotherapy radiographers delivering daily treatment (Gillan, Li & Harnett, 2013, p. 242). Adapting to this change and the introduction of advanced techniques into mainstream practice requires the possession of well-developed 3-D spatial visualisation skill to support mental model building and the mental manipulation and transformation of complex 3-D visual information which is based on CT data and knowledge of anatomical relationships in order to maximise optimal patient position.

New learners in radiotherapy are likely to be involved in the delivery of advanced treatment techniques from a very early stage of their education. This requires the integration and mental transformation of visual information relating to structural and relational anatomy and tumour size and shape into a 3-D structural framework or mental model. This framework provides the foundation to support the decision making skills required for the localisation of the tumour position, its relationships with adjacent normal anatomy and the subsequent accurate delivery of treatment. However, students must also develop and relate the complex psychomotor skills required for the manipulation and safe operation of the linear accelerator to the accurate positioning of the patient as identified by Bridge et al., (2007, p. 482). Linking these mental models to, for example, the relationships between the position of the target volume and the proposed X-ray beam path to real case interpretation of 3-D soft tissue data calls for well-developed spatial visualisation skills.

2.6 Radiation errors and their consequences

While patient safety incidents related to delivery of radiation dose in radiotherapy are rare, their consequences may have serious life-long consequences for patients. The "Towards Safer Radiotherapy" Report (The Royal College of Radiologists, Society & College of Radiographers, Institute of Physics & Engineering in Medicine, National Patient Safety Agency & British Institute of Radiology, 2008, p. 19) proposed a classification matrix to identify the level or severity of a radiotherapy error (RTE), together with a process code, hereafter referred to as a trigger code, which would identify where in the radiotherapy pathway the error had occurred. The radiotherapy workflow (section 2.4, p.) has been identified as having two primary phases, trigger codes 10 and 11 relate to activities in the pre-treatment phase and trigger code 13 relates to treatment verification and delivery activities. Recent analysis of RTE data, covering the period of December 2018 and March 2019 released by Public Health England (2019 p. 4) indicated that across the UK, 2,960 RTE`s were reported, analysed and categorised using the trigger codes summarised in table 2.1. Of these, 30 (1.0%) were classified as reportable radiation incidents under the IR (ME) R regulations, while 31 (1.1%) were non reportable. Of the remaining 2,899 RTE`s, 1,169 (39.5%) were related to treatment delivery processes. During the reporting period the estimated number of prescriptions (the number of treatment courses planned) was 49,148 equating to RTE`s being detected in 6% of those prescriptions.

Table 2.1: Radiotherapy error analysis and trigger codes (December 2018 – March 2019)

Error Type	Definition	Number (%)	Trigger codes and examples
Level 1 Reportable radiation incident	Significant accidental and unintended exposure	30 (1%)	10c, 11i, 13aa, 13g Incorrect localisation / delineation of tumour target volume
Level 2 Non reportable radiation incident	Any error which does not fit the definition for a reportable incident, but of potential or actual clinical significance. (NB while there is no legal requirement to report, notifying the statutory authority is viewed as good clinical governance)	31 (1.1%)	13aaa, 13l Examples include cases of under dosing for an entire treatment course as a result of incorrect delineation or daily positioning inconsistencies
Level 3 Minor radiation incident	A radiation incident posing no potential or actual adverse significance	958 (32.4%)	13aa, 13g, 13l, 13q Incorrect patient and / or equipment position
Level 4 Near miss	A potential radiation incident that has been detected and prevented before treatment delivery	753 (24.4%)	10f, 13aa Incorrect patient position
Level 5 Other non-conformance	Any non-compliance in following documented procedures not fitting any of the above criteria and not directly affecting treatment delivery (but if repeated may have an impact)	1188 (40.1%)	13aa, 13g, 13l, 13q As level 3 and covers failure to comply with treatment plan instructions

While the investigation and analysis of errors places emphasis on process failures and causative factors, those activities requiring aspects of spatial visualisation are not currently identified. It is possible, however, to link trigger codes to those areas of practice (known as process codes) where 3-D spatial visualisation skill has a role (table 2.2.

Table 2.2: Towards safer radiotherapy trigger codes and examples of associated errors

Trigger Code	Examples
10b, 10c, 13g	Incorrect positioning of patient for localisation and / or treatment delivery
10e,f	Production of images using inappropriate field coverage
10k	Incorrect translation of positioning marks on patient
11i	Incorrect identification and delineation of tumour target volume and organs at risk
13j,13l	Incorrect identification of & movement from reference marks
13m-13v	Incorrect setting & positioning of equipment parameters or failure to check automated processes
13aa	IGRT image approval (misinterpretation of normal /abnormal patterns

Failures in the treatment verification phase may occur due to incorrect organ delineation arising from poor or misinterpretation of left- right discrimination. While treatment delivery related errors may arise as a result of incorrect movements from tattoo or reference point to the isocentre position or inaccurate assessment of field placement leading to imaging errors. During this programme of research, the introduction of automated processes has led to the designing out of errors due to human factors engineering (Robson, Clark & White, 2014, p. 129). This has been evidenced most recently by the development and introduction of Halcyon 2 (described in chapter 2.3.5, p. 34). Automation, discussed in more detail in section 2.7 below, offers the potential to optimise and streamline workflows and improve performance, but it is important to also recognise that automated systems can fail and without experience it is difficult to recognise that an error has been made. This is likely to be the case if the system has been seen to be safe and reliable in the past. This over reliance on automated processes has the potential to impair expertise in specific tasks, for example, the process of setting treatment machine parameters. These include field size and beam shaping where automation has replaced manual setting so that radiotherapy radiographers are no longer required to perform these tasks on a regular basis. It is incumbent on educators, therefore, to instil in learners a moment to moment appreciation of the potential for failures within the pathway, a condition that Mazur et al., (2018, p. 198) refer to as safety mindfulness. They also identify that this training can be provided through simulated interactions which replicate the scenarios and cognitive effort and procedural compliance required in the clinic. This can allow learners to gain the necessary awareness of procedure expectations and to recognise where failures may occur, a theme which has been explored by Beavis and Ward (2014, n. p.) and discussed further in section 2.8, p. 59 below.

2.7 The impact of automation

In addition to 3-D spatial visualisation, optimal patient care in radiation delivery requires radiotherapy radiographers to have the knowledge and skill to support independent clinical judgement and decision making. However, from a review of the literature, Lozano (2011, p. 1) identified that developments in technology during the last two decades has resulted in a widening chasm between conceptual knowledge and radiotherapy practice. Two forces that have tended to act against each other have driven this, the first being that linear accelerators now require fewer hands on operations compared to previous generations because of automation. This has resulted in radiotherapy radiographers now performing tasks more aligned to system programming. The second is that increasing the level of computerised control has increased the level of precision and an associated reduction in the margins between the tumour target volume and the surrounding healthy tissue. During the course of this programme of research, the evolution of automated processes has continued. So while automation has the potential for safety improvements by removing the risk of human error, it has also moved the emphasis away from hands on (psychomotor) positioning of the patient at the linear accelerator isocentre, to a focus on organ position and target volume movement which requires pattern recognition skills. As automation now covers more of the radiotherapy work functions, Lozano (ibid) suggests that there is a risk that individual radiographers may abandon those functions traditionally done by hand and the repeated rehearsal of the tasks associated with them. This may result in a loss of the principles and understanding of these activities over time, VERT™ provides one way of ensuring that these automated processes are understood. However, the use of and experience with technology not only has an impact on the individual but also learning outcomes and the process of learning (Cilesiz, 2011, pp. 487-488). Therefore the following section will review the evidence base for the use of VERT™ and its role in teaching, learning and assessment.

2.8 The VERT™ research base

The research conducted with VERT™ falls into three distinct phases, the development phase, the introduction and evaluation phase and the research conducted during this programme of research. This section will review VERT™ publications in date order and will predominantly cover reports of primary research, findings from national and international user surveys and educational notes. The section will conclude with a summary of the key themes, findings and recommendations to date.

Bridge et al., (2007)

Conducted during the development phase of the VERT™ prototype, the primary aim of this study was an initial evaluation of the platform. Taken from the perspective of the impact of VERT™ on student confidence, it sought to determine the extent to which VERT™ might enhance students' knowledge and understanding of complex radiotherapy concepts associated with a 3-D positioning task. This task involves the manipulation of the treatment couch and linear accelerator along and rotated around the X, Y and Z planes so that the radiation beam will enter the patient perpendicular to the surface of the skin. Success in this positioning and the achievement of a clinically acceptable set-up calls for a combination of good spatial awareness and 3-D spatial visualisation skill. The study also sought to gain information concerning the ease of use and realism of the application from a learner perspective in order to guide further improvements. Finally, the study aimed to make recommendations regarding the platforms impact on future curriculum design and teaching, learning and assessment strategies.

Forty-two first year pre-registration radiotherapy students, (male = 14 [33.3%], female 28 [66.7%]) with a mean age of 29 (range 19 – 51) and five weeks experience of working on linear accelerators in the clinical environment took part in the study. They completed a pre task self-assessment questionnaire using a 5-point Likert response to provide quantitative baseline data for age, gender and confidence relating to three learning outcomes: understanding the skin apposition technique, understanding how to apply it and confidence to assist clinically. The use of the same learning outcome questions in the post task questionnaire would gauge if learning had taken place. Additional Likert-style questions were included to collect data regarding platform realism, ease of use, level of interaction and enjoyment. Open questions were included to collect qualitative data regarding suggestions, problems and perceived benefits of the application. Comparison of responses in the pre & post questionnaire for impact on achievement of learning outcomes demonstrated that students felt that they had improved their understanding and confidence in their technical skills after using the platform. The mean student confidence with the skin apposition technique overall was reported as 51.8% before using the application and rose to 73% after using it ($p<0.00001$). From the perspective of realism and ease use, 37 participants (88%) reported that they found the application to be realistic and 29 (69%) indicated that the controls were easy to master with six (14%) being undecided and seven (17%) reporting difficulty with them. Other participant comments suggested that to take their time without fear of worrying about harming the patient or delaying the treatment machine.

The level of perceived realism correlated with both student performance and enjoyment irrespective of age and gender, leading to the conclusion that the virtual linear accelerator had increased understanding and confidence, with the authors suggesting that prior practice of these skills in a virtual reality environment would enable students to be able to set patients up with increased confidence in a shorter time. They also advocated its use in orientation for clinical placement and to support academic teaching. They did however add the observation that, for students who are unwilling to engage with the platform or who have difficulty mastering the controls, would be at a disadvantage and may perform less well in the virtual environment than they would in the actual clinical environment.

Flinton and White (2009)

While the authors acknowledge the potential for VERT™ to realise several benefits within radiotherapy education, evidence from other sectors where simulation platforms are used, for example flight simulators and IMAX™ cinemas, would suggest that simulator sickness may be a limiting factor for participant engagement. As users of VERT™ are immersed in a 3-D virtual environment with a wide field of view and moving images, they may experience illusory feelings of self-movement and symptoms which parallel those of motion sickness minus emesis. The prevalence and severity of these symptoms, known as simulator sickness, are linked to the presence of the environment (the extent to which participants' senses are engaged by the experience). Participants who encounter more symptoms tend to be more distracted, less engaged and experience less presence than those who are fully immersed and involved. Therefore, the primary purpose of this study was to determine if a relationship existed between the side effects experienced by users of VERT™ and its presence.

The study recruited 84 participants from two English HEI`s who use the back projected, immersive 3-D VERT™ system who were evaluated prior to using VERT™ for their current state of health, medication, sleep patterns and previous history of travel sickness, all factors which may influence simulation sickness.

Appleyard and Coleman (2010)

Conducted as part of the Department of Health (England) VERT™ evaluation project, 103 pre-registration radiotherapy students participated in a stratified randomised controlled trial study to assess the influence of both VERT™ tracking technology and 3D stereoscopy on performance of skin apposition techniques. Performance was determined using what the authors refer to as an objective measures schedule and an accuracy tool integral to the VERT™ software. The Vandenberg and Kuse Mental Rotation Test was used to determine what the authors refer to as spatial ability although no detail was provided regarding its administration or scoring convention. A post-experience questionnaire was used to determine the students' experiences of the virtual environment and covered the extent to which they felt that it had enhanced their clinical practice and any adverse effects that they may have experienced. Follow-up interviews after relevant clinical placement experience also explored the extent to which practice in virtual reality was transferrable to the clinical environment. As the recruitment strategy was not explained, it is not clear where these students were studying or to determine their level of study.

Participants were randomised (although the method is not stated) to one of the three operating modes of VERT™, namely 3-D stereoscopy on and tracking on (referred to as full immersion), 3-D stereoscopy on but tracking off (3-D immersion) and its 2-D mode with no immersion (3-D stereoscopy off and tracking off). All participants then received guided practice in the technique prior to objective assessment of their ability to accurately and efficiently complete the task. Accuracy and efficiency were determined by economy of movement (the number of equipment moves required to achieve what the student deemed to be an acceptable set up), the degree of skin apposition (the standard deviation of the distance between each of corners of the applicator and the surface of the virtual patient the number of errors (collisions between equipment and patient and incorrect beam alignment) and the time taken to achieve a set-up acceptable to the student. Observations were made by an experienced radiographer whose score was used as a benchmark, against which student performance was normalised.

The findings demonstrated that there were differences in the mean set up scores between group two, who used 3-D immersion and group one, the full immersion group (Mann-Whitney Test 9.56 (± 12.2 at 95%CI), p=0.17). The difference in the mean set up scores between group two (3-D immersion) and group three (2-D with no immersion was 9.54 (± 13.2 at 95%CI), p=0.22. While the authors reported that there were no significant differences identified in efficiency across groups, those students in the full immersion group were significantly worse at aligning the light field to the skin marks compared to those in the 3-D immersion group (p<0.002). Students attributed this to difficulty in being able to position themselves closely enough to the virtual patient in order to visualise alignment accurately. Qualitative analysis indicated that students in this group found completion of the set-up more challenging although there was no statistically significant difference in set-up scores between the three arms of the trial. However, it was acknowledged that the students in groups two and three had their view manipulated by the observer. Although no verbal guidance was offered, they used their own experience to intuitively adjust the view for the student. This guided the student as to where they should be looking and provided clues regarding action required.

Follow-up interviews explored the extent to which the type of VERT™ experience influenced the transfer of skills to real world set-ups. Students from the full immersion group were more positive about the speed with which they felt able to put their VERT™ experience into practice. However, the key theme from the interviews was the concern, expressed by most students, that the experience had not fully prepared them for the real set-ups. While all students enjoyed the virtual experience and recognised that it helped them to achieve acceptable set-ups, the majority reported that the situation in the real clinical environment was very different. While confidence increased as a result of VERT™ experience, anxiety in real world situations only lessened through real world practice. They highlighted daily variations in position and patient breathing as examples of where VERT™ had not adequately prepared them for the need to adapt. This observation echoes the theme of realism identified by Bridge, Appleyard, Ward, Philips and Beavis (2007) above. Students also identified that objective assessment of their performance helped to improve their skills substantially whether they had experience of these set-ups or not. Many students, but particularly those with some clinical experience of skin apposition techniques, suggested that practice in VERT™ facilities before and during relevant placements would be very beneficial. The authors considered that comments from students with poorer spatial ability were pertinent in relation to the benefit of VERT™ in developing strategies for skin apposition techniques. The authors report the following comment as typical:

> *"My spatial awareness is terrible and that probably explains why I've shied away from getting more actively involved with electron set-ups. I just can't see how gantry and couch need to be moved. Spending time in VERT has really helped. I wish it had been there when I was in year 1."*

It is interesting to note, that given the participant observation quoted above and the weak positive correlation between the MRT score and the positioning score ($r=0.494$, $p<0.01$), no further analysis of mental rotation performance was conducted. This would have been helpful given that one of the key recommendations from the evaluation project was the routine measurement of the inherent spatial ability of pre-registration students to determine those students who are likely to benefit most from using VERT™. Study findings also suggested that the strategies and psychomotor skills required for achieving good skin-apposition could be learnt effectively in VERT™ and that the objective assessment of this technique using VERT™ could also lead to improved skill. In relation to the different visualisation modes available, 3D stereoscopy and user tracking did not appear to influence student performance or experience. While tracking appeared to more accurately reflect the actual clinical situation, the authors concluded that it may have detracted from students' ability to accurately visualise the alignment of the light beam with the skin marks on the surface of the virtual patients.

Shah and Williams (2010)

The increasing use of CT anatomy in radiotherapy treatment planning and IGRT and the introduction of VERT™ as a visualisation platform into radiotherapy clinical departments led the authors to explore the role of VERT™ in facilitating CT anatomy refresher training for qualified radiotherapy radiographers. In small groups (size not specified), 29 staff attended a one hour practical session during which they were first asked to label hard copy transverse plane CT slices taken from the head and neck thorax and pelvis regions. Following this exercise, the same slices and structures were viewed on the VERT™ screen together with a combination of multi choice questions. Interactive cue cards were used by each participant to indicate their answers. While the method of session evaluation was not identified, the authors reported that all participants found the session beneficial in relation to visualising the size and location of anatomical structures, especially organs at risk. The location of organs in relation to surface anatomy and CT slices was also reported to be useful, with interest centring on the head and neck region particularly. The conclusion was that VERT™ was a useful continuous professional tool in the post graduate setting, with the authors proposing that further resources and case studies be developed to link radiotherapy treatment side effects to anatomical visualisation to multidisciplinary training.

Green and Appleyard (2011)

The stated aims for this factorial design, randomised controlled study were to determine the impact of the VERT™ visualisation modes on psychomotor skill, skill in applying skin apposition techniques, level of student confidence in setting up these techniques in VERT™ and transfer to the clinical situation. The methodology was similar to that reported by Appleyard & Coleman (2010), with the authors reporting that factorial design facilitated the comparison of the three visualisation modes available in VERT™. These were identified as 3-D stereoscopy on and tracking on (referred to as tracking [T]), 3-D stereoscopy on but tracking off (non-tracking stereoscopy [NT, S]) and 3-D stereoscopy off and tracking off (this is the 2-D mode identified as NT, NS). It should be noted that the visualisation modes are the same as those studied by Appleyard & Coleman but have a different notation. Spatial ability was measured using the Vandenberg and Kuse MRT to determine its impact on performance in the technique. All year 1 and 2 pre-registration radiotherapy students (n=93) at a single UK HEI were initially invited to take part. A total of 44 students (response rate 47.3%) were recruited, with 23 (52.3%) from year one and 21 (47.7%) from year 2. The imbalance between male n = 11 (25%) and female (n = 33, 75%) participants and the small overall population was acknowledged.

All participants had an initial group demonstration of the principles of the skin apposition technique using the VERT™ system prior to carrying out the same scenario to enable even comparison. Performance outcome measures for the number of equipment movements, time taken to complete the setup, accuracy of skin apposition and accuracy of overall performance (to include number of alignment errors and collisions) were recorded on an outcome sheet during the respondents' individual sessions and on completion of task via an accuracy tool integral to the VERT™ software.

Results from this phase of the study reported that, using the Mann Whitney U test, no statistical significant difference between set-up score across the three arms was established. The mean set-up score differences between the differing randomised arms are as follows: T and NT, S = 7.82 (with a 95% confidence interval of 19.76), p = 0.87 and the NT, S and NT, NS = 11.42 (with a 95% confidence interval of 21.80), p = 0.50 (Mann Whitney U -Test). Participant performance in the MRT was not reported for the study cohort as a whole or by gender, but by randomisation arm. This is likely to make comparisons with the findings of other mental rotation studies challenging. The authors did report that there was a significant statistical difference p = .018 (Kruskal - Wallis test) was shown between the 3 arms although the unequal size and gender grouping in each arm should be noted. Further analysis of the differences between the mean mental rotation score across the three arms showed that there was non-significant difference between the tracking and non-tracking stereoscopic arms = .2 (with a 95% confidence interval of 1.99, p = .87, Mann - Whitney U Test). However, there was a statistically significant difference between the NT, S and NT, NS = 4.5 (with a 95% confidence interval of 1.97, p = .03, Mann - Whitney U Test). However details relating to the MRT timings and scoring method were not reported and there were no details relating to how participants were randomised to each arm. In addition, a comparison between mental rotation score and positioning performance produced a moderate positive linear correlation was established with Pearson r test score $r = .343$, $p = .023$.

As the study also aimed to determine transfer of skills from VERT™ to the clinical setting, six participants took part in post clinical experience interviews. While the small number of interviewees is acknowledged, four of the six respondents indicated that VERT™ had improved their skills and confidence with electron set-up. Key phrases identified by the authors such as:

> *'Practice on a patient who wasn't actually a patient'* and *'no worries about endangering a patient in the real world'*

This led the authors to conclude that the safety factor of using a simulated patient was a key issue in how skills and confidence were improved. As this study is similar in design to that of Appleyard and Coleman (2010), discussed above, a comparative summary of the findings from both studies is presented in tables 2.3 a, b and c, below.

Table 2.3a: Demographic profile comparison for the VERT™ studies conducted by Appleyard & Coleman (2010) & Green and Appleyard (2011)

Demographic Profile			3-D Tracking	3-D No Tracking	2-D
Appleyard & Coleman	Participants		36	35	32
Green & Appleyard			13	15	16
Appleyard & Coleman	Gender		M = 11	M = 8	M = 9
			F = 25	F = 27	F = 23
Green & Appleyard			M = 5	M = 3	M = 3
			F = 8	F = 12	F = 13
Appleyard & Coleman	Age	Mean	23.7	25.1	23.3
		SD	7.2	7.9	6.4
Green & Appleyard		Mean	23.5	25.7	24.9
		SD	6.3	7.4	8.1

Table 2.3b: Comparison of mental rotation scores and VERT positioning task performance for the studies conducted by Appleyard & Coleman (2010) & Green and Appleyard (2011)

MRT Performance Scores	(24 items)	3-D Tracking	3-D No Tracking	2-D
Appleyard & Coleman	Mean	10.8	11.5	9.0
	SD	4.3	5.1	5.6
	Range	55-151	68-204	43-158
	Conversion	45%	47.9%	37.5%
Green & Appleyard	Mean	11.3	11.0	7.0
	SD	4.7	5.8	5.1
	Range	Not reported		
	Conversion	47.1%	45.8%	29.2%

Table 2.3c: Comparison of VERT positioning task performance for the studies conducted by Appleyard & Coleman (2010) & Green and Appleyard (2011)

Positioning Task Performance Score		3-D Tracking	3-D No Tracking	2-D
Appleyard & Coleman	Mean	98.5	108.6	98.5
	SD	23.2	28.9	26.3
	Range	55-151	68-204	43-158
Green & Appleyard	Mean	99.3	107.1	95.7
	SD	18.2	33.8	27.6
	Range	70-135	68-204	43-138

These findings will be revisited as part of the triangulation and discussion of the results of study five in this programme of research (chapter 6.6.4, p. 234).

Nisbet and Matthews (2011)

In a review article, the authors discuss the development and introduction of a VERT™ clinical workbook across six radiotherapy departments. The overarching aims behind its introduction were identified as the need to ensure parity of clinical education, to enhance the learning experience and to integrate theory with clinical practice. The authors identified this integration as an essential component for connecting the underlying theoretical principles with day to day clinical practice because positioning a patient and the equipment prior to radiotherapy treatment delivery does not involve thinking about the theory before putting it into practice. Rather, it is an integrated process of knowledge-in-action, much of which Nisbet & Matthews identify as spontaneous and tacit. While they did not provide a definition for tacit knowledge, Reinders (2010, p. 32) has differentiated it from what was termed "the knows what" and referred to a personal dimension involving statements such as "I have a feeling that………" or "something tells me that………" (also see chapter 3.5.1, p. 83). The reasoning behind such statements may not be explained, but can be part technical and partly based on beliefs, perceptions and mental models. This personal dimension means that tacit knowledge cannot be easily transferred in the same way that procedural knowledge can be. As such the VERT™ workbook was designed to provide practice problems so that learners could begin to solve them in a safe, structured and positive environment where mistakes could be made, corrected and learned from them with no external pressure. This environment would also support the learning and practising relevant technical skills; development of independent thinking and problem solving approaches; and the supporting skills of team working, collaboration and communication.

The review identified the key components and aims for each year group, which aligned to Bloom's taxonomy of learning, beginning with the development of technical and psychomotor skills for the manipulation of radiotherapy equipment and the knowledge and understanding of commonly used radiotherapy techniques in year one. As learning progresses from the lower order knowledge building skills to the higher level domains of comprehension, application and analysis, the discussion of routine cases techniques supports critical reflection, clinical reasoning and evaluative skills which in turn assist in the development of confidence and competence when undertaking commonly performed radiotherapy techniques. Students with an established knowledge of a technique can then be challenged with further case-studies that develop the skills of critical analysis and evaluation, for example, evaluating different techniques for the same sites of disease. Finally, in year three, emphasis turns to professional development and the transition from learner to practitioner by providing a range of problem based clinical decision making exercises. While the authors reported that early indications of student experience via a questionnaire were very favourable, with high student evaluations for session content and learning gained, no further detail about how many students provided feedback or analysis of their demographic profile and level of study was provided.

James and Dumbleton (2013)

The authors report on a survey undertaken during 2011 to evaluate the utilisation of VERT™ in clinical radiotherapy centres across the UK. Using a quantitative methodology, a 45 question online survey requiring yes/no responses, multiple choice responses and the submission of numerical data was circulated via an email link to all 67 radiotherapy service managers. A total of 53 centres (82.8%) responded, with 43 from NHS England, six from Scotland, Wales and Northern Ireland and four independent providers responded with one replying on behalf of its four clinical sites. Eleven centres across the UK failed to provide data. There were 27 centres (51%) with Seminar VERT™, three centres with Immersive VERT™ and 20 centres (38%) across the UK with no VERT™ installation with fourteen of these in England. The results indicate that the use of VERT varies considerably across radiotherapy centres. This ranges from centres not using the system at all which was implied from their declaration that they did not use any software licences regularly, to multi-purpose usage covering induction and training, introduction of new treatment technologies for radiotherapy staff, education of patients, carers, GPs, commissioners and other hospital staff and the promotion of radiotherapy at careers fairs and staff recruitment events.

The survey identified that the most frequent use of VERT™ was for the training of staff specifically to support the training of pre-registration therapeutic radiography students. The authors reported that this was expected since funding was provided by the Department of Health England as part of the strategy to improve the retention rates of pre-registration therapeutic radiography students during their training programmes in the longer term. However they also highlighted a concern that just under a third of centres (number not specified) were not using their VERT™ system for this training purpose, when remembering that the funding of these VERT™ installations in England came from public monies provided by the DH. The report concluded that the varied use of VERT™ in radiotherapy centres across the UK, while supporting many of the findings of the initial VERT™ evaluation project, resulted in maximum benefit of the VERT™ installations across the UK was not being fully achieved in clinical radiotherapy centres. In the light of these findings the authors recommended that radiotherapy service managers should review the use of VERT™ in their centres, consider increasing the level and diversity of their VERT™ activities and to commit adequate resources to develop and implement VERT™ fully and effectively.

In doing so, the authors suggest that its full potential may be realised for the benefit of the profession and the service in its entirety. It is interesting to note that a follow up survey to determine progress has not been conducted and reported.

Beavis and Ward (2014)

In this article the authors discuss the use of VERT™ for the modelling of a range of linear accelerator calibration error conditions and the visualisation of the impact on dose delivery for the patient. This was achieved through what they refer to as user case scenario activities with correct and incorrect parameter settings. While they indicated that this was primarily aimed at trainee physicists and those working in dosimetry, the platform can also assist with the understanding of the underpinning concepts of treatment planning. They conclude that simulation training with VERT™, in addition to modelling radiotherapy concepts and workflows, the platform can be used to simulate errors and process failures and allow participants to examine such scenarios with zero risk to patients or staff even if a miss-calibration is intentional. This approach can also facilitate the sharing of experiences gained over many years of clinical work by experienced professionals in order to develop safety awareness in trainees.

Flinton (2015)

In a single centre, mixed methods, randomised cross over study Flinton compared the performance of 52 pre-registration radiotherapy students in a complex simulated positioning task conducted in VERT™ and in the radiotherapy department using a real linear accelerator. The interim analysis of quantitative performance data, published in (Flinton, 2013, p.172), indicated that the accuracy of set-up favoured the real situation with students gaining higher performance scores in the clinic compared with those gained in VERT™. Performance under the two conditions was significantly different with a mean performance score of 5.23 using the real machine and 3.62 ($p < .0001$). However no specific detail on how performance was scored was provided. Further analysis using McNemar's test demonstrated that the two tasks are performed differently with a difference of 35.6 (CI 18.6 - 39.89), $p = .0001$.

Further analysis of the performance of these 52 participants (14 male, 26.9%) who were recruited from all three year groups at a single UK HEI, demonstrated that over two consecutive set ups on the real linear accelerator performance scores increased but the same effect was not observed when using VERT™ with no significant difference between year group suggesting that experience with the technique was not an influencing factor. Additional qualitative feedback from focus groups suggested that the lack of tactile feedback in VERT™ coupled with a limited viewing angle and poor fit of the 3-D glasses were reasons offered for a less favourable perception of the platform. Participants did however indicate positive feelings about its use as a training tool to support assessment preparation by allowing them to work in an unhurried environment and to learn from mistakes. Other observations related to the ease of use of the real linear accelerator controls compared to VERT™ and the lack of a sense of reality in VERT™, it didn't feel, look and sound like the real department. Because the virtual patient was not recognised as a real patient participants did not consider vigilance and safety to be important. In regard to the learning opportunities delivered by VERT™, observations related to the ability to interact with a real hand pendant and confidence gained led to an improved performance in the clinic setting. However participants felt that the lack of realism precluded the use of the platform for assessment of competence. Other performance measures demonstrated that there were no significant differences in completion time although males were three minutes quicker in the real setting. Conversely, performance scores for females were better in VERT; this was an interesting finding, given the commonly reported gender differences favouring males in visualisation tasks. This led Flinton to conclude that, in this study, while the low number of males is acknowledged, there was no support for gender differences which the author suggested may have been removed through training.

Kirby (2015)

Kirby identified that some of the technical aspects of radiotherapy physics can be difficult to acquire since they are not practical experiences usually encountered by students' first-hand during clinical placements. Using a combination of small group revision lectures, tutor led demonstrations and practical experimentation with second year undergraduate and first and second year postgraduate pre-registration radiotherapy students, the study aimed to evaluate and share experiences of virtual dosimetry experiments using the VERT™ Physics module. In small groups, with a maximum of seven participants, the revision lectures were designed to assist the recall of foundational physical concepts. This included coverage of the inverse square law and the dose delivery consequences of positioning a patient at the incorrect distance from the radiation source, measurement of percentage depth dose and the data required for dose calculation. This session was immediately followed by a tutor led demonstration of the quality assurance tools available within the VERT™ physics software. This provided instructions for the use of the equipment alongside the virtual linear accelerator. Following this session, each group was then split into a measurement group who would conduct each experiment and a calculation group who would conduct manual calculations for that experiment. On completion of each experiment, the subgroups would swap roles. The measurement group would do manual calculations to check experimental findings, with the calculation group conducting the experiments to confirm their manual calculations.

Once all the experiments had been completed by all groups, post session feedback for the most positive and least positive aspects of the sessions and suggestions for future sessions was collected. Although the number of respondents was not indicated, the responses are reported as being heavily weighted towards the positive side, with 10 times as many positive comments as less positive comments.

Montgomerie, Kane, Leong and Mudie (2016)

In this editorial note, the authors discuss the approach to supporting the development of conceptual knowledge of radiotherapy principles with VERT™. The focus of the article is the integration of VERT™ across the curriculum and how it may be used to support traditional delivery methods. While the authors point to the link between conceptual knowledge and application of practical skills by referring to anatomy and imaging and that many students find the building of a 3-D understanding of anatomy to be challenging, their evidence is anecdotal. They also identify that the development of spatial awareness to support mental visualisation of 3-D perspectives in learners takes time to develop. They proposed that the linking of VERT™ to a radiotherapy treatment planning system can be a method to conceptualise the appropriate choice of radiation field size, shape and direction in relation to the target volume and organs at risk.

They concluded that the introduction of VERT ™ to their institution in 2013 did not change the taught curriculum but it did enhance delivery through the visualisation of complex ideas and techniques in an environment comparable to the clinical environment.

Stewart-Lord (2016)

In this educational note aimed at sharing the experiences of one HEI, the author provides an overview of the integration of the VERT™ platform into the radiotherapy training curricula. The challenge of embedding the platform into existing teaching was highlighted, although the author pointed out that initial student feedback on experiences supported the development of new training resources. One of the areas of integration focused on the viewing of CT images displayed in axial, sagittal and coronal planes for second year students. Each practical session was developed for a specific anatomical region, for example head and neck, thorax and pelvis with a focus on relational cross sectional anatomy, critical structures and radiotherapy tolerance doses. Each session was supported with practical work sheets for structure labelling and feedback from end of year module evaluations was reported to be positive. The next phase focused on the use of VERT in pre-clinical induction weeks during which a range of practical activities were undertaken in preparation for student placements. These included: hand dexterity with the pendant; adjustment of gantry positions; understanding the relationship between couch movements and digital displays, calculating and making isocentre shifts from the reference marks; reading FSDs; setting up a virtual patient in groups and avoiding collisions. Workbooks and activity sheets facilitated the sessions and supported individual work, one-to one tutorials as well as peer-supported group sessions focussed on problem solving activities. Additional developments for second and third year students included the use of VERT to evaluate treatment plans produced by individual students in the treatment planning laboratory. Completed plans were exported from the treatment planning system to VERT for viewing. Group evaluation and critical reflection were reported to improve confidence in plan interpretation and evaluation during new patient treatment set-up in the clinical department. While the conclusions suggested that introduction of VERT as an education tool had enabled academic staff to develop a range of teaching methods, much of the supporting evidence for its use was drawn from other studies. Qualitative evaluation from student feedback was alluded to, with comments such as "improved communicating and engagement of year two and three students", but not reported in detail. This would have been interesting to see and to compare with the experiences of other users.

Bridge et al., (2017)

The aim of this study was to establish the pedagogical role of VERT™ and the potential for its future role in collaborative research and development. An 18 item Survey Monkey® questionnaire comprising multi-option and short answer open questions covering hardware and software provision, current use and suggestions for future role was circulated to 52 worldwide users. Quantitative data was collated and analysed using the descriptive tools within Survey Monkey®, while responses to the 11 short answer questions were subjected to a conceptual thematic analysis.

The overall return of 47 surveys showed a response rate of 90% and responses to the 11 short answer questions provided 105 clarification comments for triangulation against the quantitative data. The most common activities identified were use of the platform with pre-registration students to support their knowledge building and understanding of fundamental radiotherapy concepts. Other less common themes or activities included treatment delivery plan evaluation, physics principles and equipment quality assurance and multidisciplinary teaching and continuous professional development. For a typical 30 week academic year, 37 respondents (78.5%) indicated a usage of one day per week or less, which would suggest an estimated 8000 hours worldwide annual use. Most activities were conducted in small group seminars or paired peer to peer learning.

Of note and perhaps somewhat disappointing, was that 15 respondents (32%) indicated that there was a perception within their institutions that VERT™ was not useful. While it was not clear how this view had arisen, one respondent did report that a lack of use had left staff with low confidence when engaging with the platform. Another view suggested that VERT™ was being perceived as the answer to everything and that students were just being taught differently. That said, the authors reported that 30 users (64%) were supportive of collaborative resource development. A key area for research collaboration was that of the platform's impact on learning and assessment strategies using both quantitative and qualitative measures.

Overall the findings suggested that there was a difference between VERT™ usage and its perceived value and that new users would value the assistance and support of more experienced users. While some users identified barriers with resource issues related to access to data and development time, there was enthusiasm for ongoing collaboration in both resource development and research. Themes which the authors concluded were vital for the successful implementation of evidence based VERT ™ resources.

Jimenez, Hansen, Juneja and Thwaites (2017)

This educational note outlined an Australian university's 18 month experience with VERT™ in a medical physics Masters programme. VERT™ was employed to supplement classroom teaching to enhance student knowledge and skills specific to medical physics equipment used for linear accelerator calibration and linear accelerator operation. By introducing students to a virtual clinical environment prior to real world radiotherapy experience would alleviate some of the issues relating to access to equipment in the radiotherapy department. The authors report on the development of three VERT™ practical sessions, each of three hours duration. The first was designed to provide an introduction to radiotherapy planning and the relationships between anatomy and dosimetry for a simple thorax treatment delivery plan. The second part of this session supported the visualisation and evaluation of individual student plans using VERT™ and interactive group discussion relating to accuracy and the advantages and disadvantages of different delivery plans. The second workshop introduced students to the association between anatomy and medical physics theory by using a whole body male CT dataset with the outlines of 43 anatomical structures visualised in VERT™ and supported by Microsoft PowerPoint® presentations to systematically discuss for example, anatomical regions, organ specific cancers and organ relationships with other organs. The final workshop focussed on aspects of linear accelerator quality assurance measurements, supported by the VERT™ physics features.

Student feedback was collected via pre and post session questionnaires for the first two sessions and a post session questionnaire for the final one. These employed Likert type rating scales, with 1 representing poor and 5 indicating excellent, combined with open-ended questions. The mean self-score rating for pre session knowledge of radiotherapy planning systems for 13 respondents was 2.4 (SD = 0.96, range 1-4) increasing to 3.1 (SD = 0.64, range 2-4) post session, no further inferential analysis was reported. For the CT anatomy session, the mean score of eight respondents, for understanding the connection between anatomy knowledge and medical physics, increased from 3.0 to 4.0 ($p = .02$, Wilcoxon sign rank test). While the post questionnaire ($n = 8$) for the linear accelerator quality assurance measurements had a rating of 3.8 to 4.4 for a range of statements. The authors concluded that the practical sessions enabled clinical education prior to entering the radiotherapy department which provided a more flexible way of teaching and they argue that it may create a deeper level of understanding. They also reported that the ability to see, hear and interact with simulated patients and equipment facilitated spatial understanding although this does not appear to have been evaluated. Overall they indicated that VERT™ has the potential to replace up to three in-hospital sessions, thereby freeing up clinical resources and time.

Chamunyonga et al., (2018)

In an educational review, the authors outline VERT features and their potential benefits to support the teaching of IMRT, VMAT, treatment plan evaluation and QA in undergraduate radiotherapy education. In doing so they begin by identifying that, in spite of the extensive use of VERT as a tool to facilitate learning in radiation therapy education institutions, the evidence for the use of VERT to support dose visualisation and virtual delivery of IMRT and VMAT plans is limited. When students transition from university settings to clinical environments, they are expected to have an understanding of the application and adaptation of these techniques. But understanding the theoretical planning concepts at undergraduate level can be challenging so the article focuses on how when Australian HEI has developed and constructively aligned VERT activities to the learning outcomes for these techniques. While they report that their experience has demonstrated the potential of VERT to enhance teaching in this area, it would have been helpful to see examples of the evaluation of this potential from a learner perspective.

Leong, Herst, and Kane (2018)

In setting the scene for this study, the authors identified that the scope of practice for radiation therapists (radiotherapy radiographers) in New Zealand includes the generation of treatment plans (dosimetry) in addition to pre-treatment localisation and treatment delivery. To support this requirement first year students learn the fundamental concepts of dosimetry and treatment planning, with the application of conceptual knowledge to the generation of basic plans being developed in the second year, while the third year focuses on more complex plan generation and critical evaluation. Whilst this scope is not consistent globally, it is internationally recognised that an understanding of plan dosimetry, and the ability to apply this understanding clinically, is essential for all radiotherapy practitioners.

Following the development of a VERT™ based teaching module that compared the technical and dosimetric features of conventional 3-D CRT and IMRT interactively, a mixed method crossover design study aimed to evaluate whether or not the VERT™ module enhanced students' perceived understanding of treatment planning concepts. The standard teaching module demonstrated a 3-D CRT plan for the treatment of prostate cancer using a proprietary treatment planning system. The clinical, technical and dosimetric aspects of the plan were discussed with the students. Isodose levels were demonstrated primarily on transverse CT slices in relation to beam arrangement, target volume coverage and dose to organs at risk. 'Beams eye views' were used to illustrate concepts of conformity to the Planning Target Volume (PTV). Following this, examples of both Intensity Modulated Radiation Therapy (IMRT) and Volumetric Arc Therapy (VMAT) plans for the same patient were demonstrated and the same aspects discussed. The VERT™ based teaching module utilised a CT dataset containing treatment fields and isodose volumes of three separate 3- DCRT, IMRT and VMAT plans for a prostate cancer patient with similar characteristics to those used in the standard teaching module. Anatomical volumes were also shown in conjunction with cross-sectional CT anatomy to reinforce the link between 3D anatomy and their representation on 2D imaging planes.

The cross over design ensured that the learning opportunities for both student cohorts (A and B) were not compromised. During teaching period one, cohort A completed the standard teaching module and cohort B completed the VERT™ teaching module. The cohorts were then crossed-over to complete the alternative teaching module 3 days later during teaching period two. The content of the two modules did not change regardless of the order that they were delivered in. Student evaluation was primarily limited to self-reporting of understanding and confidence using Likert-scale questionnaires administered at three different time points: 3 days prior to completing the teaching modules (baseline: Q-BL); following the first teaching module (Q-PM1); and following the second teaching module (Q-PM2). From a total first year group of 29 students, 20 students gave consent and took part in Q-BL (69% response rate). Within cohort A, seven students completed QPM1 and Q-PM2 following the standard and VERT™ teaching module respectively. Within cohort B, eight students completed Q-PM1 following the VERT™ teaching module, whereas seven students completed QPM2 following the standard teaching module. Qualitative data was also collected from the two staff members delivering the module via single, semi-structured interview to explore their experience using the VERT™ teaching module and their perceptions of its effect on student understanding. Interviews were conducted by an interviewer experienced in health practitioner education but independent of the Department of Radiation Therapy.

Results from teaching period one showed both modules improved students' perceived understanding of radiotherapy planning concepts to a similar extent. Improvements in understanding were reported more frequently in IMRT, VMAT and treatment technique comparison, relative to 3DCRT. Student's confidence in dose volume assessment improved more frequently on completion of the VERT™ module (38%) compared to the standard module within teaching period one (38% and 14%) respectively), however, this difference was not found to be statistically significant. Within the same teaching period, the standard module improved students' perceived confidence at assessing planning CT scans more frequently than the VERT™ module (100% and 38%, respectively, $p = .026$, Fisher's exact test). When asked about their preferred teaching module 86% of students expressed a preference for a combination of both, with 36% preferring an equal combination, and 50% preferring a combination weighted more heavily towards VERT™ content. Students also commented on how the

VERT™ teaching module allowed them to visualise technical aspects of treatment techniques in a simulated clinical context. The standard module was valued for the core information it provided, with VERT™ offering a more practical and clinically applied perspective on the content.

Both lecturers valued the ability of VERT™ to visualise the conceptual content of the module within a simulated clinical environment highly. While the content itself was not novel, they felt that VERT™ allowed them to connect different technical levels of planning information (such as contoured structures and a planning CT scan) with the reality of the treatment room. In addition, VERT™ could demonstrate the motion of linear accelerator components for the different treatment techniques. They also indicated that the students appeared to be more actively engaged in discussions and questioning during the VERT™ module sessions. This increased interaction resulted in students themselves extending the scope of the lesson to cover additional material not originally planned by staff.

Bridge, Kirby and Callender (2019)

As radiotherapy planning practical experience is an integral aspect of pre-registration training, the knowledge and skills necessary to produce a clinically acceptable plan are vital preparation for both clinical treatment planning and delivery. This is especially important for complex, dynamic and adaptive techniques. The authors state that treatment planning offers a useful format for integrating student understanding of anatomy, radiotherapy technique and radiobiology. The study involved 24 students who were enrolled on a pre-registration Post-Graduate Diploma course in radiotherapy to determine the potential role of VERT™ in a radiotherapy plan evaluation workshop through a comparison with conventional tools available in a commercial radiotherapy treatment planning system (TPS). All students attended a 3-hour workshop which gave them plan evaluation experience with three lung plans for the same patient dataset. The plans comprised a conventional 3-D conformal plan, a static gantry intensity-modulated radiotherapy plan and a VMAT plan for comparison.

Participants were split into small groups with an experienced tutor available to give individual and group guidance. All students had previously undertaken at least 20 hours of tutor-guided practical planning with Eclipse but had little experience of using

VERT™ other than in treatment set-up simulation. They were asked to use both the Eclipse™ TPS (Varian Medical Systems, Palo Alto, CA, USA) and VERT™ to help with their plan evaluations and comparison. Each evaluation session took up to one hour, with the order of evaluation tools randomised. However no detail is provided relating to the randomisation method and process or the demographic composition of the group.

Following the session, all students were invited to provide feedback on their experience via an anonymous online Survey Monkey™ questionnaire. A total of 14 participants completed the questionnaire (58% response rate). The majority of students (13 out of 14, 92.9%) enjoyed the plan evaluation session and expressed a desire to use VERT™ as an additional plan evaluation tool in the future. Most students (11 out of 14, 78.6%) found the session to be useful. Participants rated the extent to which the two modalities helped them to evaluate their plans using a 0–9 scale to gather data relating to how helpful each modality was for understanding and evaluating tumour target volume dose objectives, organs at risk dose constraints, ease of beam arrangement set up and ease of plan delivery. Following testing for normality, a paired t-test across all 14 students (70 datasets) showed that, for all five dose constraints, there was a mean increase of 3.1 points in favour of Eclipse™ in terms of helpfulness compared to VERT™ (8.3 and 5.2 respectively, $p < 0{\cdot}001$). In addition, a paired t-test across the same 14 students demonstrated a mean difference of 2.4 across 28 datasets comparing each modality, in favour of VERT™ for ease of set-up and delivery (7.5 and 5.1 respectively, $p < .001$). To identify any differences between group perceptions of the usefulness of VERT™ based on the order of evaluation, an independent t-test was performed between the groups that accessed VERT™ or Eclipse™ first. This showed a statistically significant difference in student scores of usefulness for evaluation of constraints in favour of VERT™ for the group that used Eclipse™ first (mean score =6) compared to those who used VERT™ first (mean score = 4.3, $p = .001$). Interestingly, the students that used Eclipse™ first as an evaluation tool all stated that their preferred method for evaluating each dose constraint or objective was to use both systems together. In addition the value of each modality for assessing individual dose constraints and the value placed on Eclipse™ were similar for both groups. While there was a higher reported value of VERT™ from the group using it second, it was clear that, despite the relative inexperience of the students with

VERT™, they had all managed to access the necessary functions. This led the authors to suggest that the VERT™ software functionality is intuitive and training requirements were minimal.

Student free text comments for the open ended questions were collated into themes relating to which tools within VERT™ they found the most useful. The key themes identified that being able to visualise the 3-D dose and volume relationships in VERT™ and to see the actual machine deliver the plan helped students to understand clinical delivery issues related to the choice of beam angles, dose homogeneity and volumes of over or under-dosage. The authors reported that it was interesting to note that VERT™ was perceived to be useful when evaluating dosimetric factors such as target volume and OAR doses which is primarily the remit of a TPS. They concluded that while there was a clear acknowledgement that while VERT™ provided a useful overview of the plan and potential delivery issues, a TPS was essential for formal plan evaluation. They also proposed that it would be instructive to repeat this exercise with a larger sample of students as an interim plan evaluation tool and to measure what changes, if any, are made to plans as a result of visualisation in VERT™. Inclusion of quantitative analysis of performance would also provide useful insight into the specific impact of VERT™ on plan evaluation. It would be interesting to repeat this study with experienced planners to gain their perspective on the specific value of VERT for clinical plan evaluation

In addition to the literature and research reported above, there is growing interest in the role of VERT™ for patient education through the use of a dedicated Patient Education and Radiotherapy Learning module (PEARL). This has resulted in an emerging body of research relating to the wider aspects of education for patients, family and carers and the wider community. While this research domain is recognised as an important facet in the wider role of the VERT™ platform, an evaluation and impact of the research base was considered to be outside the remit of this programme of research.

2.8.1 VERT™ research discussion and conclusions

The research reported and discussed above relates, in the main, to building the confidence of learners in radiotherapy and medical physics. In addition, the development of their understanding and visualisation of the fundamental principles of radiotherapy can support and guide clinical decision making. While it has been adopted worldwide, much of the research has focussed on the development of learner confidence and psychomotor skills as an alternative to developing these in the clinic, where opportunities are often time pressured. It is suggested that while this research has had a positive impact, the evidence base to support the importance of developing 3-D spatial visualisation skills in radiotherapy learners and the role that VERT™ may play in this development remains under researched. Therefore, the key questions of how novices in Radiotherapy develop their 3-D spatial visualisation skills and whether VERT™ has an impact on this development remain unanswered. This is similar to the conclusions drawn by Reedy (2015, p.355) in relation to the role of simulation in health professions education in general where there is a lack of theoretical grounding for the design and implementation of learning and teaching strategies.

2.9 Chapter summary

By adopting a narrative review approach, the chapter began with a discussion of the evolution of radiotherapy practice. This focused on the transition from predominantly 2-D techniques employed in the early and mid-20th century to the development of current 3-D and 4-D treatment planning and delivery methods in the 21st century. It continued with an exploration of the radiotherapy workflow and sub processes of tumour target volume delineation, dosimetry, verification and treatment delivery. As part of this exploration, the importance of accuracy, safety, the impact of automation and the importance of well-developed 3-D spatial visualisation skills was emphasised. The chapter concluded with a review of the research that has been conducted with VERT™ from the platform's initial development to the present.

Chapter 3

Review and critical evaluation of the spatial visualisation literature

3.1 Introduction to chapter 3

This chapter will provide the literature review for spatial visualisation skill and its measurement and will be presented in four parts. It will begin by exploring the evolution of theories pertaining to the components of spatial visualisation skill and how they may be developed and how spatial and visual information is processed. In doing so, links to the external beam radiotherapy pathway, workflows and processes will be established. The chapter will then provide a description of common spatial visualisation test instruments that have been reported in the literature. It will continue with a critical evaluation of the spatial visualisation measurement literature. From the findings of this critical evaluation, the chapter will conclude with a justification for the selection of appropriate test instruments for the measurement of the 3-D spatial visualisation skill of radiotherapy learners conducted during this programme of research.

3.2 Location and selection of spatial visualisation literature

To understand the nature of visual information processing, to define spatial visualisation and its development and to determine how it may be measured, four themes were identified:

1. The definition of spatial visualisation skill;
2. The processing of spatial and visual information;
3. The development of spatial visualisation skill;
4. The measurement of spatial visualisation skill.

The first three themes were used for the development of keyword search terms for the location of published material relating to spatial visualisation in general and included meta-analyses, literature reviews, and primary research studies. Searching of specific electronic databases available via the EBSCO host search engine including CINAHL, EBSCO, ERIC, IEE Xplore, INFORMIT, JSTOR, MEDLINE, PsychINFO, Science Direct, Web of Knowledge and Web of Science was employed.

For the fourth theme, focused on the measurement of spatial visualisation skill, key words, thesaurus terms, inclusion filters and Boolean operators which were applied to the literature relating to primary research and studies, these can be found in appendix 2.

3.3 Defining spatial visualisation skill

Research relating to spatial visualisation as a component of human intelligence has been conducted since the early part of the twentieth century yet there is still debate about the nature of spatial visualisation and its components. While there is broad agreement about general intelligence, usually referred to as g, as proposed by Carroll (1997, p. 33) there has been less agreement about individual sub factors. Central to the debate was the Cattell Horn theory which identified fluid intelligence (g: F) as a problem solving ability which is not influenced by previous experience and crystallised intelligence (g: C) which relies on the application of consolidated knowledge (McGrew, 2009, p. 5). In an attempt to link spatial ability with human intelligence, Johnson and Bouchard (2005, p.397) suggested that spatial perception aligned with g: F and verbal and general visualisation skills align with g: C.

However an earlier review by Lohman (1979, p. 2) had already identified that confusion existed in the field of spatial ability research during the latter part of the twentieth century. This was related to identical tests being identified with different names in different studies and uncertainty about what these tests were measuring due to subtle changes in their format and administration. Through an analysis of previous general intelligence research, the findings of the review, which Pittalis and Christou (2010, p. 195) later referred to as influential, provided spatial ability researchers with a model of three major spatial ability factors which Lohman (1979, p. 189) identified as:

1. Spatial relations, defined as:

   "The ability to turn or rotate a given figure or part of that figure in one plane (or around an imaginary axis) to see if it corresponds to another figure in the same plane";

2. Spatial orientation, identified as an ability to imagine how an object will appear from a different perspective or:

   "The ability to comprehend the arrangement of elements within a visual pattern with reference to the human body";

3. Spatial visualisation was considered as a combination of the first two and defined as:

> *"The ability to see the relationships of objects when the subject has to imagine that the object or objects involved have changed their position in space relative to one another"*.

The conclusions drawn by Lohman were supported by a later meta-analysis of the spatial visualisation literature by Linn and Peterson (1985, p. 1482) who proposed the following nomenclature:

1. Mental Rotation –an ability to rotate 2-D or 3-D figures rapidly and accurately (Lohman's spatial relations);

2. Spatial Perception – the ability to determine the relationships of objects with respect to an individual's own body orientation (Lohman's spatial orientation);

3. Spatial Visualisation which summarises all those tasks involved in the multi-step process of manipulating symbolic (non-linguistic) information.

While the interpretation of and definitions for spatial visualisation provided by Lohman and Linn and Peterson are similar, the literature continued to report an inconsistency in the application of the definitions. For example, D'Oliveira (2004, p. 20) and Mohler (2008, p. 2) have identified conflicting perspectives for the names and numbers of general intelligence factors and Yilmaz (2009, p. 84) indicated that, while spatial ability is an important component of general intelligence, the application of definitions is inconsistent. For example, in relation to spatial visualisation and spatial skill, Goldstein, Haldane and Mitchell (1990, p. 546) refer to visual-spatial ability and visual spatial skill interchangeably. However, Sorby (1999, p. 21) makes the differentiation between ability and skill by suggesting that visualisation ability is something that an individual is born with and visualisation skill can be developed through training. An alternative view has been proposed by Sutton and Williams (2007, p. 3) who made the distinction between the mental rotation of objects and the relationship between objects in space, a view supported by Peters and Battista (2008, p. 260) who suggested that:

> *"Mental rotation should be considered as a separate entity to perception and visualisation"*

There remains, however, a lack of consistency in the application of the terms ability and skill which are frequently used interchangeably, for example, Velez, Silver and Tremaine (2005, p. 512) defined spatial ability as:

> *"Those skills involving the retrieval, retention and transformation of visual information"*

However, Liu, Tendick, Cleary and Kaufman (2003, p. 609), suggested that ability is:

> *"A relatively stable capability that supports performance in a task while a skill is something that can be learned through training"*

While Terlecki and Newcombe (2005, p. 433) stated that it is:

> *"A skill in representing and transforming symbolic information in space"*

The lack of a clear taxonomy may present a challenge for any programme of research which seeks to investigate the development of spatial visualisation skill. Therefore, as one of the aims of this programme of research was to determine the spatial visualisation skill of learners in radiotherapy, a combination of definitions provided by Lohman and Linn and Peterson was employed. Using these definitions and applying their component factors to the radiotherapy process outlined in chapter 2.4, p. 35, the application of spatial visualisation theory in radiotherapy workflow processes is summarised in table 3.1.

Table 3.1: The relationships between spatial visualisation skill domains and their application in radiotherapy

Spatial Domain	Radiotherapy Application
Spatial visualisation	Building a mental model of the radiation beam path as it travels to the tumour target volume via relational organs (visualising a beams eye view)
Orientation / Perception	The relationship between the tumour target volume and the anatomical organ it is located in and the relationships between the organ and those in close proximity to it
Spatial Relations / Mental rotation	Changes in position of the target volume due to patient rotation in the X (horizontal), Y (longitudinal), & Z (vertical) planes, the impact of this rotation on the position of the tumour target volume and related organs and manipulation of the patient's position to correct for this rotation

3.4 Influencing factors for the development of spatial visualisation skill

In the same way that the components of spatial visualisation skill have been the subject of debate in the literature, so have the factors which may influence its development. The literature identifies two areas of influence; those of biological and environmental factors, thus prompting a nature versus nurture dialogue. For example, Rust and Golombok (2009, p. 12) suggested that 50% of the variation in intelligence test performance is related to biologically inherited characteristics. While a study by Casey, (1996, p. 246), employed Vandenberg & Kuse 20 item MRT to determine influences for individual differences in a cohort of 433 North American College students. Using a scoring convention of one point for each correct answer choice, the study found that right handed females, with one or more non-right handed relatives and majoring in maths or science outperformed males in the same category with mean scores of 66.7% compared with 64% for males and 47.5% for female non-maths or science majors.

From the findings of a meta-analysis of 268 studies, Voyer, Voyer and Bryden (1995, p. 260) identified that solution strategies for the subcomponents of spatial visualisation vary by spatial task and that gender differences are more significant in some tasks, for example, mental rotation. They also proposed that brain lateralisation (left brain-right brain interplay) and differential spatial experience as explanations for performance difference. In a later meta-analysis to determine if mental rotation tasks are mediated by motor stimuli in the brain, Zacks, (2008, p. 2), identified that, as each rotation is made, activation was detected with magnetic resonance imaging and positron emission tomography in the superior parietal cortex of the cerebral hemispheres, the pre-central and posterior frontal cortex motor areas. This led Zacks to conclude that that brain organisation particularly in the cerebral hemispheres was a factor in mental imagery and object transformation.

However, based on a review of the literature, Plomin and Petrill (1997, p. 60) identified that genetically related children, growing up in the same family, demonstrate different cognitive development and intelligence which suggested an environmental as well as a genetic influence. This led them to propose that a link existed between genetic disposition and choice of environment. A before and after study involving 110 psychology undergraduates conducted by De Lisi and Cammarano, (1996, pp. 356-7) reported that playing a 3-D block rotation computer game contributed to both male and females improving their mental rotation performance scores over baseline compared with those who played 2-D computer games. While Terlecki and Newcombe, (2005, p. 436), in another study of 1278 psychology students, found that those students with high spatial experience achieved higher mean scores on the Vandenberg and Kuse 20 item MRT, achieving 57.5% compared to those with lower spatial experience who scored 42.5%. But Newcombe and Stieff, (2012, p. 960), add a note of caution in that cultural perception and influence, particularly in developed countries, where spatial tasks are viewed as being predominantly masculine females will perform less well than their male counterparts because they believe that they will have less well developed spatial skills.

Given the above, there are still questions that remain unanswered. The first relates to the observed male performance advantage for mental rotation. What impact, if any might this have for predominantly female education programmes and professions such as radiotherapy? In relation to spatial experiences and the impact of training, if an individual has low spatial experience at the beginning of a programme of study and if visualisation training is effective, then how can those individuals who may benefit from training interventions be identified. Finally, how does the interplay between biological characteristics and choice of environment impact on spatial visualisation performance?

Linking the themes of nature versus nurture to radiography education, the published evidence relating to the development and measurement of spatial visualisation skills of both diagnostic imaging and radiotherapy radiographers is limited. Additionally, there appears to have been little emphasis placed on the development of specific learning outcomes for 3-D spatial visualisation in pre-registration radiography education. Without this focus, learners may be left to develop these skills from ad-hoc opportunities whilst viewing 3-D treatment plans and cross sectional X-ray images. Given the complexity of radiotherapy treatment planning and delivery workflows and processes, it can no longer be assumed that clinical observation on its own will automatically stimulate the development and improvement of skills in 3-D spatial visualisation. Evidence from other technology and engineering disciplines has also indicated that baseline spatial visualisation skill may be a predictor for the mastery of complex contextual skills (Hegarty, Keehner, Khooshabeh & Montello, 2009, p. 68). It has also been reported by Berney, Bétrancourt, Molinari and Hoyek, (2015, p. 453) that a relationship exists between spatial visualisation and the ability to learn structural anatomy. Therefore one of the questions which underpinned the development of the conceptual framework for this programme of research related to the possibility that spatial visualisation skill may change over time as a result of focussed training and specifically designed instructional methods as reported by Wang, Chang and Li (2007, p. 1194).

These differing views have led to a discussion relating to whether spatial visualisation skill is innate and therefore unchangeable over time or whether it is malleable and can be developed with focussed training. If spatial visualisation skill can be developed as proposed by Wang, Chang and Li (2007, p. 1194) then it may be enhanced through specifically designed instructional methods. It has also been suggested by Uttal and Cohen (2012, pp. 175-177) that the development of spatial skills may respond differently to different kinds of training including video game playing, semester-long instructional courses and training participants on spatial tasks through targeted practice, instruction or computerized lessons. They also identify that enhanced education could pay substantial dividends in professions where a high level of spatial visualisation skill is required. If this is the case, then the employment of visualisation platforms such as VERT™ should have an impact on the development of 3-D spatial visualisation skill and performance of tasks with complex visualisation components.

3.5 Knowledge building, visual information processing and memory systems

The way in which a learner organises new information and knowledge in relation to their prior knowledge and experience is fundamental to successful learning (Rutherford–Hemming, 2012, p. 130). The following sections will explore the theories relating to knowledge building, visual information processing and memory systems and how these support the novice to expert transition in radiotherapy. It will continue with a discussion of the twin theories of cognitive load and cognitive fit which may have an impact on the effectiveness of learning. It will conclude with an examination of how these factors may have on a learners approach to learning.

3.5.1 Conceptions of knowledge and schema building

In an attempt to provide a systematic approach and description of knowledge, two alternative viewpoints have been proposed by de Jong & Ferguson-Hessler, (1996, p.106). The first is epistemological and implies that knowledge is characterised by the role it plays in task performance. The second is cognitive and suggests that knowledge recognises the demands of a task. In relation to task performance, Mann (2002, p. 70) referred to a stepwise progression for key milestones of knowledge development. This development begins with the gathering of facts that are elements of general concepts and principles. The learner will see these facts as non-reducible into smaller elements and it is these elements which form the basis of a learner's factual and conceptual knowledge. Otherwise referred to as declarative knowledge, it is the knowledge that an individual knows and can recall and report (Baartman & De Bruijn, 2011, p.127). The organisation of this declarative knowledge into a cohesive framework for application and integration into practice is known as procedural knowledge and covers the skills and strategies that support the application of knowledge in practice. This progression requires the combination and connection of a number of elements into increasingly complex chunks of information, referred to as schema. At their lowest level these schema may relate to a single part of a more complex procedure. For example, when considering the advanced radiotherapy techniques, a low level schema could relate to the correct alignment of a patient in their immobilisation devices on the linear accelerator treatment couch. A higher level schema may relate to the correct positioning of the tumour target volume at the linear accelerator isocentre. This would be followed by additional schema covering the mental visualisation of a beam's eye view prior to the employment of daily pre-treatment imaging, image approval and treatment delivery.

This progression of knowledge building can be linked to work by Miller (1990, p. s63) who proposed a hierarchical, pyramid, structure for knowledge development and the acquisition, demonstration and application of clinical skills. Declarative knowledge and its translation into procedural knowledge to support application and integration into practice was referred to by Miller (ibid) as "the knows and knows how" and "shows how and does" respectively. The progression of a learner's knowledge development and application to the apex of Miller's pyramid reflects the students' awareness of discrepancies or variations from normal or expected situations. Referred to as conditional or meta-cognitive knowledge, it serves as a cognitive control mechanism during problem solving and is synonymous with tacit knowledge. This progression of knowledge development and the relationships with practice application and skill mastery are summarised below in figure 3.1. Reinders (2010, p. 31) has suggested that "knows what" knowledge is founded in scientific empiricism, but the practice of caring also concerns the holistic management of the patient and the dimension of personal tacit knowledge comes into play. In relation to scientific discovery, intuition also forms part of that process. By way of illustration, Polanyi and Grene (1969, p. 123) described a scenario in which a psychiatrist and his students observed a patient having what appeared to be a mild seizure. Following a discussion as to the type of seizure, the psychiatrist concluded that:

> *"….you have seen a true epileptic seizure. I cannot tell you how to recognise it; you will learn this by more extensive experience".*

This experience was referred to by Polanyi and Grene (ibid) as tacit knowledge and in his conception; he talked about characteristics that the speaker cannot explain or identify where the feeling or supposition comes from. Reinders (2010, p. 32) also proposed that tacit knowledge is part technical and partly composed of beliefs, perceptions and mental models. While scientific knowledge is contextual and embedded in specific practices, tacit knowledge is embodied in the individual professional – patient dimension of knowing how to interpret patient behaviour. This is often based on previous experience or observations and while it can be demonstrated it may not be easily explained. Henry (2010, p. 292) has also suggested that Polanyi's concept of tacit knowing provided a starting point for constructing the epistemology of knowledge in clinical practice in that tacit knowing refers to knowledge that functions at the periphery of attention and makes the recognized explicit domains of human knowledge possible. However, as Puusa and Eerikäinen (2010 p. 308) have identified, all knowledge is either tacit or based on tacit knowledge and that tacit and explicit knowledge are not counterpoints to each other, they are two sides of the same thing. So the process of schema building requires the instructor to apply the formal theory supporting each task to a description and explanation of a process (Nisbet & Matthews, 2011, p. 73). However, as Horii (2007, p. 370) has indicated, experts tend to use intuition and may not include the vital early steps of organising this theoretical underpinning. This may result in the novice, with little clinical experience, being at risk of not seeing or making the links between the theory gained in the classroom and the procedure as observed in the clinic.

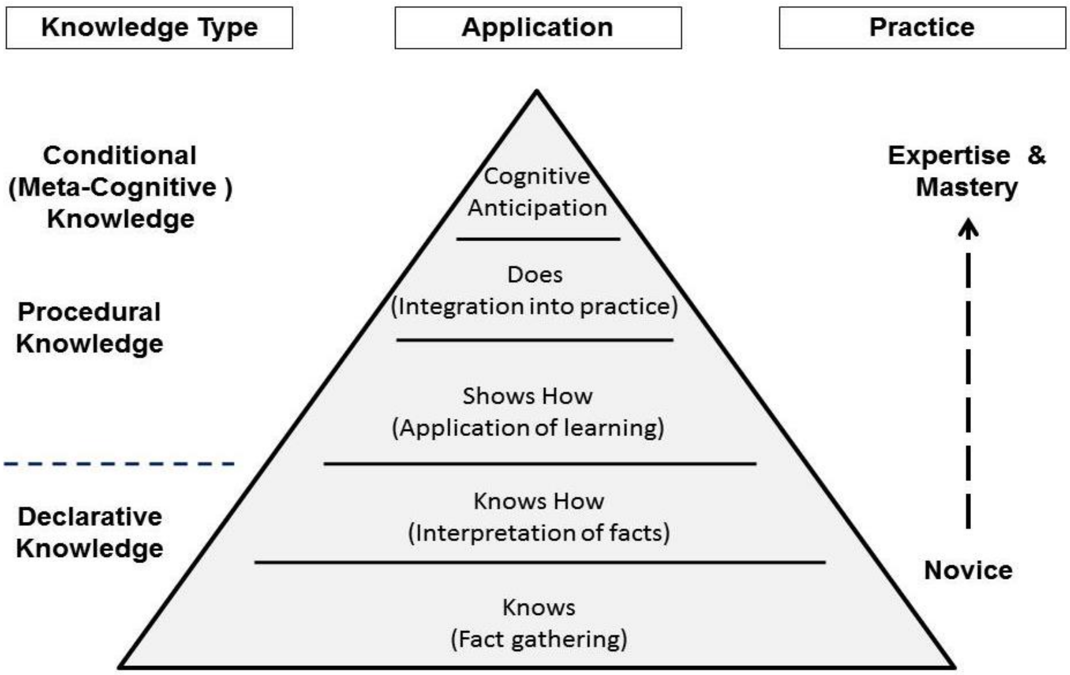


Figure 3.1: The relationships between knowledge domains, clinical application and expert practice

3.5.2 Working memory capacity, cognitive load and cognitive fit theory

The two theories of cognitive load and cognitive fit can be applied to the way in which knowledge is organised and stored and how memory functions. Central to both theories is the accepted limitation of working-memory capacity. Proposed by Baddeley (1998, p. 235) and Baddeley and Hitch (2000, p. 129), the theory of working memory states that it consists of visual-spatial and auditory channels which receive and process information from sensory memory. The limited capacity of working memory means that it can only hold in the region of seven ± two items for 30 seconds, but only process one item at a time which, when processed, will be transferred to long term memory. These items consist of auditory and visual cues which can be linked to an individual's declarative and procedural knowledge and support task performance. For example, the mental transformation of objects takes place in the visual-spatial sketchpad of working memory through the processing of visual cues, while the processing of auditory cues such as verbal instructions are processed in the phonological loop. Working memory therefore, can be considered as a buffer system that can hold temporary information about an object while it is being mentally manipulated (refer to figure 3.2).

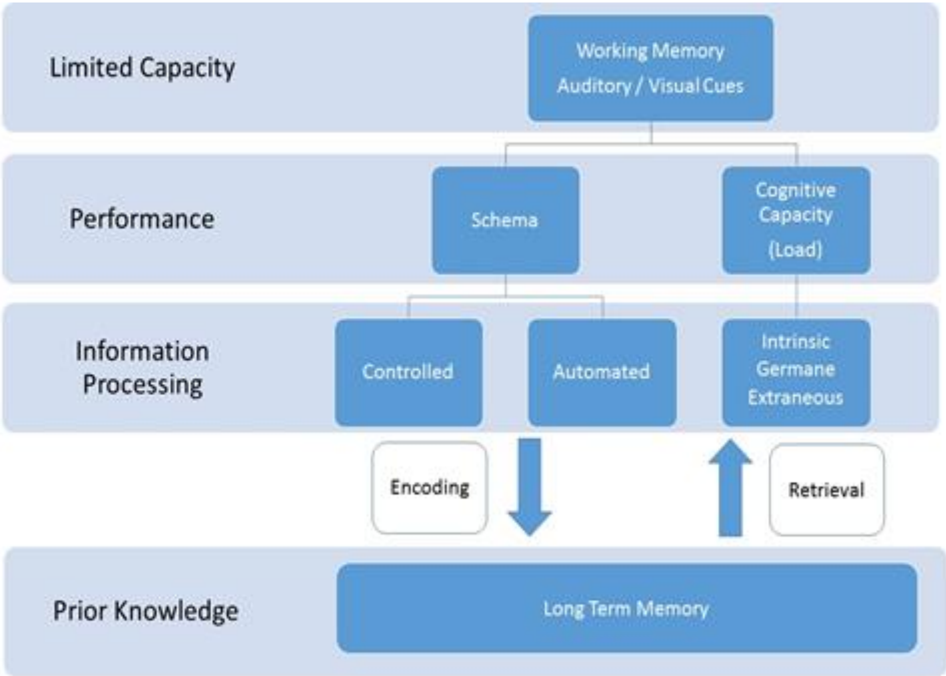


Figure 3.2: The interplay between working memory, task performance, information processing and long-term memory

Because complex schema can be treated by working memory as single entities, the limitations of working memory disappear for more knowledgeable learners when they are dealing with previously learned information stored in and recalled from long-term memory. As a result, once information is stored in long-term memory, working memory can handle complex material that exceeds its capacity prior to the information being stored. Expert performance therefore develops through the building of increasing numbers of complex single schemas from the combination of many lower level schemas. If the learning process has occurred over a long period of time, a high level schema may incorporate a large amount of information. This can be processed unconsciously in what Paas and Sweller (2012, p. 29) refer to as schema automation which reduces the load on working memory. However, for some learners the processing of information from multiple cues imposes a load on an individual's cognitive capacity and can lead to the potential for working memory capacity overload.

Cognitive load theory can provide a way of understanding the impact that learning environments can have on the ways in which people learn and the potential for learners to become overloaded with information. The theory argues for a model of cognition based on information processing. New information is first dealt with in working memory, which is optimized for dealing with new information and recalling existing knowledge from long-term memory. The theory also proposes that immersing learners in a learning environment that completely replicates the realistic world of clinical practice can make learning more difficult. A difficult task which is presented in an unstructured way can result in cognitive overload for a learner. Extraneous load refers to the ways in which the task is presented or designed and arises because of the increased cognitive load required by the multiple inputs of the environment. Extraneous load can be minimized by careful instructional design. The inherent difficulty of a task is known as intrinsic load, with some of the load being appropriate to the task at hand and thus referred to as germane load. Learners can become overwhelmed by all the inputs into their working memory and are not able to process or make sense of what they need to learn. In addition, the way in which information relating to a particular task is presented, can also be a factor in promoting learning. Known as cognitive fit theory, it provides an insight into how 3-D virtual environments may support understanding through the provision of visual clues relating to the nature and significance of a task (Van Der Land, Schouten, Feldberg, Van Den Hooff & Huysman, 2013, p. 1055). Conversely, cognitive load theory would also suggest that the rich environment may impede understanding due to the distraction caused by multiple cues. So, for tasks with strong visual components, the two theories would appear to lead to diverging assumptions relating to the contributions of 3-D virtual environments to understanding and performance.

In addition to cognitive load, the process of learning itself will also have an impact on the novice due to the amount of concentration involved in performing the task and is referred to as germane load (Naismith, 2015, p. 806). Novice learners in radiotherapy could experience working memory capacity overload as they deal with new concepts such as mentally visualising the position of individual organs and their relationships with other organs. This can be compounded by the associated requirement to mentally visualise the intended beam path. But for advanced learners who have the capacity to recall previously developed schema the task becomes less demanding and requires less concentration. Therefore the aim of any learning and teaching intervention will be to manage intrinsic load, minimise extraneous load and optimise germane load.

The impact of cognitive load in virtual environments was investigated in a study conducted by Van Der Land et al., (2013, p. 1058) involving 192 undergraduate students (mean age 19.9, SD = 1.55, range 18 – 28) who were given the task of selecting an appropriate apartment for sharing with two others. Sixty four groups, each comprising three students, were randomly assigned to one of three experimental conditions of a 2-D static floor plan, a 3-D static room view and a 3-D immersive virtual environment utilising a tracked avatar. In the first phase, individual participants made a selection of an apartment and then worked as a group to reach a consensus. The findings indicated that the three conditions were different in their realism and interactivity. The 2-D condition was identified as the least realistic and interactive compared to both the 3-D conditions which showed no difference between them. However cognitive load was shown to be the highest in the 3-D immersive condition and lowest in the 2-D condition $F(2,189) = 12.78$, $p < .001$. This may go some way to explaining the findings from the VERT ™ studies reported by Appleyard and Coleman (2010), Green and Appleyard (2011) and Flinton (2015). However, Van Der Land et al., (2013, p.1060) also reported that individual understanding of visual representations of physical space is more effective in 3-D virtual environments.

3.5.3 Linking knowledge building to the novice to expert learner transition

Novice learners in radiotherapy will require rules which can be followed clearly and have their performance monitored to ensure that tasks are completed successfully. As they progress to the advanced beginner stage they can begin to apply these rules to similar situations, but they may still have trouble with determining the importance of events and problem solving. At the competent level, learners can plan their work and take responsibility for the outcomes, while proficiency confers the ability to see the wider context, which supports the prediction of events. At the highest level, experts are viewed as masters, because each situation will trigger an intuitive response based on previous personal reference points and evaluations. The radiotherapy clinical environment is an information rich environment in which the declarative and procedural knowledge of the principles and concepts supporting radiotherapy processes must be integrated with the psychomotor skills required for the control of the linear accelerator and associated equipment. Successful integration should result in the safe and accurate delivery of dose to the tumour target volume, avoidance of unnecessary dose to the surrounding organ and limitation of error consequence. Attainment of expert level performance in such an environment is predicated by the successful organisation of tasks, visual information and decision-making skills (Patel, Yoskowitz, Arocha & Shortliffe, 2009, p. 177).

One of the challenges for academic and clinical educators is to support the transition of the learner through this knowledge building progression. Due to the varying exposure to and experience with particular radiotherapy procedures, a learner may observe and assist in a procedure (or part of it) before covering the underpinning theoretical base. Clinical experience is not always equitable across a cohort of learners due to rostering constraints and clinical opportunity. The scheduling of periods of deliberate practice in radiotherapy department can provide a systematic approach to knowledge building supported by the systematic bottom up approach to skill acquisition of Miller's model. However there is always a risk that learners, at any level of study, may be at a different stage of skill development during any given period of clinical placement. This transition can be demonstrated by considering the example of cancers arising in the male pelvic organs. A learner gaining experience on a linear accelerator treating those patients immediately after receiving the theoretical principles supporting the treatment of these tumours will get the opportunity to become immersed in the workflows of specific treatment delivery processes. They will be able to apply their declarative knowledge and construct their procedural (clinical) knowledge in a stepwise and progressive fashion. By the end of this specific rotation they may find themselves at the threshold of the advanced beginner and competent levels, using the five stage model proposed by Dreyfus and Dreyfus and discussed by Benner (1984, pp. 22 – 25). If the learner then spends time in the pre-treatment area where they may have no previous experience on which to build, they could be seen to be at the novice stage. At the same time, those learners who have placements on a linear accelerator treating patients with breast or lung cancer (before receiving the underpinning theory to support declarative knowledge) are unlikely to find themselves at the same point on the novice to expert continuum.

3.5.4 Approaches to learning and deconstruction of task

The way in which a student approaches their learning will also have a role to play in the way that they process information and build their knowledge base. As identified by Virtanen and Lindblom-Ylänne (2010, p. 355), they may come to a learning situation (opportunity) with prior experiences, perceptions of learning, motivation and self-regulation. The quality of learning, therefore, can only be partly influenced by instructor activities which are designed to get students to apply the level of cognitive processing required to achieve the intended learning outcomes (Biggs, 2011, p. 91). He goes on to propose that the learning activity must be seen by the student as having value and that they can succeed in it (ibid, p.92). However individual learners will adopt different approaches to their learning, characterised by Marton and Säljö (1976, p. 7) as a surface or deep approach. Those who adopt a surface approach tend to see tasks as externally imposed and attempt to cope with course requirements by memorising facts. Alternatively, those who adopt a deep approach will focus on understanding and meaning and will demonstrate an intrinsic motivation to learn (Virtanen & Lindblom-Ylänne 2010, p. 357). As they also point out (ibid, p.358), there are two different approaches to teaching which may influence approaches to learning. The teacher centred approach focuses on the transmission of knowledge, where the learner is more likely to have a passive role as opposed to the student centred approach where the teachers role is to assist students in focusing on the understanding and practical application of information and knowledge. This approach is linked to a further aspect of teaching and learning; that of the parts-to -whole (PTW) and the whole-to-parts (WTP) approaches. Described in relation to engineering drawing by Akasah and Alias, (2010, p.81), the PTW approach emphasises a sequential, teacher centred, process which begins with 2-D representations of objects and moves to 3-D transformations of those same objects. The WTP approach is the reverse of PTW, as students start with 3D object representations and deconstruct them into their 2-D component parts. The authors suggest that the WTP approach provides students with more immediate understanding of the relationships between 2-D drawings and their related 3-D objects.

The ability to deconstruct and visualise radiotherapy processes and workflows can be achieved with VERT™. However, a syllabus and curriculum structure which provides a one size fits all approach to clinical skills preparation and beam geometry visualisation. If it is possible to differentiate between those learners with well and less well developed 3-D spatial visualisation skill by measuring performance at the commencement of a programme of education then the potential for the development of additional 3-D visualisation resources arises and this may assist in the improvement of these skills. To achieve this, a review of available spatial visualisation skill test instruments is necessary.

3.6 The measurement of spatial visualisation skill

As the theories relating to the components of spatial visualisation were developed from the theories of general intelligence and its sub factors, so the test instruments for the measurement of spatial visualisation have evolved in a similar way. This has resulted in a lack of agreement on how each of the domains identified by Lohman and Linn and Peterson (see section3.3, p. above) may be measured. For example, Johnson, Bouchard, Krueger, McGue and Gottesman, (2004, p. 97) have indicated that the test batteries employed for visual, verbal and numerical mental ability tasks have been developed by different groups based on their conception of the structure of these abilities. Through the middle of the 20th Century, tests for spatial tasks were based on tests for general intelligence with specific loadings for these visual, verbal and numerical factors. The tests tended to be an unreliable measure of all three, since inferences needed to be made about one factor based on the results of the others and did not adequately measure either visual or verbal components of working memory. This led Johnson and Bouchard, (2005, p. 413) to conclude that a study using standard attainment tests was likely to be less reliable for spatial factors because the tests were targeted towards numerical and verbal factors and were lacking in spatial test components. They also proposed that domain specific tests were needed for each of the spatial components. Kozhevnikov and Hegarty (2001, p. 745) also made a distinction between the mental processes involved in rotation, orientation and perception of objects and their relationships. This distinction has been supported by Peters and Battista (2008, p. 260) who suggested that mental rotation is a separate entity to perception and visualisation. Mental rotation tasks require an object to object manipulation and transformation with respect to an environmental frame of reference, while the individuals` viewing position remains static. An individual's viewing position has been referred to by Hegarty and Waller (2004, p. 176) as the egocentric reference frame, which for perceptual and visualisation tasks, needs to change with respect to the object. The difference between egocentric spatial transformations (imagining the results of changing one's egocentric frame of reference with respect to the environment) making object-based transformations (imagining the results of changing the positions of objects in the environment, while maintaining one's current orientation in the environment) led Hegarty and Waller (2004, p. 188) to suggest that there was a dissociation between tests of perception taking and mental rotation.

Test instruments for the measurement of spatial rotation, orientation and perception / relationships can be divided into those requiring 2-D or 3-D manipulation. They can also be considered in relation to the type of individual reference frame transformation required. The following sections will provide an overview of the different types of tests reported in the spatial visualisation skills measurement literature.

3.7 Examples of 2-D test instruments

3.7.1 Paper folding tests

Paper folding tests described by Eliot and Smith (1983, p. 334) and Logan (2015, p. 428) utilise objects similar to those shown in figure 3.3. Participants are required to visualise the folding action of a square sheet of paper (image 1). When that sheet of paper is then folded and has a hole punched in it (image 2), participants are required to identify, from a choice of five options, how the sheet will appear when fully reopened.

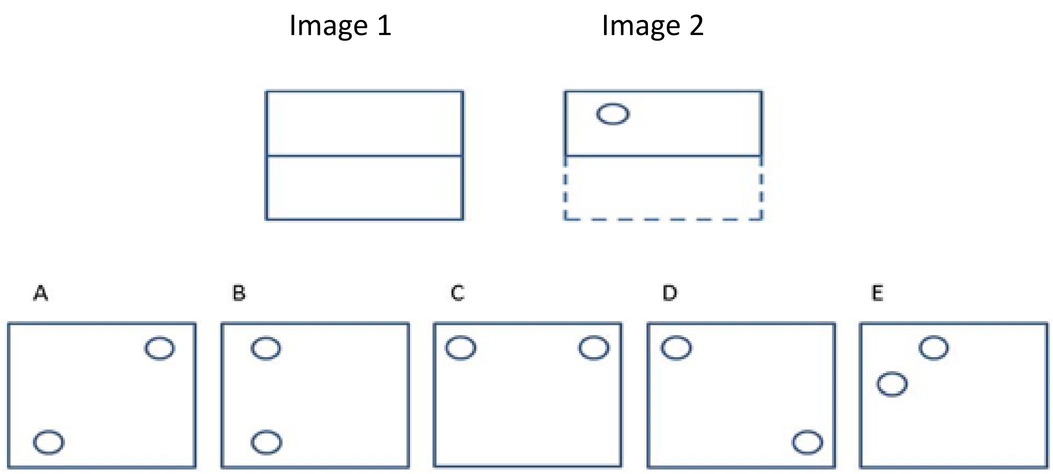


Figure 3.3: Facsimile example of a paper folding test item where option B is the correct answer

3.7.2 Pattern recognition tests

Pattern recognition tests have been described by Ekstrom, French, Harman and Dermen (1976, p. 21) and Linn and Petersen (1985, p. 1425). Also referred to as hidden figure tests, they are designed to test the ability to identify a simple line drawing of a shape within a more complex background pattern of lines or shapes and shown below in Figure 3.4.

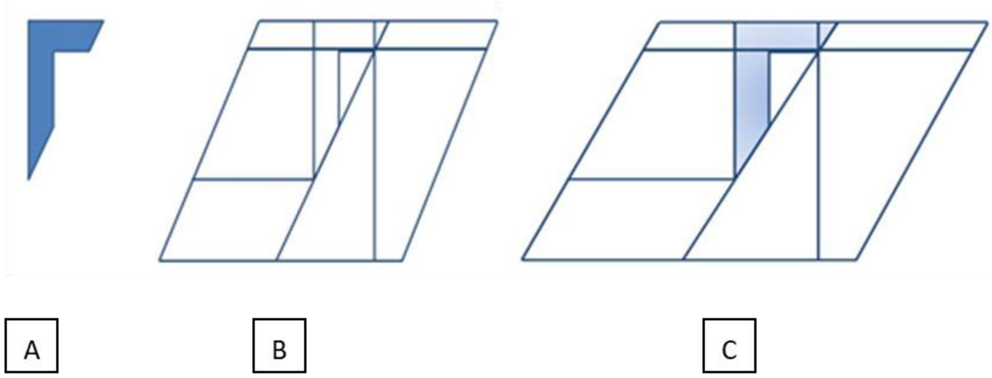


Figure 3.4: Facsimile of a hidden figures test where object A must be correctly matched to pattern B to obtain pattern C

3.7.3 Block rotation tests

These tests involve shape rotation as described by Ekstrom et al., (1976, p. 150). In the example shown in figure 3.5 below, an irregular shape (the reference image) is shown at the left hand side of a black line, with eight rotated or mirrored versions of the same shape being shown on the right. Participants must decide whether each of the eight versions are the same as the reference image or different.

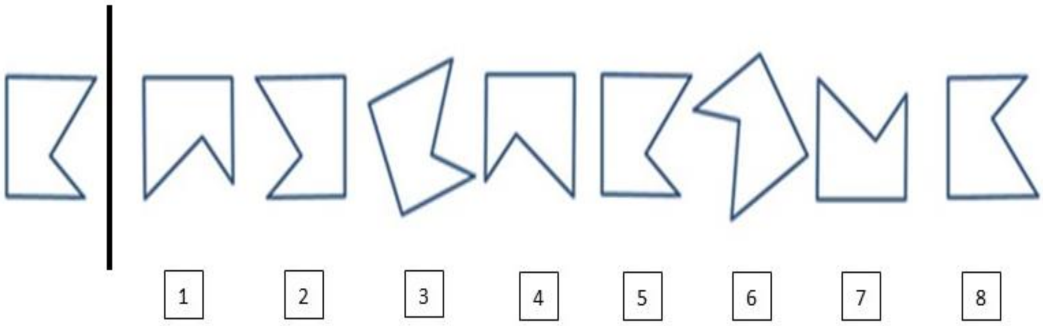


Figure 3.5: An example of a block rotation test where answer option 5 is the same shape and options 1, 2 and 8 are different

3.7.4 Flag rotation tests

The flag rotation test shown in figure 3.6 was originally developed by Thurstone and Thurstone in 1941 and described by Eliot and Smith (1983, p. 208), it is designed to test visual manipulation. From the initial position (figure 3.6 a), the flag is rotated and turned over on either its long edge, or in the case of the example shown here, its short edge. Participants must determine whether the resulting image (figure 3.6 b) is the same as the original view or whether it is different. Since its original development, a number of similar tests, both timed and untimed and of varying complexity have been developed.

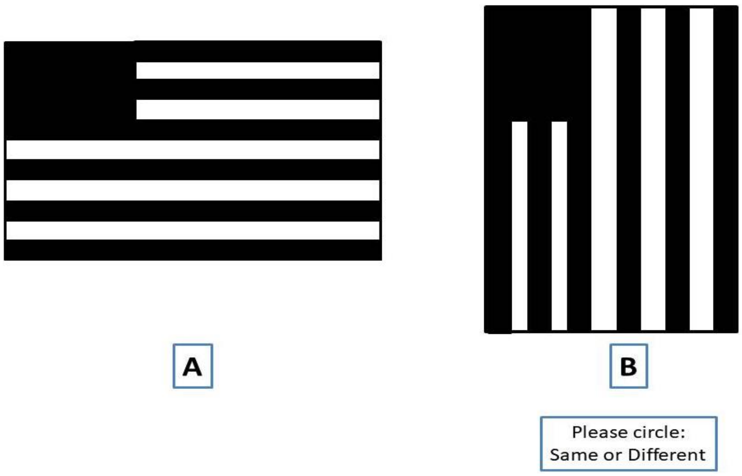


Figure 3.6: Example of a flag rotation test where the rotated image B is the same as image A

In all of the tests described above, the conventional scoring method set out in the test administration instructions, awards one mark for each correct answer identified. In addition to block and flag rotation tests, Eliot and Smith (198, p. 238) also refer to card rotation tests. Criterion test objects, referred to as cards, have a similar design to the block rotation objects described above in section 3.7.3 but have a series of holes punched in them. From the six answer choices, participants need to select three cards which can be directly overlaid on the criterion object. Alternative versions of the test use letters of the alphabet, for example E or L, which have a whole punched in one corner.

3.7.5 Form board tests

One example of this type of test is the revised Minnesota Revised Form Board Test described by Evans and Dirks (2001, p. 876). The test requires participants to view sets of geometric shapes which contain a disassembled shape and five assembled shapes and by mental manipulation, select the correct assembled shape that corresponds to the disassembled shape. The test can consist of 24 or 48 items split into two parts with a time allocation of 8 minutes per part. Scoring can be done in one of two ways, either a count of the total number of correct selections within the allotted time (Evans & Dirks 2001, p. 876) or the number of correct items minus number of incorrect items (Ekstrom et al., 1976, p. 175). While the reason for this has not been explained, some researchers apply this method in other tests to minimise the effect of guessing. In the examples shown below in figure 3.7, the correct answer for both test objects is option D.

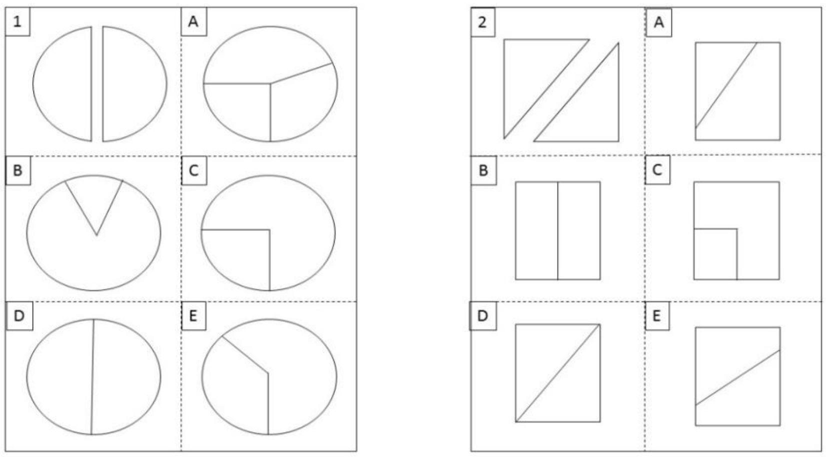


Figure 3.7: Facsimiles of two form board test items

3.7.6 Surface development test

These tests are designed to test the ability to mentally construct a 3-D object that would be formed by folding a 2-dimensional pattern. In the example shown in figure 3.8 (a),described by Ekstrom et al., (1976, p. 177), participants are required to visualise how the shape may be folded along the dotted lines to create a 3-D object and to imagine which of the numbered edges correspond to the lettered edges of the 3-D object shown in figure 3.8 b. The side marked with the letter X will be the same as the side of the object marked X. Therefore the paper must be folded so that the side marked X is on the outside.

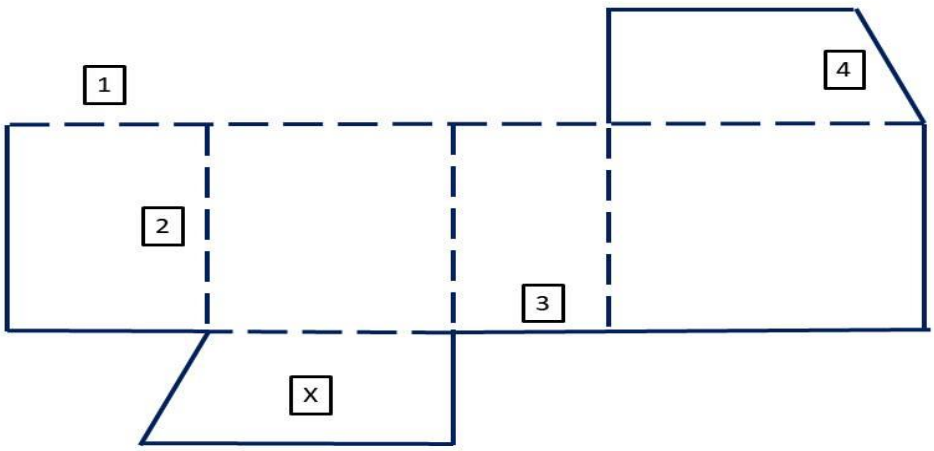


Figure 3.8 (a): Two-dimensional surface development plan

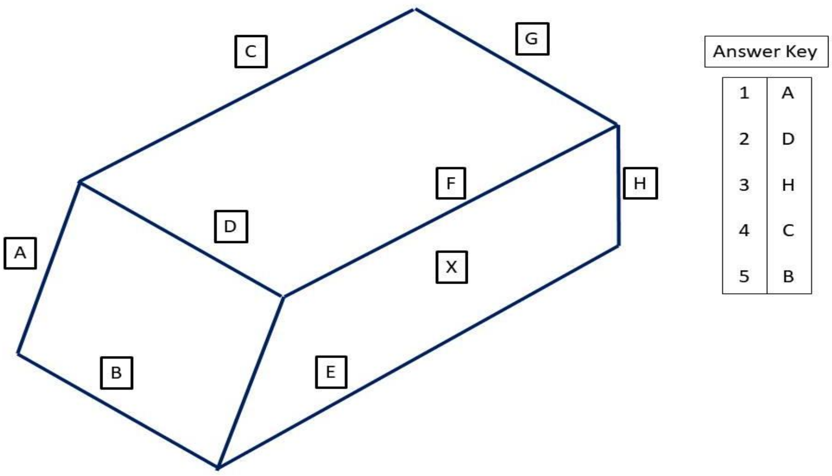


Figure 3.8 (b): Resultant three- dimensional object achieved by folding figure 3.8 (a) along the dotted lines and the answer key for matching sides

3.7.7 The Guilford-Zimmerman orientation test

The orientation test was first described by Guilford and Zimmerman (1948, p. 27) and formed one component of larger general aptitude surveys composed of a number of tests, with each being designed to measure a specific primary intellectual ability. The orientation test consists of a series of images representing two pictures of a lake and shoreline as viewed from the prow of a boat, which has moved slightly between pictures. The task is to select one of five diagrams that represent how the boat has moved. In each diagram (figure 3.9) the dot represents the old position of the prow and the dash represents the new position. Changes can include any combination of heading (i.e., rotation) and forward and sideways translation of the boat.

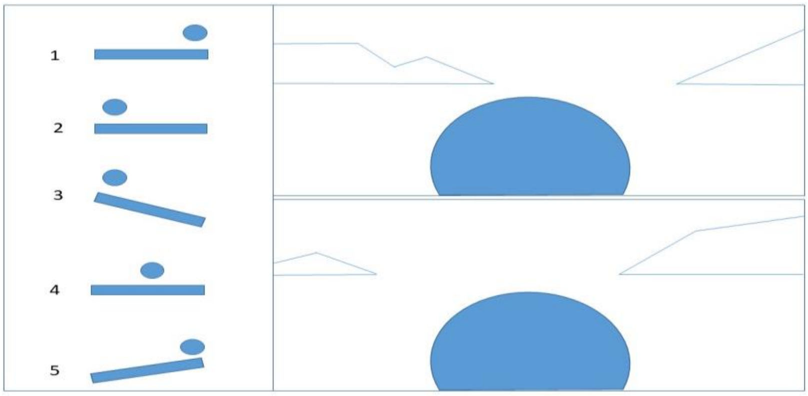


Figure 3.9: Facsimile of the Guilford-Zimmerman orientation test showing two positions of the prow of a boat. Option 1 indicates the correct change in direction

In general, the 2-D tests outlined above object to object manipulation within an environmental reference frame, whereas the more complex 3-D tests described below require egocentric reference frame transformations.

3.8 Examples of 3-D Tests

Complex 3-D tests involve the construction, comparison, cutting or rotation of cubes with a range of complexity as identified and described below.

3.8.1 Cube construction test

The block subtest of the Wechsler Adult Intelligence Scale (WAIS) test (figure 3.10) requires the construction of nine coloured blocks to recreate a design which is depicted as a 2-D plan. Participants must translate and rotate 3-D blocks to replicate this pattern on the top surface of the blocks. This test has been described by Waywell

& Bogg, 1999, p. 90) in the measurement of the spatial visualisation skill of qualified radiotherapy radiographers (see table 3.3, p. 108).

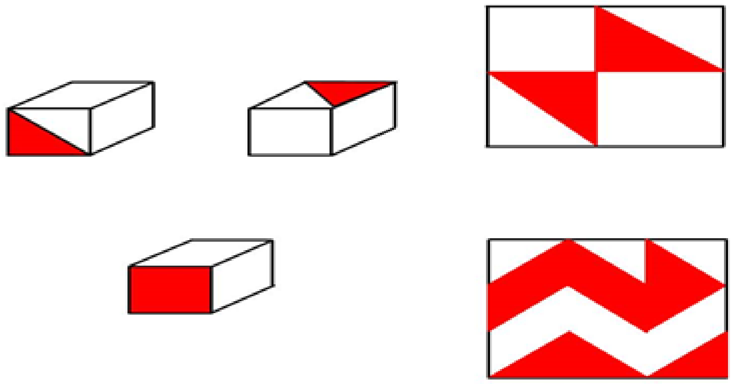


Figure 3.10: Facsimile of a WAIS block test item where the 3-D coloured blocks, shown on the left, have to be rotated so that their top surfaces patch the 2-D patterns shown on the right

3.8.2 Block comparison tests

An example of a block comparison test is the Purdue spatial visualisation test developed by Guay in the 1970`s and described by Bodner and Guay (1997, p. 8). Participants are shown a single complex 3-D object in its original and rotated form and are then asked to determine the correct rotation of a similar object rotated in the same direction as shown in Figure 3.11.

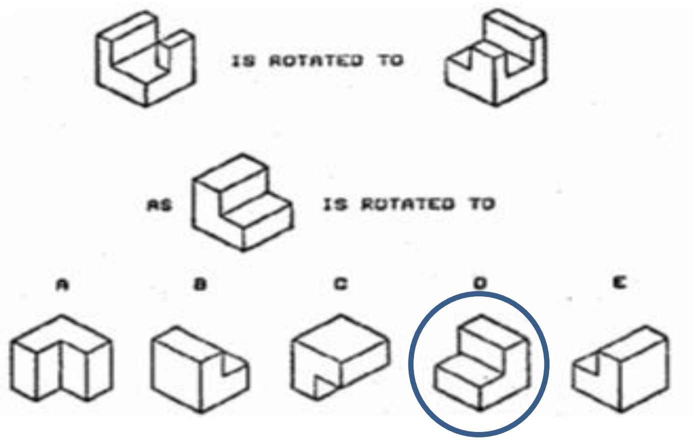


Figure 3.11: An example of a Purdue spatial visualisation test object where D is the correct answer

3.8.3 Block cutting (cross section) tests

Mental block cutting tests have been described by Titus and Horsman (2009, p. 243). In the example shown in figure 3.12 participants are shown an image of a 3-D block which has been sliced by a plane. The requirement is to identify the resultant appearance of the block at the intersection of the plane when viewed orthogonally.

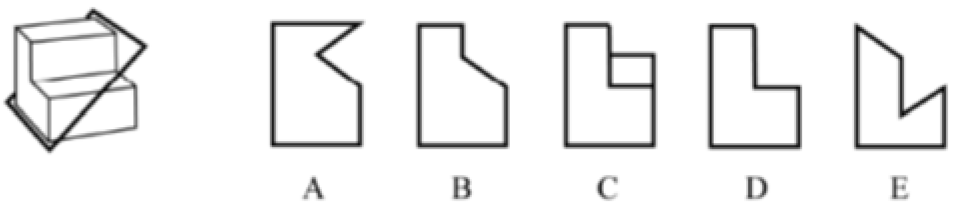


Figure 3.12: An example of a cutting planes object to test visual penetrative ability where option D is the correct answer

A more recent interpretation of the cutting planes test is the Santa Barbara Solids Test (SBST) developed by Cohen and Hegarty (2007, p. 180). The test consists of 29 test items made up of simple (single), joined or embedded objects cut by a vertical, horizontal or oblique plane. Figure 3.13 (a) shows a single object cut by a horizontal plane, (b) shows a joined object cut by a vertical plane and (c) shows an embedded object cut by an oblique plane.

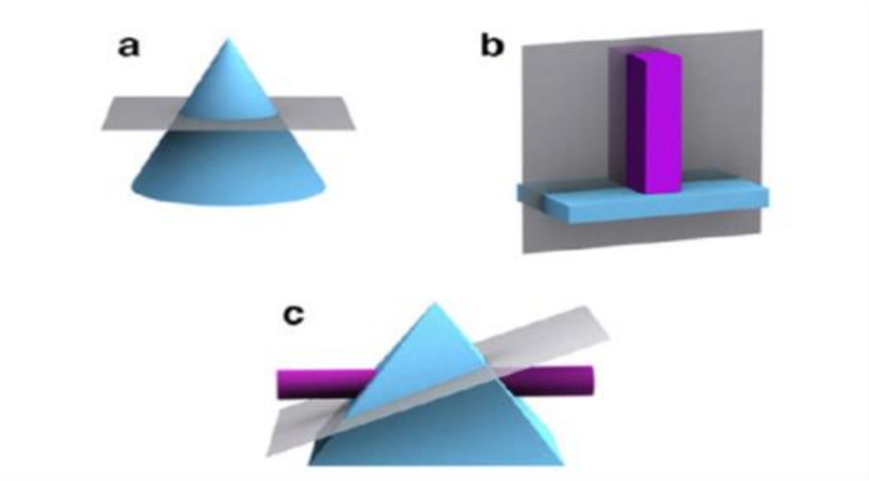


Figure 3.13: Santa Barbara Solids Test showing examples of (a) horizontal, (b) vertical and (c) oblique cutting planes. (Reprinted from Learning and Individual Differences, 22, Cohen & Hegarty, Inferring cross sections of 3D objects: A new spatial thinking test, 868 - 874 Copyright (2012), with permission from Elsevier)

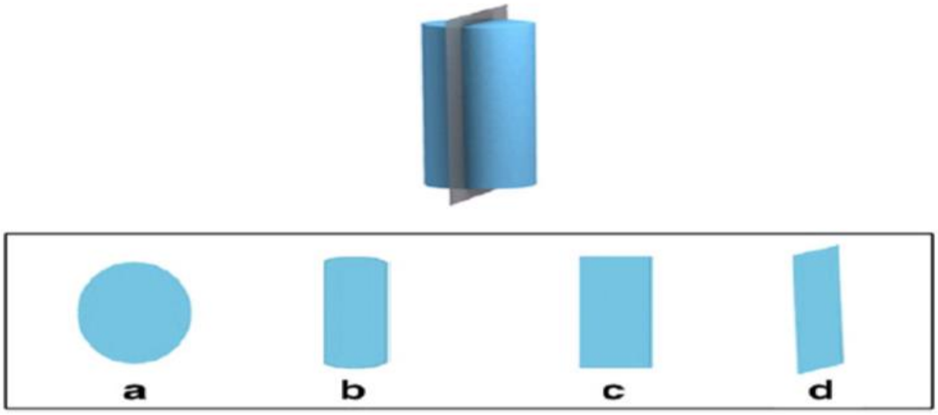


Figure 3.14: An example of a Santa Barbara Solids Test item with answer choices (Reprinted from Learning and Individual Differences, 22, Cohen & Hegarty, Inferring cross sections of 3D objects: A new spatial thinking test, 868 - 874 Copyright (2012), with permission from Elsevier)

Participants are required to identify the cross section that would be obtained if they were to mentally orientate their viewing position, referred to as their egocentric reference frame by Pittalis and Christou (2010, p. 195). By imagining this change in orientation they should view the cutting plane face on (as looking into a mirror). Each test item has four answer choices one of which is correct. As shown in figure 3.14, a single object is cut by an orthogonal vertical plane and the correct answer is c. The other three answer options are known as distractors and can be categorised as alternate (answer option a), combination (answer option b) or egocentric (answer option d).

3.8.4 Block rotation tests

The most widely reported test for mental rotation of blocks is the Vandenberg and Kuse MRT developed in 1978 (Vandenberg & Kuse, 1978, p. 599). Derived from objects developed by Shepard & Metzler (1971, p. 702), the test is designed to measure an individual’s ability to construct a mental image of a 3-D object as it is rotated in space. The target figures are composed of joined 3-D cubes which are displayed with a 15^{o} tilt from vertical and rotated around their vertical and horizontal axes by varying degrees. The answer choices (referred to as criterion figures) are composed of two correct rotations of each target figure and two mirror images of other target figures which are rotated around the X, Y and Z axes in 5^{o} steps. The test is available in its original format with 20 target figures, or in a redrawn format with 24 figures described by Peters et al., (1995, p. 42). The example shown below in figure 3.15 is taken from the original 20 item test, criterion figures one and three are correct rotations of the target figure.

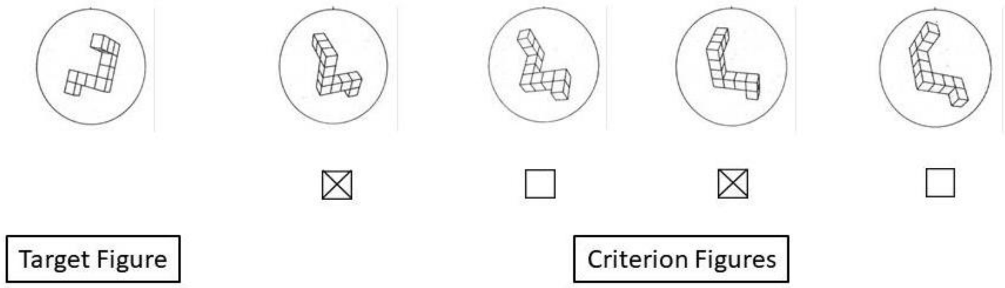


Figure 3.15: Example of a Vandenberg & Kuse Mental Rotation Test item, criterion figures 1 & 3 are the correct rotations

When Lohman (1979, p. 189) summarised the spatial factors referred to in chapter 3.3, p.76, he also identified how they may be measured. By applying Lohman's principles to the subsequent definitions for spatial visualisation skill proposed by Linn and Peterson (1985, p. 1482) it is possible to link the specific spatial visualisation domains to instruments for their measurement and to radiotherapy workflow processes as summarised in table 3.2.

3.9 Development of the critical evaluation checklist

The identification of individual performance in 3-D spatial visualisation tests can be considered as a type of cognitive diagnostic assessments. However, specific checklists for the systematic review of the quality of the spatial visualisation literature have not been reported in the literature. Therefore the first step in the evaluation of the literature required the identification and development of a suitable checklist from other sources. In a systematic review of the content of critical appraisal tools, Katrak, Bialocerkowski, Massy-Westropp, Kumar and Grimmer (2004, np) found 108 published papers reporting diagnostic testing studies. Of these, a total of 121 different appraisal tools were used, of which 104 (87%) were developed for a specific review, while just 16 (14.8%) were tools of a generic design. The review also identified 173 different checklist items and reported that the most frequently used items covered sample size, justification for the study, inclusion and exclusion criteria, randomisation details, reporting of study design, methodology and statistical analysis. Of all the reported tools only 11 (9.1%) included any reference to reliability and validity and 52 (43%) provided guidelines for completion of the checklist and scoring.

Table 3.2: Comparison of spatial visualisation components, application and measurement

Components of spatial visualisation, application in Radiotherapy & measurement			
Lohman (1979)	Lin & Petersen (1985)	Radiotherapy Application	Measurement Tools (From Lohman 1979)
Visualisation	Visualisation	Mental modelling & visualisation of internal anatomy and beam path (visual penetrative ability)	Paper Folding Form Board Surface Development WAIS Block Design Hidden Figures
Orientation	Relations and Perception	Relationships between & perception of normal anatomy & position of tumour target volume	Guilford-Zimmerman Orientation Test
Relations	Mental Rotation	Relationships between external positioning coordinates and position of internal anatomy (& changes in position). Supports manipulation of patient position	Card Rotations Flag Rotations Block Rotations

From these findings the authors recommend that researchers should select a critical appraisal tool to suit their needs and that evidence of the empirical basis of the tools` construction and guidelines for the interpretation of each item are included in any review. An initiative to improve the quality of clinical trial reporting led to the development of the Consolidated Standards of Reporting Trials (CONSORT) statement (Begg et al., 1996, p. 637). Subsequent revisions in the area of diagnostic test accuracy, resulted in the development of a 25 item checklist proposed by the Standards for Reporting of Diagnostic Accuracy (STARD) working group (Bossuyt et al., 2003, pp. 8-16), and the 14 item Quality Assessment for Diagnostic Accuracy Studies (QUADAS) checklist (Whiting, Rutjes, Reitsma, Bossuyt & Kleijnen, 2003, n. p.). Since their development, both STARD and QUADAS have received a growing acceptance as identified by Fontela et al., (2009, p. 2). As reported by Cook, Cleland and Huijbregts

(2007, p. 98) the checklists differ in their intent, STARD is a prospective tool with a structure that reflects the accepted structure of an article, while QUADAS is considered to be a retrospective instrument for the assessment of methodology and results. QUADAS has been recommended by the Cochrane Collaboration (Reitsma et al, 2009, p. 5) and the National Institute for Health and Clinical Excellence (Meads & Davenport, 2009, n.p.).

The Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy also indicates that up to 20 items may be used to meet individual review and evaluation requirements (Reitsma, Rutjes, Whiting, Vlassov, Leeflang & Deeks 2009, p. 7). As Whiting et al., (2006, n.p.) also propose, reviewers should consider how each checklist item can be applied and the importance of tailoring the guidelines and scoring to their review. This is a proposal supported by Leeflang, Deeks, Gatsonis and Bossuyt (2008, p. 892) who stated that consideration should be given to the inclusion of additional items which should be included in a QUADAS list. Based on the reports and recommendations outlined above, the decision to develop a checklist based on both STARD and QUADAS items for the evaluation of the spatial visualisation measurement reporting literature was made. Therefore, for this programme of research, the construction of the checklist for this critical evaluation was carried out by identifying all the possible QUADAS checklist items and aligning them as closely as possible to their STARD equivalents. The results from this alignment exercise and the resulting checklist can be found in appendix 3.

3.9.1 Critical evaluation of spatial visualisation test instruments

Using the search strategy identified in appendix 2 (p. 300), studies were selected for evaluation if they reported the measurement of spatial visualisation of cohorts with a similar demographic profile to those of pre-registration radiography cohorts. A total of 25 studies were identified (refer to PRISMA flow chart in appendix 2, p. 307) and divided into those reporting baseline or single time point testing only (n = 16, 64%) and those reporting pre and post intervention testing (n = 9, 36%). The justification for a cut-off point of December 2010 was related to the plan to commence the pilot phase of this programme of research in April 2011. A summary of the literature included in the evaluation findings can be found in table 3.3 below. It should be noted that when reviewing table 3.3, the nomenclature used by some of the authors when reporting their studies does not match the definitions provided for spatial visualisation skill by Lohman (1979, p. 189) and Linn and Peterson (1985, p. 1482) and proposed for this programme of research. When examining the test instruments that were used in these studies it became clear that spatial ability, spatial cognition, spatial skills and visuo-spatial ability were all being used in place of spatial visualisation. They are reported here as they have been employed by the authors of the studies evaluated. It is acknowledged that in doing so there is a risk of confusion which will do little to contribute to greater clarity. It does however serve to highlight the fact that, while clear definitions exist, the application of these definitions is inconsistent across the field as identified by Yilmaz (2009, p. 84).

Table 3.3: Summary of the literature reporting spatial visualisation skill measurement

Authors & Study title	Type of study	Population	Article aims, outcomes and recommendations	Comments / Relevance to this programme of research	Country
Appleyard and Coleman (2010) Virtual environment for radiotherapy training (VERT™) final project report	Randomised Factorial Study	Undergraduate Radiotherapy Students (n=103, male = 28, 27%)	**Aim:** Evaluation of VERT™ tracking technology and 3-D visualisation capability in a radiotherapy positioning task. **Outcome:** Moderately positive correlation between mental rotation ability and positioning performance (r .494). **Recommendations:** Measurement of baseline spatial visualisation skill to determine those most likely to benefit from using VERT™, but made no recommendation as to how this might be achieved or which test instruments were most appropriate.	Spatial visualisation skill measured using the Vandenberg & Kuse MRT.	UK
Clem, Anderson, Donaldson and Hdeib (2010) An exploratory study of spatial ability and student achievement in sonography	Case Control Study	Year 1 Ultrasound students (n=17, male = 2, 12%)	**Aim:** To determine the level of spatial ability and its impact on scanning performance following 30 hours of sonography training. **Outcome:** Reports Pearson-Product-Moment correlation between spatial ability and scanning performance after 30 hours tuition as .60 and that spatial ability can account for 36% of the variation in scanning performance. **Recommendations** the use of spatial ability testing for admissions screening.	No post intervention testing to determine change in spatial ability with experience / training.	USA
Cohen and Hegarty (2007) Sources of difficulty in imagining cross sections of 3D objects	Cohort Study	Psychology Students (level of study & demographics not stated)	**Aim:** To validate a new cross section cutting planes test (SBST) against two recognised tests (mental rotation and Visualisation of Views) and to determine individual differences in the visualisation of complex objects cut by orthogonal and oblique planes. **Outcomes:** Correlation between standard visualisation tests ($r = .47$, $p < .01$) and "significant" correlations with cutting planes test ($r = .50$, $p < .01$). Performance on SBST, mean correct scores 54%. Also analysed number of incorrect choices (egocentric foil), selected in 50% of incorrect answers.	Justification for use of non-standard scoring method for Vandenberg & Kuse MRT not explained. Is it possible to use the egocentric foil as an indicator for low spatial visualisation skill?	USA

			Recommendations: Further testing to determine impact / benefits of spatial training.		
Geiser, Lehmann and Eid (2006) Separating “rotators” from “non-rotators” in the Mental Rotations Test: a multigroup latent class analysis	Exploratory Cohort Study	High School & University Undergraduates	**Aim:** Study designed to explore the type of solution strategy employed for the 24 item Vandenberg & Kuse MRT. Utilises conventional (1 point for both items correct) & ratio scoring (total attempted ÷ total correct) strategies to compare differences between males and females. **Outcomes:** male: female performance difference is reduced by ratio scoring. Cronbach α is reported as 0.87 indicating good internal reliability. Also indicates possible solution strategies other than mental rotation. **Recommendations:** Investigate influence of experience and training on performance in mental rotation tasks.	Mean age of participants is lower than this programme of research but in the age range identified by Sorby as having potential difficulty in visualising. Solution strategy not as important in Radiotherapy / Imaging provided correct decision is made.	Germany
Green and Appleyard (2011) The influence of VERT™ characteristics on the development of skills in skin apposition techniques	Randomised Factorial Study	Year 1 (n = 23) & 2 (n = 21) Radiotherapy Undergraduates (Male = 11, 25%)	**Aim:** To determine the impact of virtual reality characteristics of VERT™ on the psychomotor skills & levels of confidence in a virtual representation of a complex radiotherapy positioning task. **Outcomes:** 39/44 participants (89%) reported increase in confidence. Mean Spatial visualisation skill score measured by 24 item MRT = 9.93/24 (41.4%) shows a moderate positive correlation with performance score on positioning task (r=.343, p .023). **Recommendations:** Explore the value of testing potential students to determine potential clinical competency.	No post intervention measurement of SVS to determine impact of VERT™.	UK
Hedman, Klingberg, Enochsson, Kjellin and Felländer-Tsai (2007) Visual working memory influences the performance in virtual image-guided surgical intervention	Self-Controlled Cohort Study	Medical Students n=28, (male = 14, 50%), age range 23-36)	**Aim:** To test the hypothesis that visual working memory correlates with performance in surgical instrument navigation. **Outcomes:** Pearson r correlations between “visual-spatial ability” and economy of movement (r = -.433, p = .021), time taken (r = -.543, p = .003). **Recommendations:** Study impact of working memory training on enhancement of performance.	MRT - standard scoring, but timing convention was not standard – provided a 1 minute break between subsets rather than 3.	Sweden

Hegarty and Waller (2004) A dissociation between mental rotation and perspective taking spatial abilities	Cohort Study	Psychology Undergraduates (no demographic data)	**Aim:** To test a new perspective taking tool & establish that perspective taking skills (egocentric transformation of own position) can be differentiated from those for mental rotation (object transformation). **Outcomes:** Reliability of MRT reported as 0.8 (from test manual & not determined by this study). Identifies that object perspective test requires different skills but may be some overlap. **Recommendations:** Need to use more than a single test to differentiate between perspective and mental rotation.	Goes some way to justify use of more than one test but used unconventional scoring for MRT (total correct – total incorrect).	USA
Kaufman (2007) Sex differences in mental rotation and spatial visualization ability: can they be accounted for by differences in working memory capacity?	Cohort Study	College students (n=100, male = 50) age range 16-18	**Aim:** To determine the difference between males & females in tests of spatial working memory & mental rotation. **Outcomes:** male performance in mental rotation tasks 22% better than females, effect size difference $t = 5.84$, $p < .0001$. **Recommendations:** Further studies to investigate sex differences should use more than one test.	Age range of students may give comparative data for that gained at baseline at the start of undergraduate study. Also covers age range identified by Sorby as having difficulty with visualisation and transformation of unfamiliar objects.	UK
Keehner, Tendick, Meng, Anwar, Hegarty, Stoller and Duh (2004) Spatial ability experience and skill in laparoscopic surgery	Not Specified – Cohort Study?	48 experienced laparoscopic surgeons: (Male = 42 [87.5%], Female = 6 [12.5%]) and 45 inexperienced laparoscopic surgeons (Male = 42 [93%], Female = 3 (7%)]	**Aim:** To determine if spatial ability is related to performance in minimally invasive videoscopic surgery & if high level cognitive function is moderated by practice based on the premise that that unfamiliar tasks rely on a high level of attention but with practice performance becomes more automatic as the tasks become proceduralised and therefore require less cognitive load. **Outcomes:** In the inexperienced group spatial ability measured by the Ekstrom paper folding test is a significant predictor of performance ($r=.39$, $p < .01$), while the experienced group shows no significant correlation (no values reported). **Recommendations:** If individuals with low spatial ability are slow to learn new skills initially but can reach an acceptable skill level, the impact of focussed accelerated training should be	See recommendations	USA

			investigated.		
Keehner, Lippa, Montello, Tendick and Hegarty (2006) Learning a spatial skill for surgery: how the contributions of abilities change with practice		22 non-medical undergraduates (programme and level of study not reported)	**Aim:** To determine the longitudinal relationship between spatial ability and performance during laparoscopic surgical skill acquisition. **Outcomes:** All individuals reported as learning the surgical task to similar proficiency, but rate of improvement decreases with practice. Performance is dependent on ability to maintain & transform spatial information but no correlation values between spatial ability and performance are presented. **Recommendations:** Explore whether focus should be on using spatial ability tests to select most able students or to develop virtual environments to train all students to proficiency.	Spatial visualisation tests were MRT & Visualisation of Views but scoring convention is not stated.	USA
Luursema, Buzink, Verwey and Jakimowicz (2010) Visuo-spatial ability in colonoscopy simulator training	Cohort Study	15 Medical Trainees with no experience of colonoscopy Mean age 25, range 21-29, 5 (Male n = 5 [33%])	**Aim:** To understand the role of "visuo-spatial" ability in the development of endoscopic surgical skills. **Outcomes:** A negative correlation ($r = .69$, $p <.01$) between visualisation skill & time taken is reported which indicates those with better visualisation perform faster, Repeated measures ANCOVA shows a significant between subjects effect for visualisation skill and time taken to complete 2 endoscopy navigation tasks $F(1,14) = 10.7$, $p <.01$ and $F(1,14) = 8.6$, $p <.02$ indicating that participants with high visualisation skill improve on the time taken to complete tasks faster than those with low spatial visualisation skill. **Recommendations: T**asks requiring visuo-spatial demands are included early in training.	Test battery includes MRT & Visualisation of Views Test (tests for visualisation) together with card rotation & figure comparisons from Edstrom's Kit of Factor-Referenced cognitive tests (tests for general intelligence). While the authors state that mean scores and SD were calculated for each of the tests no results have been presented.	Holland
Parsons, Larson, Kratz, Thiebaux, Bluestein, Buckwalter and Rizzo (2004) Sex differences in mental and spatial	Case Control Study	22 undergraduate / graduate students, age not stated. Male	**Aim:** to validate a virtual reality spatial rotation tool against the paper version Vandenberg & Kuse 20 item MRT. **Outcomes:** Performance differences on paper based MRT favour males ($t(42) = -3.27$, $p = .002$) but not the virtual reality spatial rotation test ($t(20) = -0.18$,	Timing of the paper based MRT was 5 minutes per split half and scoring was the non-standard 2 points per item resulting in a possible	USA

rotation in a virtual environment		= 10 (45.5%)	p = 0.86). **Recommendations:** Future studies using larger sample sizes to determine which spatial factors, e.g. task demands and visual working memory influence performance.	total of 40.	
Peters, Laeng, Latham, Jackson, Zaiyouna and Richardson (1995) A redrawn Vandenberg and Kuse mental rotations test: different versions and factors that affect performance	Cohort Study	636 Undergraduate students across Science (males=135 (43%), females=177) & Humanities (males= 102(31.5%), females= 222) programmes	**Aim:** To confirm the male / female performance differential in a redrawn & expanded (24 item) MRT using students registered on different programmes & to determine relationship between performance & handedness / computer gaming. **Outcomes:** Males outperform females across both programmes (Male: Female BSc mean =61.6%: 43.3%, Male: Female BA mean = 50.4%:34.2%) for sex F (1,632) = 135.75, p<.0001 & programme F (1,632) = 46.77 p<.0001. The relationship between handedness & performance accounted for 1.3% of the effect and while males played computer games more frequently than females there was no correlation with mental rotation performance for either sex. **Recommendations:** Further studies to explore links between mental rotation performance and spatial activities.	Authors report that the MRT remains the most convincing test in terms of demonstrating sex difference. Are these difference seen in Radiography, a predominantly female profession?	Canada
Smoker, Berbaum, Luebke and Jacoby (1984) Spatial perception testing in diagnostic radiology	Case Control Study	21 post graduate Radiology residents (8 1st years, 7 2nd years, 6 3rd years) and 8 faculty members. Age ranges & sex not stated	Based on observations that residents differ in ability to make judgements based on complex imaging **Aims:** To determine if perceptual ability can be a predictor of success in skill acquisition, if it is influenced by length of training time, can individual differences be detected, and can it be used for selection purposes. **Outcomes:** Ranking performance in the Lego Brick test (part 2) & comparing it with the rank order for Faculty rating of image interpretation shows “considerable correspondence” (Spearman rho = .50, p <.025) but Thurstone`s surface development test is not predictive of rank order (Spearman rho = .18, p > .05).Performance on part 2 of the Lego brick test and the surface	Used a Lego brick visuo-construction test & Thurstones surface development (mental paper folding) test. Reports that other paper based measures of spatial visualisation are 2-D tests of a 3-D ability but a 3-D construction test such as the Lego brick test requires motor skills as well as mental	USA

			development test are highly correlated (Spearman rho = .78, p <.001), t tests to compare test performance of participants at different levels of training showed no difference (no values reported). Lego brick test appears to be better predictor of performance. **Recommendations:** Need for further studies before advocating use of the tests for selection.	visualisation for the transformation from a 2-D image.	
Waywell and Bogg (1999) Spatial ability assessment: an aid to student selection for therapy radiography training	Cohort Study	54 Radiotherapy Radiographers – 1 male (mean age not stated, range 18 – 54). Mean experience 9.3 years, range 1 – 31 years. Number of students not stated.	**Aims:** To determine if Radiotherapy Radiographers have a greater spatial ability than normal population, to establish validity of new spatial ability test and investigate link between tests of spatial ability and a test of clinical competence using Lego brick test and WAIS-R Block Test. **Outcomes:** Pearson product moment correlations - WAIS-R and Lego Brick test (r = 0.57, p < 0.0005), the clinical task and WAIS-R (r = 0.128, p > 0.1) & the clinical task and Lego Brick test (r = 0.23, p > 0.1). Using a 2 sample t test, Radiographers have a significantly better spatial ability on the WAIS-R test than the normal population - age group 25-34 (t = 3.44, p <.001) & 35-44 (t = 2.92, p <.01). Between 19 & 65% of participants did not complete the Lego brick test. **Recommendations:** Examine whether psychometric testing has a role in determining clinical potential & could it be used in conjunction with traditional methods to aid selection of prospective students.	Successful participants appeared to demonstrate a range of solution strategies which appeared to be dictated by how Lego brick subtest 1 had been demonstrated. If bricks from the top view were assembled first then this was the approach that was followed by the majority – copying rather than mental transformation? Also is there any link between those who did not complete & their scores on the WAIS-R and Lego brick test? Not explored in this study.	UK
Zacks, Mires, Tversky and Hazeltine (2000) Mental spatial transformations of objects and perspective	Cohort Study	48 undergraduates (22 male (46%), 26 female, degree programme and level of study	**Aim:** To determine the relationships / disassociation for mental & perspective transformation using an object based transformation test – mental rotation & 2 perspective transformation tests, map reading & perspective taking. Based on the premise that each is mediated by different processing regions of the brain. **Outcomes:** Tests were reported as having a "pair wise" correlation with each other ranging from 0.38 – 0.42 but	Object based transformations are imagined rotations of an object relative to the reference frame of the environment while egocentric perspective transformations are	USA

		not specified)	with no further detail. It should also be noted that the scoring for MRT was non-standard (1 point for each correct item with 1 point deducted for each incorrect choice) making comparisons with other studies challenging. **Recommendations:** Need to establish a unified biological & computational framework for transformations.	imagined transformations of an individual's point of view relative to the reference frame. Each require different strategies but could there be overlap between them because some individuals may solve mental rotation tasks by imagining a change in perspective rather than an object rotation. When positioning patients for delivery of radiotherapy treatment, the correct decision is more critical than solution strategy.	
Pre and Post Intervention Studies					
Alias, Black and Gray (2002) Effect of instruction on spatial visualisation ability in civil engineering students	Case Control Study	Civil Engineering Students from 2 colleges. 1 college - control group (n=28, males = 20, 71%), other college - experimental group (n=29 males =18, 62%)	**Aim:** To determine the impact of related spatial activities in the improvement of spatial ability using a control group who engaged with standard curriculum & an experimental group who had additional object manipulation & free hand sketching exercises. **Outcomes:** Increase in post intervention test score for experimental group = 5.8%. A statistically significant gain in score between experimental & control groups using repeated measures 2-way ANOVA. **Recommendations:** Spatial skills training should be integrated across the engineering syllabus & remedial lessons made available to those students with low baseline spatial visualisation scores.	Supports evidence for identification of those with less well developed spatial visualisation skill.	Malaysia
Gorska, Sorby and	Cohort	1350	**Aim:** To determine the impact of additional graphics	Study involved 3	USA /

Leopold (1998) Gender differences in visualization skills – an international perspective	Longitudinal Before & After Study	undergraduate engineering students. Demographic split for male: female reported in terms of number taking each test rather than overall split, age not reported.	and descriptive geometry interventions on performance in a combination of the Purdue Spatial Visualisation Test: Rotations, the Vandenberg & Kuse MRT and the Mental Cutting Test. **Outcomes:** Gain scores in performance across both sexes ($p<.005$) but post intervention scores test scores of females did not reach the pre intervention test scores of males. Age and dominant hand were not found to be significant background variables. **Recommendations:** Future research should determine if there are differences in the preferred learning style of females and how these may be used to improve visualisation skill.	Institutions, not all students received the same combination of visualisation tests. Suggestion that females start engineering programmes with lower visualisation skill and while improvements are seen there is still a performance gap compared to males. Radiography is a predominantly female profession, will the findings be replicated?	Europe
Hegarty, Keehner, Khooshabeh and Montello (2009) How spatial abilities enhance and are enhanced by dental education Experiment 1	Case Control Study	Dentistry students, 2nd year n = 82, male = 47 58%) and 4th year (n = 36, male = 24, 69%)	**Aim:** To determine if there was a correlation between a new dentistry visualisation test and traditional tests (mental rotation and Visualisation of Views) & whether dental education enhanced spatial ability in general or the ability to imagine cross sections specifically. **Outcomes:** Significant positive correlations between all tests ($p<.01$) and (Cronbach's alpha) ranging from 0.63 - 0 .8 indicating reasonable to good internal reliability, the study did not provide evidence to support enhancement of spatial ability through education. **Recommendations:** This study used 2nd year students who had already received restorative dentistry training, therefore repeat study with novice learners with no prior experience.	Unconventional scoring of MRT (reported possible maximum of 80 compared to recommended 20 or 24 depending on version).	USA
Hegarty, Keehner, Khooshabeh and		Year 1 Dentistry students, n=79,	**Premise:** a) if dental education aids the representation and transformation of spatial objects then 1st & 4th	Identifies change in spatial visualisation over	USA

Montello (2009) How spatial abilities enhance and are enhanced by dental education Experiment 2		(male = 52 66%), year 4 Dentistry students, n=65 (male = 43 66%), Year 1 Psychology undergraduates, n = 62 (male =31 50%)	year dental students would achieve higher test scores than the 1st year psychology students, b) if dental education improves the ability to imagine cross sections then students would perform better in the cross sectional tests than they would MRTs, c) if the ability to imagine cross sections does improve then measures of spatial ability will be predictive of performance as knowledge and experience increases. **Outcomes:** At first testing, Psychology students performed less well than dentistry students on all cross section, visualisation & MRTs. Significant differences are reported for the novel object cross section test F (2,188) = 6.95, p<.01, the tooth cross section test F (2, 188) = 10.155, p<.001, MRT F (2,188) = 5.59, p<.01 & Visualisation of Views F (2,188 = 3.21, p <.05. Examining the scores of the 1st year dental students shows that their performance on the novel object cross section test did not differ between the 1st (M=5.4,SD=2.0) & 2nd (M=5.8, SD=2.4, t(67) = 1.81, p=.07, d = .17) but performance on the tooth cross section test was better at 2nd test (M=11.3, SD4.3) than the 1st (M=8.6, SD = 4.4, t(67) = 5.83, p<.001, d= .6) indicating that dental education improves a specific ability to imagine cross sections of teeth but not cross sections in general. **Recommendations:** Training for spatially demanding fields need to be informed by analysis of those aspects of spatial performance that are domain specific and can be learned.	time and with training interventions.	
Hoyek, Collet, Rastello, Fargier, Thiriet and Guillot (2009) Enhancement of mental rotation abilities and its effect on anatomy learning	Case Controlled Before & After Study	Group 1 (male=10, female = 6) anatomy sessions plus 12 x 20 minutes of mental rotation training. Group	**Aim:** To determine if mental rotation training has an impact on mental rotation performance & does this performance transfer to the learning of functional anatomy? **Outcomes:** An independent t test showed greatest performance enhancement at post testing for the group receiving mental rotation training compared to the anatomy only group (t = 4.14, p = <.001) and the control group (t = 4.03, p = <.001) There was no	Results support findings in previously published literature that good mental rotation skills may facilitate learning. Do these findings support transfer of good	France

		2, (male=10, female = 6) anatomy sessions only. Group 3, control group (male=11, female = 5) no anatomy training.	significant difference between the anatomy group and the control group (t = 0.3, p > 0.05). **Recommendations:** None specified.	mental rotation / visualisation skills to the performance of a clinical positioning task?	
Jansen and Pietsch (2010) Physical activity improves mental rotation performance	Cohort Before & After Study	Physical education students n = 88, male = 43 (49%) mean age 23.7, females = 45 mean age 22.5. Physical activity (male = 22, female = 22) + cognitive activity (male = 21, female = 23) groups	**Aim:** To determine the impact of physical activity on spatial cognition. **Outcomes:** No significant difference found between groups in the first MRT F (1, 87) = 0.294, but a significant main effect of group in the second test F (1, 87) = 5.03, p<0.05 indicating that there was a significant improvement in score by the physical activity group. As no improvement was observed in the cognitive activity group, improvements are not due to practice effects. **Recommendations:** Further studies to determine if improvement varies with different types of physical activity.	MRT with standard scoring but second half applied following 45 minutes of physical (running, jumping, skipping, & callisthenics) or cognitive (didactic lecture). In Radiography cohorts, is there a relationship between participation physical activity (team / individual sports) & baseline spatial visualisation skill?	Germany

Németh (2007) Measurement of the development of spatial ability by Mental Cutting Test		250 first year engineering students (demographics not stated)	**Aim:** To determine the relationship between a programme of descriptive geometry and spatial visualisation skill using a mental cutting test applied at the beginning of semester one and then repeated at the end of semester two (test-retest time period not stated in terms of weeks). **Outcomes:** Results are reported graphically as proportion of correct answers at before & after intervention testing and stated as significant at the 98% confidence interval (*t* test values not stated). **Recommendations:** Further longitudinal studies with regular testing are required.	Lack of specific detail relating to overall teaching time, duration of each session and the graphical presentation of the results rather than statistical convention makes comparison with other similar studies challenging.	Hungary
Rafi, Anuar, Samad, Hayati and Mahadzir (2005) Improving spatial ability using a web based virtual environment (WbVE)	Randomised Controlled study	Preservice undergraduate teachers, n = 98 (male = 46, 47%). Experimental group (n=49, male = 26) WbVE training, control group (n=49 male=23) conventional classroom-based teaching	**Aim:** To determine the benefit of a web based virtual environment (WbVE) in the development of spatial understanding. Before & after testing using a combined spatial visualisation / MRT with 2-D projection & standard Vandenberg & Kuse mental rotation items. **Outcomes:** Reliability of mental rotation items reported as .68 & projection images .63. Before intervention gender differences significant for mental rotation $F(1, 96) = 8.23$, $p<0.01$ across both groups, after intervention results show significant group differences for the experimental group $F(1, 96) = 8.35$, $p<0.01$ & for females in that group $F(1, 50)$ 5.32, $p<0.05$. Learning through WbVE is reported as being more effective than classroom teaching. **Recommendations:** None specified.	Supports the use of online / electronic visualisation tools for the development of spatial visualisation skills.	Malaysia
Russell and Churches (2010) What do we really want to know about spatial visualization skills among engineering	Cohort Before & After Study	89 year 1 undergraduate engineering students (demographic data not	**Aim:** To compare first year student test scores in mental rotation and cutting tests with other groups of students in the published literature and to determine the impact of previous drawing experience and additional (4 hours) drawing tuition on these scores. **Outcomes**: Descriptive statistics only. Test scores for mental rotation show an increase of 8% for both groups. Cutting test score	Other studies report that only significant time on task delivers larger gains. Small performance difference demonstrated following 4 hours tuition	Australia

students?		collected)	decreased in group with previous drawing experience, further exploration shows that time taken to complete was 5 – 10 minutes shorter in the second attempt. **Recommendations:** To determine if more structured tuition develops skills faster, investigate visualisation skills as part of a more complex (unspecified) learning context and to conduct comparative studies at different points in undergraduate skill development.	but no inferential statistical analysis conducted. Spatial visualisation skill training in the Radiotherapy clinical environment is ad- hoc & nonspecific. Will significant performance gains be demonstrated when compared to Diagnostic Imaging student performance?	
Terlecki, Newcombe and Little (2008) Durable and generalized effects of spatial experience on mental rotation: gender differences in growth patterns	Repeated Measures controlled Study	79 “Introductory Psychology” undergraduate students. Males = 28 (35%). 46 in the repeated testing group & 33 in the training group.	**Aim:** To determine if performance gains were different for males and females, to determine if training with video games has an impact on mental rotation skill, to determine whether gains from simple repeated testing (simple practice) of a mental rotation task or playing a video game (spatially relevant focussed training) are sustained over time & determine if performance gains from focussed training exceed those from simple practice**. Outcomes:** Post intervention testing on the 20 item MRT with standard scoring was significantly better than before testing (d = 3.72, p<.01) and sustained at retesting between 2-4 months (d=3.72, p<.01) indicating that mental rotation growth trajectories are sustained over long periods of repeated testing & training and show improvement over the course of 1 semester. **Recommendations:** More training studies are required to confirm the utility of video game training.	Women with lower spatial experience at the start of training may have difficulty in engaging with interventions of a spatial & visual nature. While women appear to show slower initial increase in mental rotation performance, the improvement is sustained over a 12 week period and demonstrates greater overall growth when compared to males over the same period. Is this replicated in Radiography which is a largely female	USA

				profession?	

3.9.2 Summary of methodological quality

Each study selected for evaluation was assessed for each QUADAS/STARD checklist quality item as "yes", "no", "unclear" or "not applicable". This was in line with the Cochrane Collaboration handbook recommendations for the presentation of assessment of methodological quality results in tabulated form (Reitsma et al., 2009, p. 19). The methodological quality assessment for the spatial visualisation measurement studies evaluated in this programme of research is shown in table 3.4a (single time points) and 3.4b (pre and post intervention time points).

Table 3.4 a: Methodological quality summary for single time point studies reporting spatial visualisation skill measurement

Single Time Point Testing													
Author	Participant Spectrum	Selection Criteria	Objectives Specified	Reference Test	Replication	Differential Verification	Test-Retest Time Period	Blinding	Interpretation & Review	Interpretable Results	Withdrawals	Validity & Reliability	Yes Items
Appleyard													4
Clem													9
Cohen													9
Geiser													9
Green													8
Hedman													7
Hegarty													7
Kaufman													7
Keehner 2004													6
Keehner 2006													9
Luursema													5
Parsons													5
Peters													6
Smoker													6
Waywell													8
Zacks													5
Yes	13	9	11	15	10	16	0	1	14	10	6	5	
No	3	4	5	1	5	0	0	15	0	5	4	10	
Unclear	0	3	0	0	1	0	0	0	2	1	6	1	
Not Applicable	0	0	0	0	0	0	16	0	0	0	0	0	

Table 3.4 b: Methodological quality summary for pre and post intervention studies reporting spatial visualisation skill measurement

Author	Participant Spectrum	Selection Criteria	Objectives Specified	Reference Test	Replication	Differential Verification	Test-Retest Time Period	Blinding	Interpretation & Review	Interpretable Results	Withdrawals	Validity & Reliability	Yes Items
	Pre and Post Intervention Testing												
Alias	Yes	No	Yes	Yes	No	Yes	Unclear	No	Yes	Yes	Unclear	Yes	7
Gorska	Yes	No	No	Yes	Yes	No	No	No	Yes	Yes	Yes	No	6
Hegarty	Yes	No	Yes	Yes	No	Yes	Unclear	No	Yes	No	Yes	Yes	7
Hoyek	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	9
Jansen	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Unclear	Yes	No	8
Németh	Yes	No	Unclear	Yes	Yes	Yes	Unclear	No	Yes	Yes	No	No	6
Rafi	Yes	Yes	Yes	Unclear	No	Yes	Yes	No	Yes	Yes	Unclear	Yes	8
Russell	Yes	No	Yes	Yes	Yes	Yes	Unclear	No	Yes	Yes	Yes	No	8
Terlecki	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	10
Yes	9	3	7	8	6	8	4	0	9	7	5	3	
No	0	6	1	1	3	1	1	9	0	1	2	6	
Unclear	0	0	1	0	0	0	4	0	0	1	2	0	
Not Applicable	0	0	0	0	0	0	0	0	0	0	0	0	

Key:

(green)	Yes
(amber)	Unclear
(red)	No
(grey)	Not Applicable

The calculation of an overall score for each study is one method of incorporating quality assessment into a review of diagnostic accuracy. This will combine individual checklist items from the quality assessment tool to provide an overall single score. But one of the challenges encountered when adopting this approach is the determination of the relative weight and importance of each item and may lead to different conclusions regarding the comparative quality of studies. This led Whiting, Harbord and Kleijnen (2005, n.p) to suggest that quality scores should not be used in results tables for systematic reviews. The alternative would be a component approach where the association of quality items are investigated individually and weighted equally. For the evaluation reported here, the numbers of items identified as yes were allocated one point, with those items identified as no, unclear or not applicable were awarded zero, and those items which were identified as being not applicable were not scored. A summary of the number and percentage of items identified as yes, no or unclear for all studies is presented in table 3.5. Across all studies, there were a total of 284 checklist items, of which 179 (63%) were identified as yes, 83 (29.2%) were no and 23 (7.8%) were unclear. Further analysis of individual articles demonstrated that the mean number of yes items was 7.16 (SD = 1.57, range 4-10).

Table 3.5: Summary of total checklist item scores for each category

	Participant Spectrum	Selection Criteria	Objectives Specified	Reference Test	Replication	Differential Verification	Test-Retest Time Period	Blinding	Interpretation & Review	Interpretable Results	Withdrawals	Validity & Reliability
Yes	22 88%	12 48%	18 72%	23 92%	16 64%	24 96%	4 16%	1 4%	23 92%	17 68%	11 44%	8 32%
No	3 12%	10 12%	6 24%	2 8%	8 32%	1 4%	1 4%	24 96%	0	6 24%	6 24%	16 64%
Unclear	0	3 12%	1 4%	0	1 4%	0	4 16%	0	2 8%	2 8%	8 32%	1 4%
Not Applicable	0	0	0	0	0	0	16 64%	0	0	0	0	0

In addition to providing a summary and analysis of the individual checklist items for each of the studies, the Cochrane handbook (ibid) also recommends the use of a methodological quality graph. This provides a stacked bar chart to demonstrate the percentage of studies that rate the items yes, no or unclear and provides a summary of the overall study quality across the whole review as shown in figure 3.16. Examination of each checklist item showed that item six (differential verification) had the highest score of 96% while the highest scoring no item was number eight (blinding), also scored 96%. The reporting of blinding across all studies was reported in just one of the single time point measurement studies, that of Clem, Anderson, Donaldson and Hdeib (2010, p. 166). This would indicate that there was high consistency in the application of the same test to all participants.

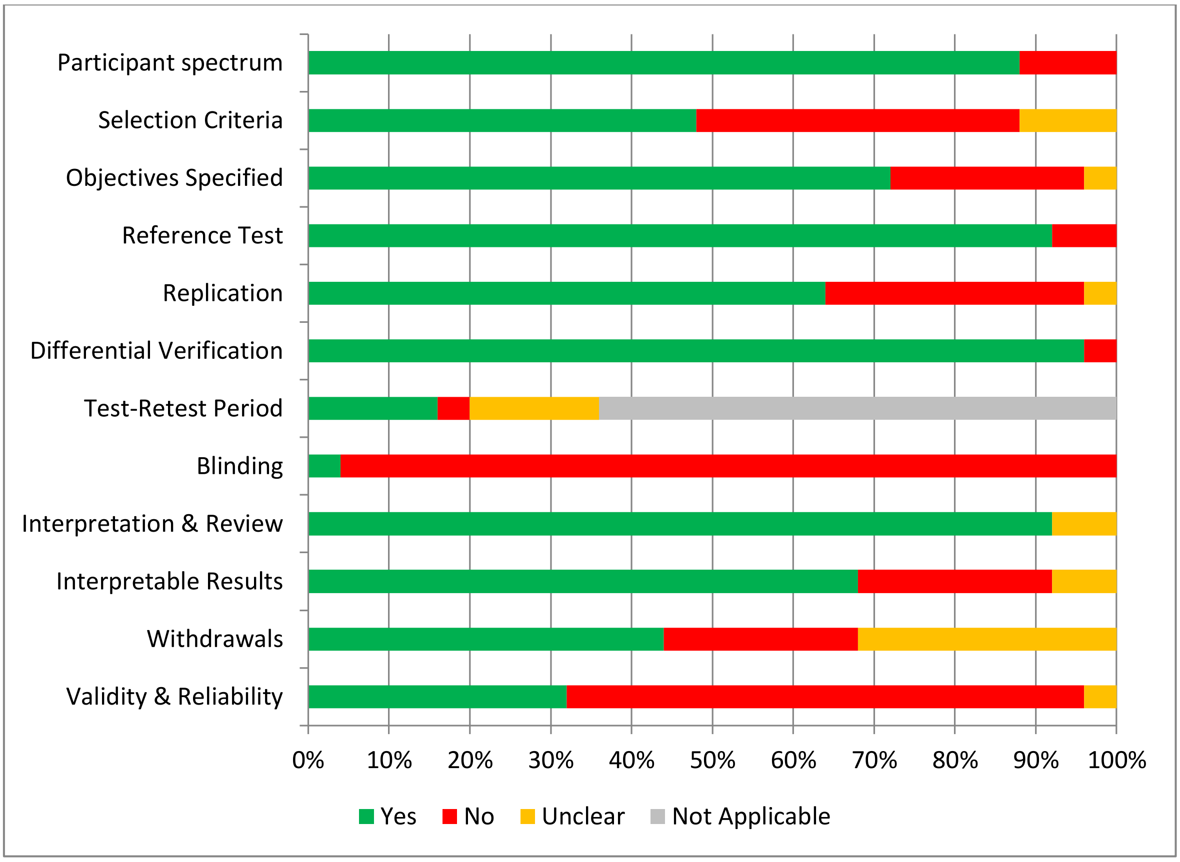


Figure 3.16: Methodological quality graph for all studies

3.10 Critical evaluation findings and discussion

The critical evaluation of the literature reporting the measurement of spatial visualisation identified three themes for discussion, these were participant profile and recruitment, test administration and scoring convention, and interpretation and reporting of results, which are discussed in turn below:

3.10.1 Participant profile, recruitment and randomisation

Study participants were recruited predominantly from one programme in a single institution. Clem et al., (2010, p. 168) recruited trainee sonographers, from multiple levels of study within a single programme while Green and Appleyard (2011, p. 179) reported on the relationship between performance on the MRT and a simulated radiotherapy positioning task (refer to chapter 2.8, p. 55 for detail). Other studies recruited participants from multiple cohorts from multiple institutions (Gorska, Sorby, Leopold, 1998, p. 11). While the spectrum of participants was described in 22 studies (88%), the justification of the selection criteria was explained in just 12 (46%).

The majority of studies (n = 19, 76%) recruited cohorts of undergraduate students with a similar demographic profile to that of the pre-registration Diagnostic Imaging and Radiotherapy cohorts envisaged for this programme of research. Two of the exceptions were Kaufman (2007, p. 213) who recruited 100 post – 16 college students (50 males and 50 females) from the UK and Geiser et al., (2006, p. 264) who recruited 1695 German High School students (male n = 843, mean age = 16.7, SD 6.9, female n = 850, mean age = 16.8 SD 6.3). These studies have been included in the evaluation because the participant age range is similar to the age group identified by Sorby (2007, p. 1) as having potential difficulty with the visualisation and transformation of unfamiliar objects.

As radiography education programmes tend to recruit from a wide spectrum of prospective students, it is possible that there will be a proportion of learners in the 18 year old category with less well developed spatial visualisation ability and therefore at a disadvantage when attempting to mental visualise the complex relationships between tumour target volumes and normal anatomy. Two studies that were also included were Keehner et al., (2004, p. 72) who reported on the links between spatial skill and performance of general surgeons and Smoker et al., (1984, p. 1106) whose study involved trainee Radiologists. Both studies explored the relationship between spatial visualisation skill and the performance of technical tasks in groups of novices and experts.

Where recruitment occurred from a specific subject pool, for example in Psychology, participants were offered course credit (Cohen & Hegarty, 2007, p. 180), cash payment (Keehner et al., 2006, p. 491) or a choice of either course credit or payment (Zacks, Mires, Tversky and Hazeltine, 2000, p. 319). While this is considered to be standard practice in some programme areas or institutions, Leentjens and Levenson (2013, p. 395) have suggested that students who receive course credit may self-select studies because they are performing less well in general coursework and therefore see participation as an easy way of making up this deficit. This recruitment approach may have the potential for selection bias which does not appear to have been considered in any of the studies evaluated. It could also be the case that those students who are struggling are doing so because they have less well developed spatial visualisation skill which may lead to the skewing of population data in visualisation skill studies. Four studies (16%) reported clearly that participation was voluntary (Geiser et al., 2006, p. 263, Hedman et al., 2007 p. 2045, Peters, et al., 1995, p. 41 & Rafi et al., 2005, p. 709), while the remaining studies just indicated that participants “were recruited” or “took part” but provided no additional details. Evaluation of the reproducibility of the studies indicated that seven studies (28%) did not provide information relating to randomisation methods or failed to provide enough detail about the conduct of the testing procedure. Two studies (8%) reported that they ensured homogeneity between intervention groups by matching participants by gender, age and mental rotation ability (Hoyek et al., 2009, p. 202) and Parsons, et al., (2004, p. 557) who matched participants based on their MRT performance score.

3.10.2 Test administration and scoring convention

The test instruments employed in the studies ranged from simple 2-D shape comparison tests to the more complex 3-D mental rotation and cutting plane tests. The most widely used test instrument was the Vandenberg and Kuse MRT in either its original or redrawn format reported in 17 (68%) of the studies. In seven studies (28%) a single test was administered while in the remainder multiple instruments were utilised. All studies, apart from one, applied the same test(s) to all participants. The exception was the study reported by Gorska, Sorby and Leopold (1998, p. 11) who employed four tests in different combinations.

3.10.3 Interpretation and reporting of results

One of the challenges encountered when interpreting study results was the range of alternative scoring conventions employed which were different from those advocated by the original test developers. This was the case in the administration of the Vandenberg and Kuse MRT particularly, with Geiser et al., (2006, p. 265) and Hegarty and Waller (2004, p. 180) employing a ratio score of the number of correct items ÷ number of incorrect items rather than awarding one point for each correct pair identified. Other scoring methods included that used by Zacks. Mires, Tversky and Hazeltine (2002, p. 322) who awarded one point for each correct item (rather than one point for each correct pair) and then deducting a point for each incorrect item identified, leading to a minimum score of -40 and a maximum of +40. Cohen and Hegarty (2007, p. 180) referred to a possible score of 80 with no explanation or justification of how this score was derived. These alternative scoring methods mean that any comparisons with the findings and conclusions of other studies need to be treated with some caution.

Results were presented and interpreted using a combination of descriptive and inferential methods in 23 (92%) of the studies evaluated. One of the exceptions was Németh (2007, pp. 125-7), who presented graphical data to show the percentage of correct answers gained in the mental cutting planes test at the start and end of a semester. The other, (Russell & Churches, 2010, p.571) only reported comparative, before and after data for mean, standard deviations and 95% confidence intervals for the mental rotation and cutting planes tests.

When reporting reliability, Schuwirth and Van Der Vleuten (2004, p. 805) refer to the calculation of Cronbach Alpha coefficients, but also point out that the achievement of a high coefficient is not a goal on its own. The results of any reliability analysis should inform the decisions that are made based on the results of the test. Based on the test results, some decisions can be made with certainty, however if the reliability coefficient is low then more prudence is required when drawing conclusions. Therefore, what is important is the reproducibility of the decisions that are made on the basis of the test results. The validity and reliability of test instruments was reported in eight (32%) of the studies evaluated. The test instruments employed in these studies were the Vandenberg & Kuse MRT, Guay`s visualisation of views, Thurstone's surface development test, the revised Minnesota form board test and the SBST. The remaining studies referred to previously published reports to support the claims of instrument validity and reliability.

3.11 Selection of authentic spatial visualisation test instruments for radiotherapy

The general psychometric approach to assessment as a scientific model has been identified as playing a major role in improving the quality of skill assessment (Schuwirth & Van der Vleuten, 2006, p. 296). But they also acknowledged that the practical feasibility of testing is also an important consideration alongside with the intended educational goals of testing and its context. Educators should also consider what they want to test and then identify and design a test that is most likely to assess learner performance in a specific domain as advocated by Srinivasan, Hwang, West and Yellowlees (2006, p. 509). This is a view supported by Lammers et al., (2008, p. 1081) who suggested that assessment should be procedure specific. These views are similar to those of Rust and Golombok (2009, pp. 32-33) who report two approaches to assessment and test design. The first, known as the functional approach, proposes that the design of a test instrument should be determined by its use and that what it measures has little value. This viewpoint suggests that a good test will be defined as one which will differentiate between people who perform well in a task and those who do not. This approach may have a tendency to lead to an all or nothing interpretation. The alternative, referred to as the trait approach, proposes that individuals are not entirely good or entirely bad at a task, therefore goodness and badness exists on a continuum. The two approaches lead to different types of testing known as functional and trait based testing. A functional test would be designed to determine if an individual has the skill requirements needed to carry out a specified role. Test performance will provide an indication of an individual's match to the job requirements. The functional approach therefore focuses on a justification of the test rather than the justification for its use. Conversely, a trait based test may be used to identify areas of strength or weakness in a specific task in order to determine training needs. This will require several different measures in order to provide an overall indication of performance, in this way each subtest item will present a profile for each participant. Of relevance to this programme of research are the observations of Tavakol, Mohagheghi and Dennick (2008, p. 77) who identified that as skills become more complex, so the challenge of assessment also increases. In addition, Harris, Snell, Talbot and Harden (2010, p. 647) indicated that assessment will have clinical authenticity when it relates to the performance of clinical tasks. So tests employed to determine a learners 3-D spatial visualisation skill in an educational setting would need to align as closely as possible to the clinical tasks requiring those skills.

The critical evaluation of the spatial visualisation skill measurement literature identified the use of 17 test instruments across a range of undergraduate and postgraduate settings. Therefore the next step for this programme of research was to select and develop appropriate test instruments for the measurement of the spatial visualisation skill of radiotherapy learners. Setting the above observations within the context of planning and delivering 3 and 4–D radiotherapy, clinical and academic educators would benefit from having an understanding of their learners` spatial visualisation skills so that they can identify those who may be at risk of not being able to visualise and therefore understand these 3 and 4-D concepts. Test instruments should therefore be designed to replicate as closely as possible those mental spatial visualisation skills which are related to patient positioning and imagining proposed beam directions. But, as Rust and Golombok (2009, p. 37) point out, it is also important to recognise that the transfer of a test developed for one purpose to another and its ability to achieve the same differentiation of performance cannot be assumed. From the findings of the critical evaluation of the spatial visualisation measurement literature it was evident that the most widely employed test with reported reliability and validity is the Vandenberg and Kuse MRT. While it can be applied on its own, chapter 3.3, (p. 76) identified that spatial visualisation skill is not a unitary construct, but has three subcomponents so the use of a single test may not provide information about the other domains. Therefore, the use of more than one instrument is recommended. Also, the employment of 2-D paper folding tests, while providing an indication of participant performance in mental transformation tasks, may not be indicative of performance in the more complex 3-D transformations required in a complex clinical setting such as radiotherapy. The use of a complex 3-D rotation test, however, could test the relationships between several objects and replicate the transformation in position and appearance of a tumour target volume and the relational normal anatomical position required in the clinical setting. The Purdue Spatial Visualisation Test could determine rotational skill but only uses transformation of a single object, therefore the more complex block associations in the Vandenberg and Kuse MRT would be more contextually meaningful in the representation of the rotation of a tumour target volume and normal anatomy. For the measurement of the spatial visualisation and perceptual skill required to imagine the X-ray beam path and its relationship to the tumour target volume and normal anatomical position, a cutting planes test would be most closely aligned. While the traditional cutting planes test (chapter 3.8.3, figure 3.12, p. 98) may replicate the visualisation of a plane (the central axis of an X-ray beam for example) the cutting plane only intersects a single object. Therefore, in order to provide an indicator of the perception and spatial visualisation of the relationships between anatomical structures then a more complex combination of objects is proposed for this programme of research. Given these considerations, the test instrument of choice to measure mental rotation will be the Vandenberg and Kuse MRT and the SBST will be employed to determine perception and spatial visualisation skill. Figure 3.17 illustrates the relationships between a tumour target volume in a normal position and in a position which is rotated in the same orientations as MRT objects.

Figure 3.17: Comparison of tumour target volume rotations in the same planes as the Vandenberg & Kuse Mental rotation Test

The relationships between the three SBST viewing perspectives, cutting planes and linear accelerator gantry angles are shown in figure 3.18. The relationship between cutting planes and beam's eye view is demonstrated in figures 3.19 and 3.20 which show test object six and the expected view of a dose distribution in a quality assurance phantom. If an individual can translate their egocentric frame of reference then view perspective is shown below the correct answer choice, at the bottom of the figure. However, if an individual cannot change their view perspective and views the object as if they are looking directly at the object, as if looking at a PC monitor, they are likely to see and select the egocentric foil.

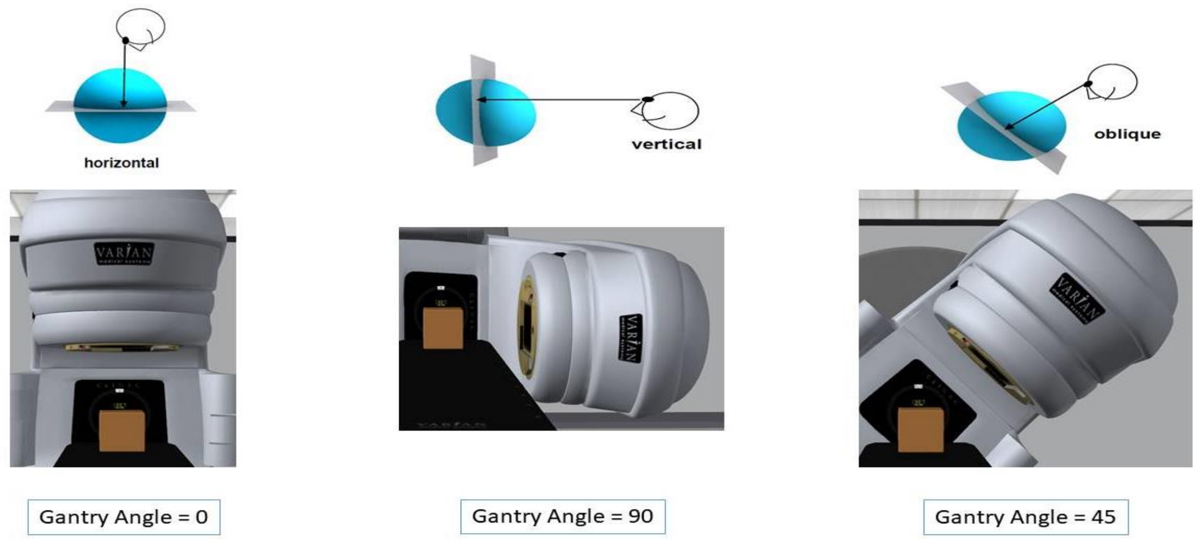


Figure 3.18: The SBST viewing perspectives, cutting planes and corresponding linear accelerator gantry angles

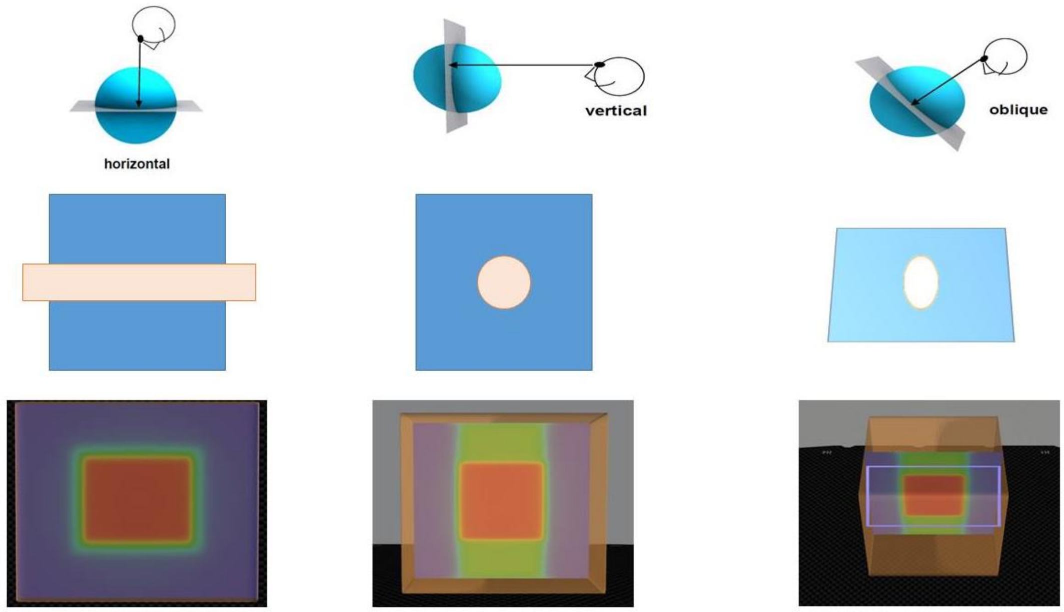


Figure 3.19: The SBST cutting planes with correct answer choices for object six and corresponding expected beams eye views

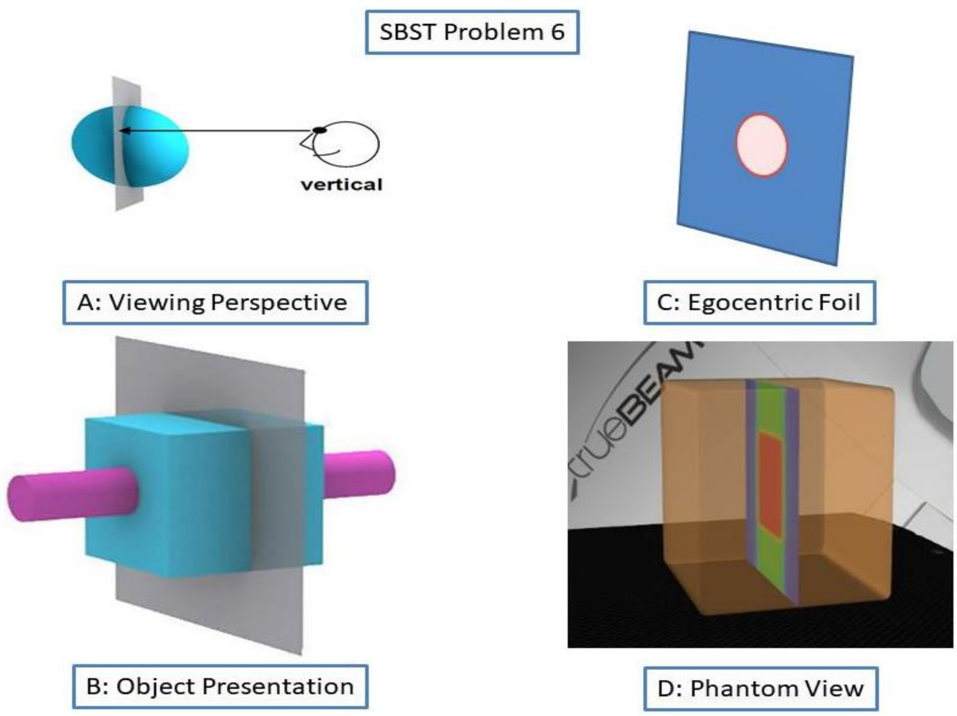


Figure 3.20: SBST object six showing the shape of the egocentric foil and associated beam's eye view

3.12 Research questions for this programme of research

The development of the research questions summarised below in table 3.6 were guided by the principles outlined by Agee (2009, p. 432). These suggest that the questions which arise during the early stages of research development and planning are likely to be driven by initial curiosity. They will exist in draft form initially and be based on provisional questions such as “what is going on here?” or “what processes are at play?” and will generate early thinking. The researchers’ interest in how VERT™ might support the learning of complex radiotherapy processes through visualisation, the concept of measuring baseline 3-D spatial visualisation skill and how it may change over time evolved and drove the formulation of questions for studies one and two. Research during the development phase of this programme of research indicated that spatial visualisation skill might be mediated in part by environmental factors such as the activities or games that children play and this underpinned the development of the question for study six.

Table 3.6: Research questions for the pilot and experimental study phases

Research Questions and Hypotheses
The Pilot Phase
Study 1 1. Can the SVS of Diagnostic Imaging & Radiotherapy students be measured using paper based & online test instruments? 2. Will the online test instrument produce equivalent (non-inferior results) when compared with the paper based test? *H^0: There will be no significant difference between performance scores with each platform*
Study 2 1. Does spatial visualisation change over time? *H_0: No change will be identified between measurement time points* 2. If it does, can it be detected by paper and online test platforms?
Study 3 1. Does the acceptability (suitability) and utility (usefulness and fitness for purpose) of the online platform compare with the paper based test?
The Experimental Phase
Study 4 1. To what extent can the spatial visualisation skill of pre-registration radiotherapy students be measured? 2. Does spatial visualisation skill change during the programme of study? *H_0: No change will be identified between measurement time points*
Study 5 1. To what extent does baseline visualisation skill have an impact on the performance of a complex positioning task using the 3-D virtual environment for radiotherapy training (VERT™) platform? *H_0: Baseline visualisation will not have a significant impact on task performance*
Study 6 1. What factors may affect the development of spatial visualisation skill?

Within the context of the emerging evidence base for the role of the VERT™ platform, the overarching aims and objectives for this programme of research were summarised in table 1.3, chapter 1.11, p. 21. Following the identification and selection of validated 3-D spatial visualisation test instruments, the research questions and their associated hypotheses, shown in table 3.6 were formulated in order to fulfil the aims and objectives.

3.13 Chapter summary

The first part of the chapter provided a technical narrative that examined current radiotherapy practice. The second part opened with an exploration of the evolution of the theories supporting spatial visualisation skill and its development. This was followed by a discussion relating the definitions and components of spatial visualisation, how they may be applied to radiotherapy and how they may be measured. In order to fulfil the research questions detailed above it was important to identify, via a critical evaluation of the literature reporting spatial visualisation skill measurement, appropriate tests that could measure the components of visualisation pertinent to the radiotherapy process. Following the identification of the Vandenberg and Kuse MRT and the SBST, a justification for their selection was provided through an examination of their alignment with radiotherapy patient positioning and beam's eye view visualisation activities. The chapter concluded with a statement of the research questions which were developed and formulated for this programme of research.

Chapter 4

Research design and methodology

4.1 Introduction to chapter 4

Chapters one and two identified the role of advanced radiotherapy treatment modalities in the management of cancer and the importance of 3-D spatial visualisation skill in supporting the accuracy of delivery of these techniques. Chapter three provided definitions for the components of spatial visualisation skill and their applicability to the radiotherapy process were derived from the general spatial visualisation literature. This was followed by the identification of validated tools for the measurement of these spatial visualisation components. Finally, the research questions relating to the measurement and development of 3-D spatial visualisation skill in radiotherapy pre-registration learners have been identified.

Chapter four will provide an overview and justification for the research approach, design and methods employed for the studies undertaken during this programme of research. The chapter will begin by examining the philosophical underpinning of this programme of research via a discussion of research worldviews and paradigms. It will continue by discussing the evolution of the nature of radiotherapy knowledge from its epistemological context. Following an examination of research design and methodology the key ethical considerations for the research will be considered. The chapter will continue with a personal reflection relating to the position of the researcher within the context of the programme of research before concluding with a summary of the research phases and the methodologies employed for each of the six studies undertaken.

4.2 Philosophical underpinnings of this programme of research

4.2.1 Introduction to research worldviews, approaches and methods

The term worldview has been used by Creswell (2009, p. 6) to mean a set of beliefs that may guide actions. Others have referred to these beliefs as paradigms, with Giuliano, Tyer-Viola and Lopez (2005, p. 244) referring to them as the researcher's general philosophical orientation. Research paradigms have been defined by Doyle, Brady and Byrne (2009, p.1 76) as worldviews that are determined by the elements of epistemology (how we gain the knowledge of what we know) and ontology (the nature of reality

4.2.2 The positivist / post positivist worldview

The positivist worldview is based on assumptions that tend to represent the traditional view of research and is referred to as the scientific method (Creswell, 2009, p. 6) It has been identified by Scotland (2012, p. 10) as making predictions and generalisations about the nature of knowledge. The epistemology is one of objects having an existence which is independent of the researcher while the ontological position is one of realism and positivism. Phenomena will therefore have an independent existence which can be discovered by the researcher's chosen methods. These methods will generate quantitative empirical data derived from, for example, standardised tests and the answers to closed questions. Researchers will therefore seek to make predictions and generalisations about the nature of knowledge by identifying the causal relationships between variables. For the studies planned for this programme of research, this would involve the measurement of 3-D spatial visualisation skill, to determine if any previous activities of a spatial nature would have any bearing on test performance and to examine any relationships between test score and performance of a complex radiotherapy task.

According to Creswell and Plano Clark (2011 p. 40), the post-positivism worldview is associated with quantitative approaches based on cause and effect and the reduction of the data based on the narrowing of the focus on the interrelationships between select and distinct variables. The post -positivist tendency is therefore to prove or disprove a hypothesis by working from the theory to the findings in a top down approach to add to or disprove the hypothesis. In this way the philosophical assumptions and the theoretical paradigms are crucial to gaining understanding. Therefore the positivist and post-positivist approaches view the social reality of research as being external to the individual. Because objects have an independent existence, the basis of knowledge is objective and places the researcher in an observer role as identified by Cohen, Manion & Morrison (2007, p. 7). The objective position therefore proposes that scientific investigation is predominantly quantitative and the analysis of data is related to theory testing.

4.2.3 The constructivist worldview

The focus of the constructivist worldview is that the understanding of social reality is gained as a result of individual perception and is mediated by the senses of the individual, i.e. internal to the individual. Reality is constructed by that individual's awareness so that knowledge becomes subjective and unique to that individual and will differ from researcher to researcher (Creswell, 2009, p. 8). Meaning is developed through interaction with others and the interpretation of findings is based on an understanding of context and culture. This places the researcher in a participant role and the research approach can be aligned with both quantitative and qualitative methods. The methods employed will therefore yield insight into and understanding of behaviours from the participants' perspective using, for example case studies, phenomenography and ethnographic methods. The contrasting and traditional views of positivism and constructivism have tended to polarise between quantitative and qualitative approaches throughout the 20th century (Johnson, Onwuegbuzie & Turner, 2007, p. 117). As Onwuegbuzie and Leech (2007, p. 240) observed, quantitative researchers aim to generate statistical evidence from which generalisations to a wider population may be made. Qualitative researchers, however, may not wish to make wider generalisations because the research goal may be to gain insight into a particular social context. The qualitative researcher will purposefully select individuals or groups in order to increase understanding in interpretivist studies. But, as Scotland (2012, p. 11) has suggested, the ontological position and the methodology of the interpretivist paradigm is directed at understanding phenomena from an individual perspective through the investigation of interactions.

4.2.4 The transformative worldview

The positivist / post positivist and constructivist worldviews seek to impose structural theories and laws which may not fit with all individuals in society. The transformative worldview, therefore, tends to speak to issues of social justice, discrimination of marginalised groups and empowerment for change (Creswell, 2009, p. 9). Inquiry is therefore linked to an action agenda for reform and tends to focus on inequalities within the study group from which an understanding of needs can be developed.

4.2.5 The pragmatic worldview

The worldview of the pragmatist is similar to that of the transformative view and arises from a concern for actions, consequences and solutions. But as Doyle et al., (2009, p. 179) reported, the perspective is informed by the notion that the practicalities of the research cannot be driven purely by theory or data exclusively. Therefore, it supports a process of abduction in which the researcher can move backwards and forwards between deduction and induction. The authors go on to suggest that when deciding on a methodology, the first consideration should be to ascertain which approach would best suit the research questions asked. This aligns with the philosophy of Johnson and Onwuegbuzie (2004, p. 21) who proposed that researchers should be encouraged to determine which method will work best in terms of answering the questions. This allows the researcher to adopt needs based or contingency based approaches in which the consequences of the research are more important than the process and lead to the end justifying the means. A summary of the key characteristics of these four worldviews is presented in Table 4. below. The historical approach to educational research in general has been located firmly in the positivist quantitative paradigm which contended that there was a single reality from which causal relationships could be identified (Doyle et al., 2009, p. 177). As already identified in section 4.2.2, p. 130, the researcher was considered to be independent and objective, conferring an outsider status in relation to the study participants. As constructivism emerged, a more qualitative form of inquiry developed, as researchers sought to examine the contextual issues of care and illuminate reality through detailed description of patient experience.

Table 4.1: Four worldview approaches to research (adapted from Creswell, 2009, p. 6)

Positivist / Post Positivist	**Constructivist**
Empirical observation and measurement Verification of theory Determination Quantitative	Socially constructed Theory generation Multiple participant meanings
Transformative (Participatory)	**Pragmatism**
Collaborative Change and power oriented	Problem centred Real world oriented

4.3 The epistemology and ontology of the approach chosen for this programme of research

During the development and planning phases of this programme of research it was important to position the context for the research and the nature of the aims, objectives and research questions within the epistemological and ontological foundations of the evolution of radiotherapy knowledge. Therefore the following section will examine this evolution from predominantly 2-D processes through to the application of 4-D adaptive techniques in current clinical practice and the impact that this has had on the scope of practice and the role of radiotherapy radiographers.

4.3.1 The nature of radiotherapy knowledge

The nature of radiotherapy knowledge and the basis on which knowledge claims are founded can be considered to be similar to those in nursing. These have been identified by Giuliano et al., (2005, p. 243) as being grounded in the understanding of human relationships and the health care environment. They point out that understanding the impact of education on the concepts of the prescribing of care and the prediction of consequences are not mutually exclusive and lead to multiple ways of knowing. From an historical perspective, there is the received view that contends that there is a body of facts and principles to discover independently of their social context. This results in a disconnection from any interactive process. The alternative is the perceived view, based on the belief that facts and principles are embedded within a particular history, that truth is dynamic and bound to a person, place and time and the interactions of individuals are related to their socio-historical context.

While the nature of care in nursing and radiotherapy is similar, the nature of knowledge in radiotherapy is also predicated by the evolution of specific scientific understanding relating to the differential response of tumours and normal healthy tissue to radiation dose. The received view of radiobiology prevalent in the development of dose and fractionation regimes during the early part of the 20th Century focussed on restriction of and interruption to growth patterns for the tumour (Delwiche, 2013, p. 121). For healthy tissue response to radiation dose, the perceived view relates to understanding normal tissue complication probability and the risk and impact of adverse long-term sequelae on patients’ lives post radiotherapy. As the techniques for the planning and delivery of radiotherapy have developed through the transition from 2-D to 3-D techniques it has become possible to escalate the dose delivered to the tumour target volume while limiting the dose to the surrounding normal tissue. The evidence base relating to late side effects has developed alongside the technology to visualise the shape and position of the tumour target volume in 3-D. Taking into account the changes in patient and tumour position which may take place over the duration of a treatment delivery course through 4-D adaptive processes has further reduced the risk of long-term effects. As the spheres of practice knowledge in radiotherapy are considered to be the biology of the tumour, technology and the patient, Gillan and Liszewski (2016, p. 5) state that it is incumbent on all radiotherapy radiographers to keep abreast of these practice trends and developments in technology to ensure that the evolving needs of the patient are met. Current radiotherapy practice is reliant on technology and more complex field arrangements with tighter margins between the tumour target volume and normal tissues have resulted in efforts to automate tasks that could be prone to human error. As this complexity grows, so does the need to consider the issues of quality and safety of procedures leading Robson, Clarke and White (2014, p. 129) to refer to the designing out of errors via human factors engineering and automation. As task automation has developed over the past decade, it has eliminated the need for certain skills and competencies which were once an integral part of daily practice. One example identified by Gillan and Liszewski (2016, p. 5) is that of the hands on positioning of the patient, based on the external coordinates on their skin surface, at the isocentre of the linear accelerator. This increased dependence on technology, however, does not override or eliminate responsibility for the task, rather the responsibility shifts from task performance to an appreciation of the degree of automation and ensuring that all the required tasks are completed according to the desired outcome. But it is also important to recognise that automated systems can fail and without experience it is difficult to recognise that an error has been made (Probst, Hutton, Collins & Adams, 2014, p.249). This is most likely to be the case if the system has been seen to be safe and reliable in the past and this over reliance on previous reliability has the potential to impair decision making expertise since radiotherapy radiographers no longer need to employ these tasks on a regular basis. This places a caretaker responsibility on radiotherapy radiographers to ensure that they are able to interpret what the technology is telling them based on their understanding and application of fundamental principles.

A further example of technology development has been seen in IGRT (Gillan and Liszewski (2016, p. 5), images can now be acquired digitally and viewed online in a fraction of the time required for previous film based procedures. Reliance on orthogonal planar imaging in 2-D has been replaced with the ability to complete a cone beam CT scan in a single rotation of the linear accelerator gantry. It is the responsibility of the practitioner to ensure that their decision making is based on fundamental knowledge, skills and best practice in order to interpret what the technology is telling them. Central to this is the gaining of skills in cross sectional image interpretation and an associated need for radiotherapy radiographers to develop their 3-D spatial visualisation skills. However, while the evidence base relating to the measurement and development of spatial visualisation skills in other science, technology and engineering learners has been widely reported, it remains relatively under researched in radiotherapy.

4.4 Methods of enquiry in quantitative research

The tenet of quantitative methods of enquiry assumes that behaviour can be explained by measurable facts which are investigated by deductive logic. The process requires the development of testable hypotheses and theories which may be generalised across different settings. The results are used to explain phenomena, as in this research, the spatial visualisation skill of individual learners, based on the statistical analysis of numerical data for test performance score. Quantitative research can therefore be described as an empirical, scientific method (Cohen et al., 2007, p. 15) designed to test a theory consisting of variables that can be measured with numbers and analysed with statistical methods (Yilmaz, 2013, p. 311). It is therefore deductive in nature and informed by objectivist epistemology which seeks to develop explanatory laws in social behaviour by emphasising the measurement and analysis of causal relationships. Methods of inquiry can include surveys to provide a numerical description of opinions, attitudes and trends within a population via cross sectional or longitudinal studies. Another method involves the use of experimental research to determine the outcome of an intervention. The random assignment of subjects to an intervention group is referred to as a true experiment that can utilise a control group (no intervention) and the experimental (intervention) group. The non-randomised design is referred to as a quasi-experiment (Creswell, 2009, p. 12).

4.5 Methods of enquiry in qualitative research

Research employing qualitative methods have been referred to by Yilmaz (2013, p. 313) as being any type of approach that will produce results that are not derived from statistical procedures. Founded in the epistemology of constructivism, it assumes that social phenomena are so complex and interwoven that they cannot be reduced to distinct variables. Unlike quantitative methods which do not take account of, or provide insight into, participants' individual experiences, qualitative methods are concerned with context and gaining understanding through emergent, inductive and interpretive reasoning. Qualitative research can include a range of approaches as outlined below together with examples from the field of health care in general and radiography.

4.5.1 Grounded theory

Grounded theory involves the researcher deriving a general theory relating to a process or interaction which is grounded in the views of study participants (Creswell, 2009, p. 13). Typical examples are the study of practitioner effectiveness in the management of clinical cases and practitioner – patient interactions.

4.5.2 Phenomenology

Phenomenology has its origins in the work of Brentano and Husserl during the latter 19th and early 20th Centuries and is identified by an attempt to describe the basic structure of human experience from a first person point of view, as described by Merleau-Ponty (2013 p. viii). It focuses on how objects or events appear in the consciousness of an individual. A number of strategies that a phenomenological researcher may use in the investigation of lived experience have been identified by Randles (2012, p. 2). These may include interviews, conversations and discussions, observations and focus group meetings but generally starts with a reflection of the researchers' personal experience to illuminate one's own thought processes (refer to section 4.14, pp. 146 -148 for a reflection on this researchers experience as a child and as a clinical radiotherapy radiographer. Analysis therefore, aims to explore the relationships between objects, acts and meanings. During this programme of research, a survey tool (based on Terlecki`s Survey of Spatial Activities, section reference) was employed to explore participants lived experience of participation in activities of a spatial nature (this needs linking to references for the literature reporting potential impact of these activities in the development of spatial visualisation skill.

4.5.3 Ethnography

Ethnography as a field of inquiry is related to the way of life or perceived identity of a specific culture. It has been described by Carthey (2003, p 13) as the observation and systematic recording of human culture. In relation to healthcare the collection of both qualitative and quantitative data related to the type, frequency and severity of adverse events relating to drug administration and departmental process is known as structured observational research. Examples include the perceptions of the concept of caring held by radiotherapy radiographers and the experiences of minority patient groups or cancer patients.

4.6 Comparison between quantitative and qualitative methods

The proponents of quantitative and qualitative research have tended to focus on the difference between the two philosophies rather than the similarities. However, as Onwuegbuzie and Leech (2005, p. 376) have suggested, there needs to be an appreciation of both methods. They refer to a continuum of perspectives ranging from a purist view that quantitative and qualitative methodological approaches cannot be mixed because they stem from different epistemological and ontological assumptions about the nature of research. They also contend that if purists are positioned at one end of the continuum, at the other end are the pragmatists who believe that the consequences of the research are more important than the process. So when considering the most appropriate research approach, the driver should be the research questions themselves. This is a view supported by Doyle et al., (2009, p. 176) who suggested that research should not be restricted by traditional methods; rather it should be guided by a foundation of inquiry which underlies the research activity and not locating it within a specific paradigm. This is the position held by the situationists who hold the single method stance of the purists but also acknowledge that certain research questions lend themselves more to quantitative rather than qualitative methods (Onwuegbuzie and Leech, 2005, p. 376).

4.7 Mixed methods

Given the above, at the core of the mixed methods approach, is an effort to integrate the complementary strengths of quantitative and qualitative methods and to provide triangulation of the findings (Morgan 1998, p. 363). The approach has been described by Symonds and Gorard (2010, p. 121) as encouraging the integration of both quantitative and qualitative methods which can encourage alternative and independent thinking and sits outside the traditional standpoints identified above. This is a view supported by Creswell (2014, p. 13) who indicated that the three approaches should not be viewed as distinct and rigid categories but framed in terms of the chosen research methods.

4.8 The planning and design justification for this programme of research

When deciding on the methodology for this programme of research, the first consideration was to ascertain which approach would best suit the research questions posed, as advocated by Doyle et al., (2009, p. 78) and Bunniss and Kelly (2010, p. 359). This aligns with the philosophy of pragmatism whereby researchers are encouraged to determine which method will work best in terms of answering the questions. This allows the researcher to adopt a needs based or contingency based approach in which the consequences of the research are more important than the process and lead to the end justifying the means. To support these decisions the principles outlined by Agee (2009, p. 432) were considered in order to support the development and planning process. Good questions may arise from initial curiosity but are likely to exist as drafts in the early stages and will change during the research process as a reflection of an increased understanding of the problem or issue. Basic questions such as “what is going on here?” and “what processes are at play?” can form the early provisional and generative questions. While the early iterations may already be influencing and determining the decisions relating to research approach and methods the questions posed will generate early thinking which will inform the approach and method.

The stratification or grouping of pre-registration learners in radiotherapy by their spatial visualisation skill would require the use of measurement tools and the analysis of performance scores, thereby positioning the research approach in the quantitative domain. However, as reported in chapter 3.4, p. 79, the development of spatial visualisation skill may be influenced by environmental factors. To gain an understanding of and an insight into how these factors may have an impact on baseline visualisation skills would require the use of qualitative survey tools. The combination of these two complementary requirements therefore positions this programme of research in the domain of mixed methods.

4.9 Priority decision and timing

In order to combine quantitative and qualitative methods, two basic decisions need to be taken. These have been identified by Morgan (1998, p. 364) as the priority decision which will determine the extent to which the quantitative or qualitative method will be the principal tool and the sequence decision which will determine the order in which each method will be used to collect data. The relative weight given to these decisions leads to a framework of four combinations in a two by two matrix as shown in table 4.2. While this framework may serve as a starting point when considering the relative contributions provided by quantitative and qualitative data Creswell (2009, p. 206) also points out that in addition to the sequence priority consideration also needs to be given to the timing, weight and mixing of the quantitative and qualitative components of the study. One of the most frequently used is that of the sequential-explanatory design as reported by Ivankova, Creswell and Stick (2006, p. 5). This method commences with the collection of quantitative data and analysis which is followed up with the collection and analysis of qualitative data as described by Creswell and Plano-Clark (2011, p. 71). The qualitative results thereby assist with the explanation of the initial quantitative findings.

Table 4.2: The priority – sequence model of complementary quantitative & qualitative research methods showing the relative weighting of quantitative & qualitative approaches

		Priority Decision	
		Principal Method Quantitative	**Principal Method Qualitative**
Sequence Decision	**Complementary Method Preliminary**	Qualitative Preliminary = qual - QUANT	Quantitative Preliminary = quant + QUAL
	Complementary Method Follow Up	Qualitative Follow Up = QUANT - qual	Quantitative Follow Up = QUAL - quant

Creswell (2009, p. 210) also goes on to indicate that a quant - qual notation would signify that qualitative methods are embedded within a quantitative design but each has equal weight and contribution. This is referred to as a concurrent embedded design and is characterised by collecting quantitative and qualitative data simultaneously. The primary driver in both phases of this programme of research was to determine if 3-D spatial visualisation skill could be measured using a quantitative measurement tool and to understand if any experiential and biological (qualitative) factors could be demonstrated to have an impact on spatial visualisation skill test performance. While data from the two methods was collected simultaneously to provide an overall composite analysis, they were kept as independent entities during the analysis stage and mixed during the interpretation stage. This has been referred to as a convergent parallel design by Creswell and Plano-Clark (2009, p. 69).

4.10 Ethical considerations

While researchers in health and social science education may generate knowledge through their research, they should also acknowledge the position and power relationships that they have with their students as study participants (Karnieli-Miller, Strier & Pessach, 2009, p. 282). There is also a requirement to prevent participant harm and uphold the standards of ethics and confidentiality. The risk of breaking confidentiality increases in what Damianakis and Woodford (2012, p. 708) refer to as small connected communities. This is an important consideration in radiotherapy given the relatively small number of registrants and until 2018 the limits placed on commissioned training places for each Higher Education Institute by Health Education England.

Therefore, the main ethical considerations and concerns identified for this programme of research were:

1. Using students as participants in research;
2. Ensuring that participants understood the purpose and requirements of the research studies;
3. Ensuring that participants had provided informed consent;
4. Ensuring confidentiality and the security of written and electronic data;
5. Researcher positionality and conflict of role.

4.10.1 Students as participants in research

Students as participants in research are protected by the Declaration of Helsinki (Bradbury-Jones & Alcock, 2010, p. 192) and the Nuremberg code (Burgess, 2005, p. 59) which refers to those individuals who may be dependent or vulnerable. In educational research the potential exists for learners to be vulnerable if the researcher can simultaneously exercise power as a teacher or examiner (Kanter 2009, p. 149; Ten Cate, 2009, p. 608). To formally safeguard participants in this programme of research, favourable ethical opinion was given by the University of Portsmouth Science Faculty Ethics Committee (refer to appendix 4 for the pilot study documentation and appendix 5 for the experimental study documentation).

4.10.2 Participant understanding of the purpose and requirements of the research

It was also important that potential participants were fully conversant with the purpose of and the requirements for the planned research studies prior to consenting to take part. All first year (FHEQ level 4) students registered on the Diagnostic and Therapeutic Radiography programmes at the University of Portsmouth were invited, via a student portal email, to attend briefing sessions prior to the commencement of both study phases. A short verbal introduction by the researcher outlined the purpose of the studies, data collection methods, participant time commitment and requirements. At the end of these sessions information sheets and consent forms were distributed by the student consultative committee representatives. In recognition and acknowledgment of the researcher / student power relationship identified above, there was no direct recruitment of students by the researcher.

4.10.3 Informed consent

The principles of informed consent lie in the participant right to freedom of choice and self-determination after being informed of the facts likely to influence their decision on whether to take part in a research study or not (Cohen, et al., 2007, p. 52). This must include identification of the benefits and risks of taking part and the right to withdraw at any time without the need for explanation or penalty. This ability of participants to opt in and opt out is likely to increase participant autonomy and feelings of control and relative power. Prospective participants returned signed consent forms to a designated drop box in the radiography academic team office or were brought to the data collection sessions and collected by the researcher.

4.10.4 Confidentiality and data security

To ensure that confidentiality was maintained throughout the programme of research, paper copies of answer booklets and consent forms were stored in a locked filing cabinet within the Radiography academic office. Data from online tests and questionnaires were exported in Microsoft Excel spreadsheet format and stored on the researchers password protected storage space on the University intranet, with a back-up copy stored on a dedicated encrypted memory stick and stored in the same cabinet. Anonymity was maintained by giving individual participants a unique numerical identifier. For paper copies of answer booklets these were numbered sequentially from one, based on the order that they were collected in. Student registration numbers were used solely for the purpose of collation of subsequent test results. Where results were automatically downloaded to Microsoft Excel, participants were identified by sequentially numbering the spreadsheet. A log of participant names and unique identifying numbers was kept as a separate document to facilitate the collation of subsequent test results. The collation and coding of participants in subsequent rounds of testing was completed before data analysis of performance score was carried out to ensure that the researcher was blinded to previous test performance.

4.11 Ensuring reliability and validity in quantitative research

Reliability and validity are the criteria used to determine the quality of quantitative measures. Reliability refers to the reproducibility of a test and the consistency of the results gained during follow up testing (Polgar & Thomas, 2013, p.107), while Walker and Almond (2010, p. 86) identify it as the ability of a test to consistently measure an attribute in practice. Validity, however, centres on ensuring that a test instrument is accurately measuring what it purports to measure.

4.11.1 Reliability

Three components of reliability have been described by Lodico, Spaulding and Voegtle, (2010, pp.95-96). The purpose of the first, referred to as stability or test-retest consistency over time, is to demonstrate that an individual can obtain the same score on a test instrument if the test is taken more than once. The second is referred to as equivalent form reliability (also known as alternative form consistency) and relates to the consistency of performance across different formats of the test instrument. It seeks to identify if asking different questions, while still assessing the same content, knowledge and skills will produce the same mean results and standard deviations. The final component is internal consistency, which is a measure of the consistency of each item within a given test instrument to test the same trait or ability.

4.11.2 Validity

The broadest definition of validity has been provided by DeVon et al., (2007, p.155) who identify it as the ability of a test instrument to measure the attributes of the construct being examined. This would suggest that the validity of an instrument is measured by the production of meaningful results. In the case of this programme of research, this would relate to the measurement of 3-D spatial visualisation skill using previously validated mental rotation and cutting plane test objects. However, in an examination of the threats to experimental validity, Lodico, Spaulding and Voegtle, (2010, p.98) identified three components of validity. The first is content validity and relates to the critical components of individual test objects and the relationships between them. For this programme of research, individual test item analysis for missing items and the pattern of incorrect answer choices from the SBST subcomponents might determine the relationships between complete and incomplete questions and correct and incorrect answers.

The second is that of criterion related validity, which examines the relationships between two tests which are taken at the same time and the degree to which performance in each is correlated. Lodico, Spaulding and Voegtle, (2010, p.99) identified and reported two forms; concurrent and predictive validity. Concurrent validity examines the degree to which performance in one test one is related to another. For this programme of research concurrent validity relates to the ability of the Vandenberg and Kuse MRT to measure an individual's skill in mentally transforming complex objects and the SBST to measure perception, visualisation and visual penetrative ability. The tests should produce results that are consistent with previous studies. However Cook and Beckman (2006, p.10) argue that these distinctions may be arbitrary and should be contained within an overarching concept of construct validity which incorporates the test content, the thought processes of the study participants (for example, does the test invoke higher order thinking). They go on to indicate that consideration should also be given to the internal structure of the test (or reliability), the consequences of the assessment (the degree to which the desired results have been achieved, the method used to determine score thresholds) and the manner in which the evidence relates back to the original theoretical construct. The second form, predictive validity, is the ability of a test to accurately predict outcomes or performance at a later date. In relation to the question posed for study five in this programme of research, if baseline 3-D spatial visualisation is a predictor of future performance in complex radiotherapy tasks, then there should be a correlation between test performance score and task performance. The third component is construct validity which involves the finding of evidence to determine that a test instrument is able to accurately measure the trait or skill being investigated.

Validity can also be considered as being internal and external. Internal validity relates to the extent to which differences in a dependent variable (performance scores in 3-D spatial visualisation test instruments) is due to the experimental interaction (mental visualisation activities undertaken during clinical placement experience and simulated environment practical activities) rather than other extraneous factors. So, at the conclusion of this programme of research, if there are positive changes in 3-D spatial visualisation test performance score, can they be attributed to the development of 3-D spatial visualisation skills? External validity relates to the degree to which the results may be generalizable to other groups beyond the study sample. While this programme of research was designed to study 3-D spatial visualisation skills of diagnostic imaging and radiotherapy students, its findings could have relevance for other medical and health care educators whose students are required to have similar skills.

4.11.3 Relationships between reliability and validity

While a valid test can be considered to be reliable, Lodico, Spaulding and Voegtle, (2010, p.101) observed that a reliable test may not be valid due to the inclusion of inappropriate individual test items. They note that in order for a standardised test instrument to preserve its reliability and validity, it should be administered and scored according to the original test instructions. They also suggest that failure to do so is likely to compromise the quality of the instrument. In addition, Cook and Beckman (2006, p.7), referred to the degree to which interpretations and conclusions drawn from any psychometric assessment are justifiable. This would imply that validity is a property of the inferences drawn from the results rather than relating to the instrument of measurement itself. One further theme that has been identified for consideration when examining validity is that of utility (Keszei, Novak & Streiner, 2010, p.322) and relates to the practicalities of the test application. They suggest that even if an instrument has been demonstrated to be both valid and reliable, it may be impractical to administer due to the training required for the administrators, the time required for its completion or the time taken for marking. They also report that longer tests, in terms of the number of individual test objects employed, tend to demonstrate stronger validity than shorter ones with fewer objects. However in the interests of greater utility, a shorter test time with fewer test objects may be more advantageous in the practical setting. When considering the time permitted for completion of the spatial visualisation tests in this programme of research it was important to remain mindful of the fact that an average radiotherapy treatment delivery time is between 12 and 13 minutes and therefore the decisions relating to patient and tumour target volume position have to be made in a time pressured environment.

When Vandenberg and Kuse (1978, pp. 601 – 602) described the development of the MRT, they reported Pearson Product – Moment correlations with the card rotation test of .62 and for Shepard & Metzler identical blocks .68. For other spatial tests, such as hidden figures and form boards these values are lower at .4 and .41 respectively. . When Cohen and Hegarty (2007, p. 181) reported on the development of the SBST, in a study of 59 psychology students, they also employed the MRT and the Visualization of Views Test. They reported that the performance in both tests was highly correlated ($r = .47$) and using averaged score from both tests, which they referred to as the spatial score, reported a correlation of .5 ($p < .01$) with all types of test figures and cutting planes in the SBST. They also reported a split half Cronbach Alpha for internal consistency for the 29 test items in the SBST of 0.86, which they referred to as a satisfactory.

4.12 Ensuring reliability and validity in qualitative research

While quantitative research employs a deductive approach using standardised instruments which do not take account of individual participant experience, qualitative studies are concerned with the interpretation of data from inductive reasoning based on the meanings that individuals attach to their experiences of the world. The criteria of reliability and internal and external validity for testing rigour in the quantitative domain are well known, but they may not be appropriate or meaningful for naturalistic studies (Lincoln & Gaba, 1986, p. 74). They proposed the alternative and overarching principle of trustworthiness (as a parallel to rigour in quantitative research), within which they identified three subcomponents as analogues for reliability and internal and external validity (ibid, pp. 76-77). They identified these as dependability (reliability), credibility (internal validity) and transferability (external validity).

Dependability has been defined by Lodico et al., (2010, p. 172) as the tracking of the processes and procedures employed to collect and interpret data and go on to suggest that it can be increased by detailed explanations of how the relationships between the researcher and study participants was managed (refer to section 4.14.2, p. 149). Credibility was identified as how well the researcher has evidenced the reality of the situations that have been studied (ibid, p. 169) and can be thought of as the fit between respondents' views and the researcher's representation of them. Transferability relates to how lessons learned from one setting may be useful to others in different settings. It does not necessarily relate to a representative sample, but identifies how other researchers may determine whether similar processes will be at work in their own communities (Lodico et al., 2010, p. 173). While the predominant methodology planned for this programme of research was quantitative, qualitative elements were envisaged for both the pilot and experimental phases (refer to section 4.8, p .137).

4.13 Triangulation

When multiple methods or sources of data are used in research, triangulation is a methodological approach that adds to the validity of results and can lead to a multi-dimensional understanding of complex issues (Farmer, Robinson, Elliott & Eyles, 2006, p. 378). The technique has its origins in the determination of the position of a point which is based on observations from two additional points. Through an examination of these multiple dimensions, complementary convergence or dissonance between data sources can be exposed. However, as Farmer et al., (2006, p. 377) also reported, there appears to be little direction in the social sciences literature relating to the nature of this analytical process. Based on the work by Denzin in the 1970`s, they proposed four techniques for triangulation. The first is methodological triangulation which involves the use of more than one data collection technique. For the studies conducted during this programme of research, this involved exploring the possible links between a qualitative assessment of biological and environmental factors (study six) and quantitative performance scores (study four). The second uses data from multiple sources, this involved, the comparison of study findings with previously published work of similar design. Theoretical triangulation examines a phenomenon from multiple perspectives; this approach is similar to methodological triangulation. Investigator triangulation which entails the involvement of two or more investigators in data analysis, an approach was not pertinent to this programme of research.

Quantitative researchers attempt to triangulate using several measures of performance to provide explanations for their findings. Qualitative researchers will triangulate via focus group and interview data, phenomenography and their own view of reality to discover meaning. In quantitative methods data is reduced via item analysis while in qualitative methods data reduction occurs through a process of thematic analysis (Onwuegbuzie &

Leech, 2005, p. 379). This approach of thematic analysis was adopted for study three in the pilot phase to explore participant observations when comparing the paper based and online spatial visualisation test instruments.

In addition to these techniques, Cohen, Manion and Morrison add time triangulation, which considers the factor of change in longitudinal or cross sectional studies. While Farmer Robinson, Elliott and Eyles (2006, p. 379) refer to methodological triangulation, Cohen, Manion and Morrison (2007, p. 142) refer to combined triangulation which uses more than one level of analysis based on the individual and group being studied. It has also been identified by Cohen, Manion and Morrison (2007, p. 143) that the most frequently used techniques in educational research are time and space (based on the number of occasions a group is studied). Relating this to the studies planned for this programme of research, one of the time factors to be considered would be the analysis of the number of participants taking part in each stage and therefore the impact of the potential for missing data would need to be considered. This is explored further in section 4.16, p.152 below.

4.14 Researcher positionality, conflict of role and power relationships

The positivist (quantitative) paradigm or the naturalistic constructivist (qualitative) tradition to which they align may influence the worldview of the researcher. Over the last 20 years there has been an increasing focus on issues concerning researcher positionality, power and representation in research. While relating primarily to qualitative methods as identified by Merriam et.al, (2001, p. 406) it does have a resonance with the approach to all research. It was important therefore for the researcher to recognise, acknowledge and reflect on their life, identity and past experiences, which may influence the research trajectory. While primarily viewed as components of qualitative research, these issues could have an impact on the objectivity of the researcher as an observer in quantitative research. Positionality refers to the relationship that the researcher has with their study and their motives for collecting data as described by (Bourke, 2014, p. 7). Therefore, the lived experience of this researcher's development, education and career as a radiotherapy radiographer in both the clinical and academic settings may have influenced the approach to this programme of research.

Much of the researcher's early school life was viewed through a succession of occlusal patches on both eyes due to being born with a strabismus amblyopia of the right eye.

Having near normal sight in the left eye but a near-sighted right eye has led to a predominantly (but not exclusively) left-handed approach. This is illustrated by considering a range of sporting activities, a football will be kicked with the right foot, but in cricket, batting is right-handed and bowling is left-handed. As a child, when not building with Lego® bricks or other construction toys, time was spent sketching (copying rather than freehand) because perception of depth was not particularly well developed at that time. Training in radiotherapy took place in a predominantly 2-D era when the position of soft tissue organs was determined by their position in relation to surrogate bone anatomy and surface landmarks. The researchers' spatial visualisation skill was developed as a treatment planning radiographer over many years of viewing and interpreting 2-D X-ray images. As radiotherapy treatment planning and delivery moved into the 3-D era, cross-sectional visualisation skill was developed by comparing 3-D CT images with cross sectional images based on anatomical dissections compiled by medical illustrators. It is suggested that the researcher's ability to build and apply complex 3-D mental models has developed and evolved in pace with the evolution of advanced radiotherapy techniques. Having had the benefit of many years of clinical experience and time on task, the transition from a 2-D to a 3-D world was relatively straightforward. Today's learners in radiotherapy do not have that luxury of time. Radiotherapy radiographers work in a 3-D visualisation environment as standard, so there is a need for educators to understand the spatial visualisation profile of pre-registration learners and engender these skills in them from the start of their programme of education.

Given the researcher's clinical role in the planning and delivery of radiotherapy, much of this work has required the manipulation of mathematical radiation dose data to achieve clinically acceptable radiotherapy treatment delivery plans for each patient. These activities have also been supported by the researcher's participation in quality improvement and clinical audits which deliver answers to the closed questions of yes or no. This clinical role therefore employed a predominantly quantitative approach which stems primarily from the standardised methods for producing a treatment delivery plan and the interpretation of the results in terms of achieving maximum dose to the tumour target volume, whilst minimising doses to normal tissue. Latterly, the majority of the researcher's academic role has focussed on clinical preparation of FHEQ level four and five learners utilising VERT™ and examining the links between academic and clinical performance. The key motivation for embarking on this research journey was based on the researcher`s observations from radiotherapy departments and the academic setting that some learners had difficulty in following a practical demonstration and explanation of the concept of the linear accelerator isocentre and then being able to apply this to the accurate positioning of a patient. From a clinical perspective, the question of how clinical and academic educators could support these students in developing this insight arose.

While reflecting on these observations, the introduction of VERT™ has provided a platform for visualising the underpinning concepts of radiotherapy treatment planning and delivery. While the platform's contributions have been evaluated and reported in chapter 2.8, pp. 50 - 73), there is still a need to identify and understand which learners are likely to have difficulty in visualisation and potentially gain more benefit. The installation of VERT™ at the University of Portsmouth in 2008 initially led to consideration of its potential as an alternative assessment platform for threshold clinical skills. While searching the evidence base supporting the role of simulated and virtual reality environments, thoughts about how students learn in virtual environments began to evolve. This coincided with the publication of the Department of Health VERT™ Project report, which recommended the assessment of students' inherent spatial ability to assist identification of individuals who are likely to benefit most from experience in VERT™ (Appleyard & Coleman, 2010, p. 33).

4.14.1 Researcher power and representation

The ability of a researcher to reflect on their position however goes beyond an understanding of self and must also acknowledge the positioning of the study participants and the power of the researcher in the research process. The nature of the identity and complexity of research groups has been discussed by Oikonomidoy and Wiest (2015, p. 55) and highlights the cross-boundary connections between researchers and the group being studied. During this programme of research, the position of the researcher as a lecturer and tutor and the students as potential study participants had to be recognised and managed accordingly. This duality of identities has been referred to as the insider – outsider position. Management of this interface was achieved by having regular discussions with the researcher's supervision team and by adhering to the principles discussed below in section 4.14.2. It has also been suggested by Oikonomidoy and Wiest (2015, p. 55) that if insiders get too close to the study group then this can lead to an inability to recognise and interpret group characteristics. Furthermore, participants may not be willing to share information due to fears about confidentiality. Conversely there is also a risk that outsiders may not understand these group characteristics and neglect important aspects of participant views and behaviours.

Within this programme of research, the researcher as a postgraduate student has an outsider status in relation to the pre-registration students as study participants. But being a programme lecturer and clinical practice link tutor positioned within a small community of practice confers an insider status and has the potential to influence the power based dynamic over the duration of the research. As the research studies were conducted using participants recruited from the researchers host institution, for some students the researcher is a programme lecturer, while for others the researcher has a role as their clinical practice link tutor. In recognition of the power relationship between lecturer as researcher and students as participants in research conducted within a small connected community (Karnieli-Miller, Strier & Pessach 2009, p. 282), the choice of a predominantly quantitative research approach aimed to reduce the risks identified above.

4.14.2 Minimising researcher conflict of role

The researcher as an HCPC registered radiographer must abide by the standards of proficiency for radiographers (HCPC, 2013) specifically regarding standards 13.8, (p. 12) and 14.5, (p. 15) relating to the principles of scientific enquiry and research methods, standards 2.7 (p. 7) and 7.1 (p. 9) gaining informed consent and maintaining confidentiality and standard 15 (p. 18) which involves the need to establish and maintain a safe practice environment. These standards are also embodied within the Society and College of Radiographers Code of Professional Conduct (SCoR, 2013, p. 8) which requires radiographers to work within the current legal, ethical, professional and governance frameworks relating to their specific occupational role. There is also a requirement for researchers to reflect on their motives for collecting data (Bourke, 2014, p. 7). The key drivers for this programme of research were the researcher's observations, from both the radiotherapy clinical environment and the academic setting, that a proportion of learners had difficulty in following a practical demonstration and explanation of the concept of the linear accelerator isocentre as discussed above and then applying this to a real case. From a clinical perspective, the question of how clinical and academic educators could support these students in developing this insight arose. The introduction of VERT™ provided a potential platform for visualising the technical concepts of radiotherapy planning and treatment delivery and has been firmly embedded into pre-clinical preparation practical tutorials. That said the role of the platform in supporting learners who have difficulty in spatial visualisation remains under researched. This in turn led to several questions centred on the spatial awareness and visualisation skills of learners and how they might be developed. From these reflections the overall concept and framework for this programme of research was derived.

4.15 Sampling considerations and management of missing data

As reported by Collins, Onwuegbuzie & Jiao, (2007, p. 269) decisions relating to sampling will stem from the research goal which may aim to have a social or organisational impact, to make predictions, to measure change and to understand complex phenomena. Four broad categories of sampling strategy have been described for use in mixed methods research by Teddlie and Yu (2007, p. 78) with the selected approach being driven by the research questions being asked. These approaches are summarised in table 4.3:

Table 4.3: Summary of sampling methods and approaches

Sampling Method	Approach
Probability sampling	Will randomly draw many cases from a population or subgroup so that each member of that population has an equal chance of being selected
Purposive sampling	Selection of a group based on a particular dimension of interest in order to reflect the characteristics of the study population
Convenience sampling	Occurs within a captive audience whose members may not be totally representative of the specific characteristics being investigated and therefore any findings may not be transferable to a specific subgroup
Mixed method sampling	Employs a combination of probability sampling to increase external validity and purposive sampling to increase transferability

While Teddlie and Yu (2007, p. 83) have identified similarities between probability and purposive sampling, in that they both provide a sample which is likely to answer research questions which may be generalised and transferred to an external context, they also identify differences in the representativeness of each sample. Probability sampling is designed to select many cases which are collectively representative of a population. An example to illustrate this within the context of the measurement of 3-D spatial visualisation skill would be to measure the performance of all students registered on healthcare and biomedical science. While this approach is likely to provide a general impression of spatial visualisation skill, the findings may not be specific enough to radiotherapy practice. Conversely, purposive sampling approaches will seek to recruit from a smaller number of cases which will yield more in depth information relating to a specific phenomenon. The sampling strategy should therefore allow transfer and generalisation of the findings and conclusions to other groups and contexts, for example other healthcare learners and practitioners working in complex clinical environments where 3-D spatial visualisation is a key component of practice. However, Teddlie and Yu (2007, p. 87) also indicate that the study sample should be of a certain size relative to the population. While it has to be acknowledged that during the data collection time period of the studies reported here, training places were limited by the number of commissions determined by the Department of Health and Health Education England, the sample size still had to be large enough to be representative.

In addition, consideration should be given to the issues occurring throughout the duration of any programme of research which result in data sets being incomplete. Known as missing data, the taxonomy for its classification was originally proposed by Rubin (1976, p. 581) and identified three types which Peugh and Enders (2004, pp. 526-527) described as missing completely at random, missing at random and missing, not at random. Data missing completely at random occurs when participants are absent from a data collection session due to other commitments, by choosing not to attend or by withdrawing from the study permanently. This is referred to by Newman (2014, p. 374) as person level missingness. Data missing at random (but not related to the value of the missed variable), is caused by participants failing to answer a specific question, failing to provide an answer to a survey item because they do not wish to divulge information or missing a complete section. Whereas if the value of the missing variable is related to the reason why it is missing, then the data is referred to as data missing not at random. From the studies conducted during this programme of research, one example would be participants with low spatial visualisation missing some test items because they are perceived to be too difficult to comprehend. Data missing at random and not missing at random involving single questions constitute item level missingness and if an entire section is missed then this is termed construct level missingness (Newman 2014, p. 374). Through an examination of the type and level of missing data the researcher can identify techniques for the handling of incomplete datasets. The following techniques, described by Newman (2014, p. 384), may be applied singly or in combination to any data set containing missing data. For item level missingness, one item is enough to represent a construct, therefore participant's responses should not be ignored if some items are missing and all available data should be used. This means that values for missing variables for one participant should not be replaced with the mean of that variable for all other participants (known as single imputation). If the construct level missingness is greater than 10%, then maximum likelihood and multiple imputations should be used. Multiple imputations provide a method for identifying and replacing missing values by a random sample of plausible values. It uses the distribution of the observed data to estimate the missing data values by repetition of a sequence of operations multiple times, rather than using a single value. For missing data at the person level and a response rate below 30%, a simple missing data sensitivity analysis should be conducted.

4.16 Data collection methods, order and timing

The primary dimension of interest for this programme of research was the measurement of the 3-D spatial visualisation skill of pre-registration learners in radiotherapy by employing a quantitative approach. The secondary dimension was to identify the impact of the factors of age, gender, dominant hand and engagement with activities of a spatial nature on 3- D spatial visualisation using a qualitative approach. The relative weighting of the contributions of each approach emerging during the planning stages as QUANT + qual. This was based firstly on the aims and objectives of the research identified in chapter 1.11, table 1.3, p. 21, secondly on the questions that were formulated from these aims and objectives (chapter 3.12, table 3.6, p. 128) and finally on the world view and experience of the researcher.

In relation to the order and timing of data collection, four main typologies have been identified by Creswell (2011, p. 69) and are designed to determine whether quantitative and qualitative results and findings converge or how follow up qualitative findings help explain the initial quantitative results and vice versa. The first, known as convergent parallel design, involves the researcher gathering both quantitative and qualitative data at the same time. Analysis of both datasets occurs separately and independently and then compares the results. Mixing at the interpretation stage will determine whether the results support or contradict each other. The direct comparison of the two datasets by the researcher provides a convergence of data sources.

The second, termed explanatory sequential design consists of collecting and analysing quantitative data first and then collecting qualitative data to assist with the explanation of findings or to elaborate on the initial quantitative results. The rationale for this approach is that the quantitative data and results provide a general picture of the research problem; with further analysis, specifically through qualitative data collection explaining, refining or extending the general picture. The alternative is the employment of an exploratory sequential approach. This has the researcher prioritising the collection and analysis of qualitative data followed by the collection of quantitative data. The purpose of this approach is to use the qualitative data to explain the initial findings or explore a phenomenon and then employ the quantitative data to test any relationships found in the qualitative data. Finally, the simultaneous or sequential collection and analysis of quantitative and qualitative data within a traditional quantitative or qualitative design, is referred to as an embedded approach. The purpose is to use one form of data to support or argue against the other form of data in order to enhance the overall research design.

4.17 Research design for this programme of research

During this programme of research, six separate studies (summarised in table 4.4) were designed, developed and employed in two phases. The first study in the pilot phase aimed to determine if the 3-D spatial visualisation skill of diagnostic imaging and radiotherapy pre-registration learners could be measured. It was designed to test paper and online versions of a combined mental rotation and cross sectional solids test (discussed in chapter 3.8, pp. 98 - 100). The aim of study two was to determine if the test could detect change over time and study three sought to compare the acceptance and usability of both paper and online versions. In the experimental phase, study four was designed as a controlled, longitudinal study which sought to determine the baseline 3-D spatial visualisation skill of volunteers recruited from cohorts of year one learners in diagnostic imaging (the control group) and radiotherapy (the experimental group) and to track any development over time. Study five examined the relationship between experimental group baseline spatial visualisation skill and performance in a complex 3-D radiotherapy task. The final study (study six) explored the relationships between participant demographic characteristics and baseline 3-D spatial visualisation skills. Both phases employed a purposive sampling technique and utilised a convergent parallel approach. In doing so, qualitative data collection was embedded within a predominantly quantitative methodology, thereby reflecting a QUANT-qual, concurrent mixed model design.

Table 4.4 Summary of study phases and research design

Pilot Study Number	**Study Design and Aims**
Study 1	Develop a quantitative study to collect demographic data and to test paper and online versions of a combined mental rotation and block cutting test instrument designed specifically for this programme of research
Study 2	Employ a quantitative study to determine whether the paper- based and online tests employed in study one could detect change in spatial visualisation skill over time
Study 3	Deploy a qualitative questionnaire to gauge the acceptability and usability of the online test instruments compared with the traditional paper-based testing method
Experimental Study Number	**Study Design and Aims**
Study 4	A controlled, longitudinal study designed to determine baseline spatial visualisation skill prior to clinical preparation practical workshops and to track any change in spatial visualisation skill that may occur over time
Study 5	An observational, quantitative study designed to determine the relationships between baseline spatial visualisation skill their performance in a complex simulated clinical positioning task in the experimental group of radiotherapy students
Study 6	A qualitative study to determine if a relationship exists between demographic profile and self-reported spatial activities and baseline spatial visualisation skill (measured in Study 4)

4.18 Chapter summary

This chapter has provided an overview and discussion relating to research worldviews and paradigms and the philosophical underpinning of this programme of research by considering the epistemological and ontological nature of knowledge in radiotherapy. It continued with an exploration of the proposed research design and methodology together with the key ethical considerations for the research. The chapter concluded with a personal reflection relating to the position of the researcher within the context of the programme and a description of the phases and the methodology employed for each of the six studies undertaken within this programme of research.

Chapter 5

The pilot phase studies

5.1 Introduction to chapter five

The narrative review of the development and current position for radiotherapy (chapter 2.3, pp. 28 - 35) and the review of the spatial visualisation literature (chapter 3.3 & 3.4, pp. 76 – 81) identified and defined the complimentary skills of mental rotation, perception and spatial visualisation. The successful combination of these skills is an important component in the successful 3-D visualisation of the spatial relationships between the tumour target volume, relational anatomy and planned radiation beam path in radiotherapy. The evidence base documenting the measurement of spatial visualisation skill in radiotherapy in general is limited with no reports of longitudinal studies using a combination of test instruments.

This chapter will begin by discussing the design and development of a 3-D spatial visualisation test instrument (3-D SVT) based on the Vandenberg and Kuse MRT and the SBST. It will continue with a discussion of the methods and materials, data collection and findings of the three studies conducted during the pilot phase of this programme of research. The role of a pilot study has been identified by Thabane, et al., (2010, p.2) as a small-scale study that can assist the planning and preparation of a more comprehensive, larger study. It may also determine the feasibility, which Hertzog (2008, p. 180) refers to as the identification and resolution of problems with the proposed methods and procedures prior to the implementation of the larger study. Finally, a pilot study can identify other requirements such as time and investigator resources and gather information which may be used to refine or modify research methodology. The chapter will, therefore, conclude with an examination of the implications of the findings of the pilot phase in relation to design refinements for the experimental phase of the research which will be covered in chapter six.

5.2 Test instrument design and development

Study one required the development and piloting of a paper based and online 3-D SVT using a combination of mental rotation and cross section cutting plane test objects. These were selected from the MRT developed by Vandenberg and Kuse (1978, p. 599) and the SBST a cross sectional perception and visual penetration test (Cohen and Hegarty, 2007, p. 180). The justification for developing an online version was based on the observations of Middleton et al., (2009, p. 301) who identified that, as radiotherapy technology has evolved; many parts of the workflow are achieved in a paperless (digital) environment.

This phasing out of paper-based systems in favour of online processes would suggest that computer based testing of spatial visualisation skill would be a more appropriate and authentic assessment platform. In addition the rationale for moving from traditional a paper based spatial visualisation test format to an online format was based on published evidence suggesting that online testing offers automatic randomisation of item order. This has the advantage of reducing the possibility and risk of any test-retest practice effects (Quaser-Pohl & Lehman 2002, p. 246) and order effects (Terlecki, Newcombe & Little 2008, p. 998) in follow up testing particularly in longitudinal studies. Online testing has also been reported by Monahan, Harke and Shelley (2008, p. 425) to reduce the impact of the widely reported male – female performance differential usually seen in the MRT. This was considered to be an important factor to explore in radiotherapy since it is a predominantly female profession.

5.2.1 The paper based test platform

The MRT is available in its original version which is composed of 20 test objects or in its redrawn format which is composed of 24 test objects (Peters et al, 1995, p. 42). Test items are presented on A4 paper with five or six test objects displayed per page. The SBST is composed of 30 objects[3] which are also available on A4 paper with two objects being displayed on each page. Both test instruments are supplied with written test administration instructions and practice test objects with answers and can be found in appendix six

5.2.2 The online test platform

For study one, the online test was developed in Microsoft PowerPoint with test objects being scanned as JPEG images and then inserted into individual slides for presentation as a PowerPoint slideshow. For study two, test objects were uploaded as JPEG images into QuestionMark Perception®, the quiz module of the University of Portsmouth virtual learning environment at the time.

5.3 Selection of test objects and randomisation of object appearance

The paper-based test (3-D SVT Set 1) employed half of the MRT objects taken from the redrawn version (12 test items) and half of the SBST test objects (15 test items). Each of the MRT objects is displayed with a 15^{o} tilt from vertical but with a varying degree of

[3] One object (object 3) was removed due to an incorrect representation of one of the answer choices

rotation around the vertical and horizontal axes. This provides a random appearance of individual object rotations throughout the test. As the order of appearance for the MRT objects is already randomised, the decision was taken to use the first 12 objects in sequential order for the first paper-based test, followed by the remaining items in sequence for the first Microsoft PowerPoint online test (3-D SVT Set 2). For the planned follow up testing, conducted in study two, the objects used for the paper-based test in 2011 were used in the online test in 2012 and vice versa, i.e. the 3-D SVT Set 1 test objects became the online test and the 3-D SVT Set 2 test objects were used for the paper based test.

The order of appearance and presentation of the SBST objects follows a recurring pattern of single joined and embedded geometric objects as shown in table 5.1 below. Examination of the table shows a pattern of the type of cutting plane used, for example, objects seven, eight, nine and ten are all cut with an oblique plane as are the triplet single, joined and embedded objects 25, 26 and 27. To ensure an equal representation of geometric shapes in each test the objects were divided into three groups according to their structure (single, joined or embedded) and placed in three envelopes. Manual random number selection from each envelope in turn provided test items from each of the three object types for use in studies one and two, with a second randomisation taking place to select to determine object order of appearance.

5.4 Test presentation

An answer booklet was designed for the paper-based tests and contained information relating to the purpose of the study, full instructions for completion of the test, sample test objects with answers and the mental rotation and solids cutting test objects themselves. For the first, Microsoft PowerPoint online test, the answer booklet contained test instructions and practice objects as above and a grid on which to indicate their answers (see appendix 6 c).

Table 5.1: Santa Barbara Solids Test object type and associated cutting plane

Object N^{o}	Object Type	Cutting Plane	SBST Test Object	Object Type	Cutting Plane
1	Single	Horizontal	**16**	Single	Oblique
2	Joined	Vertical	**17**	Joined	Vertical
3*	Embedded	Oblique	**18**	Embedded	Horizontal
4	Single	Vertical	**19**	Single	Vertical
5	Joined	Vertical	**20**	Joined	Oblique
6	Embedded	Vertical	**21**	Embedded	Horizontal
7	Single	Oblique	**22**	Single	Oblique
8	Joined	Oblique	**23**	Joined	Oblique
9	Embedded	Oblique	**24**	Embedded	Vertical
10	Single	Oblique	**25**	Single	Oblique
11	Joined	Vertical	**26**	Joined	Oblique
12	Embedded	Horizontal	**27**	Embedded	Oblique
13	Single	Horizontal	**28**	Single	Horizontal
14	Joined	Vertical	**29**	Joined	Oblique
15	Embedded	Oblique	**30**	Embedded	Oblique

(*Test object 3 was removed by the developers due to the incorrect presentation of answer choice)

5.5 Timing considerations

The recommended standard timing for the redrawn Vandenberg and Kuse MRT described by Peters et al., (1995, p. 42) sets a time limit of three minutes for each half of the 24 object test (15 seconds per test object). This timing is the same as that in the original Vandenberg and Kuse 20 item test which recommended a time of three minutes per split half of 10 objects (18 seconds per test object).

When describing the development and administration of the SBST, Cohen & Hegarty (2007, p. 181) did not specify a time limit, although in a later study the authors reported that most participants completed the test in five minutes. Given the time constraints placed on radiotherapy treatment appointments, an unlimited time for completion was not considered to be appropriate for this study. While it is important that role of research and the practical situation should not be confused, it has been acknowledged by Peters (2005, p. 177) that while spatial visualisation skills may evolve in response to the environmental demands of a particular task, the environment does not always permit the luxury of unlimited time to complete that task. For the spatial visualisation studies in this programme of research, it was also important to determine the impact of time pressure on decision-making. Taking these observations into consideration, a compromise time limit of four minutes was set for the mental rotation subcomponent and five minutes was the limit placed on the cutting planes test (equivalent to 20 seconds per test item. This decision was based partly on the timing conventions reported in the literature and the researchers' personal clinical experience in workload planning where 12 minutes was allocated per complete treatment delivery appointment. At the time of the planning of the pilot phase studies, published literature relating to workload and process timing was limited. However a study by Van de Werf, Lievens, Verstraete, Pauwels and Van den Bogaert (2009, p.138) analysed data from 324 randomly selected radiotherapy treatment sessions and reported that the mean in- room time (the time from the patient entering the room to the time the patient leaves the room) was reported as 11.6 minutes (SD = 5.9) for conventional 3-D CRT delivery. For more complex treatment delivery, such as a seven field IMRT technique, this time increased to 13.6 minutes (SD = 5.4) and with the addition of IGRT the mean was 17.3 minutes (SD = 6.8). It should also be recognised that patient positioning prior to the start of imaging and / or the start of the treatment process will take a smaller proportion of this time, but these timings support the justification for the time allowed for completion of the spatial visualisation test.

5.6 Performance scoring and interpretation

The standard scoring convention for the MRT awards one point for each correct pair identified and one point for each correct solids test item. The critical evaluation of the spatial visualisation measurement literature identified a range of scoring conventions for the Vandenberg & Kuse MRT. Although it is acknowledged that much of the literature provides no justification of why these methods had deviated from the developer's original recommended scoring convention. For the studies reported here, the conventional scoring method of one point for each correct pair identified was adopted throughout.

The SBST scoring convention awards one point for each cutting plane correctly identified. Additional analysis of the number of incorrect answer choices can be applied to determine the number of times the egocentric distractor (foil) has been selected. This has been reported by Cohen and Hegarty (2007, p. 183) to offer an additional indicator of less well-developed spatial visualisation skill, since those who select this answer may have difficulty in changing their perspective relative to the orientation of each test item.

5.7 Demographic questionnaire

Data for age and gender was collected to gain a general profile of the study cohort. As the pilot phase was conceived in order to test both paper and online platforms, it was also considered important to gauge self-rated confidence with the use of information technology as this may have an impact on the level of engagement with an online testing platform. If performance between platforms was shown to be comparable, then this would offer the opportunity for further development of the online platform as students would be unlikely to be disadvantaged by such a move.

5.8 Study 1 Comparison between paper and online spatial visualisation test platforms

The first study was designed to determine the spatial visualisation skill and demographic profile of volunteers recruited from the 2010 – 11 year 1 cohort of the BSc (Hons) Diagnostic Imaging and Therapeutic radiography (Radiotherapy) programmes at the University of Portsmouth using the paper based and online versions of the 3-D SVT described above.

5.8.1 Method and materials

The first data collection event employing the paper-based test format was scheduled to take place in April 2011 in a flat space classroom. Participants were presented with the answer booklet described above (section5.4, p. 148) and asked to follow the instructions as they were read to them by the test administrators (the researcher, assisted by one other member of the Radiography academic team). Correct answers for the practice objects for each test were given by the administrators prior to participants starting each test section. All participants were presented with the same test objects in the same order. The MRT was administered first, followed by the SBST. This order was selected because it was considered to be most representative of the patient positioning workflow in radiography. In diagnostic imaging, the first step in the process will align the patient to the image receptor (usually orthogonally); while in radiotherapy the patient will be aligned with the linear accelerator isocentre using the external positioning coordinates on the patient's skin. This will involve the mental visualisation of internal anatomy and physical rotation of the patient. The SBST would replicate the alignment of the diagnostic X-ray tube or linear accelerator gantry and the mental perception of the relationship between the proposed beam path and internal anatomy. Demographic data relating to gender, age, dominant hand, perceived confidence with computer technology and computer gaming experience, (refer to appendix 7), was collected via a self-report survey questionnaire which was completed once the time allocated for the solids test had elapsed.

5.8.2 Recruitment and sampling

There are few recommendations in the social science literature regarding sample size (Johanson & Brooks, 2010, p.395), but in clinical research, the literature identifies a range of acceptable sizes. For example, Julious (2005, p.291) makes a recommendation of 12, while Herzog (2008, p. 181) makes reference to sample sizes of between 10 and 40 for between subjects studies. As the potential recruitment pool of radiotherapy students registered on the 2010 – 11 BSc (Hons) Therapeutic Radiography year one programme was relatively small at 21, the decision was made to also recruit from the diagnostic imaging cohort. This decision was justified by the similarities between the two pathways in terms of the spatial visualisation skills required for patient positioning and X-ray beam alignment. This resulted in a total recruitment pool of 80 students made up of 59 (73.7%) diagnostic imaging and 21 (26.3%) radiotherapy students. Using a purposive convenience sampling approach, all first year students on both programmes at the University of Portsmouth were invited by email (delivered via the University intranet) to attend a briefing session at the beginning of April 2011. The session was delivered by the researcher at the end of a timetabled teaching session. Following a short introduction by the researcher which outlined the purpose of the study, data collection methods and participant commitment, information sheets and consent forms were distributed by the cohort student consultative committee representatives. In acknowledgement of the power relationships between the researcher as a programme lecturer and students as study participants, consent forms were returned individually to the designated drop box located in the radiography academic office or brought to the first data collection event. The first collection of spatial visualisation test performance and demographic data was planned for mid-April following the final clinical placement of year one and to fit in with already scheduled academic events with follow up testing scheduled for June following the end of year summative assessment period.

5.8.3 Data anonymization, and collation

Participants were requested to place their unique six digit University of Portsmouth identification number on the front cover of their answer booklet for the sole purpose of collation with any subsequent data collection. At the end of the first data collection session answer books were collected by the researcher and identified with a unique participant number commencing at one. This numerical identifier was linked to each participant's university registration number throughout the pilot phase to link all subsequent data for individual participants. This process ensured that the researcher was blind to individual participant identity prior to any data coding and interpretation, thus minimising the risk of interpretation bias.

Population data for performance in the MRT and the SBST scores were calculated manually and combined to provide an overall raw score for both components. This score was also converted to a percentage value to facilitate comparison with other reported studies which employed different numbers of tests and test objects. These were entered manually into a Microsoft Excel® spreadsheet. Following checking for accuracy they were exported to the Statistical Package for Social Sciences versions 20 - 25 (IBM Corp 2012-2019) and JMP Pro version14.0 (Statistical Analysis System Institute, 2018) for descriptive and multivariate statistical analysis.

5.8.4 Proposed methods for statistical data analysis

A combination of descriptive and inferential bi- and multivariate methods was employed in the analysis of the data from all studies. The first stage would use tools such as the mean (for central tendency), the standard deviation of the mean and the identification of maximum and minimum values to describe the characteristics of age, gender and performance. The second would use tests for significance and included analysis of variables (ANOVA) to determine if the performance of the two populations of diagnostic imaging and radiotherapy students were significantly different using the hypotheses:

H_o: There is no statistically significant difference between the populations;

H_1: There is a statistically significant difference between the populations.

The difference was identified as significant if the p-value was $< .05$ and the Tukey- Kramer Honest Significant Difference (HSD) was used as a post hoc test to determine where in the populations the significance lay. For the measurement of associations Chi-squared analysis, following the Pearson method, was employed to indicate whether relationships between data-sets (the goodness of fit) was present or whether patterns were due to independent factors. It should be noted that any relationship is based on a statistical analysis, rather than implying a real world relationship. The qualitative data from demographic questionnaires was analysed and reported using the same descriptive methods and observations from experience surveys were analysed thematically using the phased advocated by Braun and Clarke (2012, p. 58).

5.9 Study one results and analysis

This section will present and analyse the results from study one which compared the performance of a cohort of year one radiotherapy students (referred to as pathway 1) and diagnostic students (pathway 2) in paper and on-line versions of a 3-D SVT. The section will begin by describing the participant flow through the study time points prior to comparing demographic profiles for age and gender. To gain an overall indication of test performance, the results for all participants across both pathways will be presented first. This will be followed by comparative performance analysis by pathway and then by test subcomponent.

The participant flow for study one is summarised in a Consolidated Standards of Reporting Trials (CONSORT) flow diagram (figure 5.1). Developed for the reporting of randomised clinical trials, CONSORT flow diagram explicitly shows the number of participants for each intervention group who are included in primary data analysis. The use of a flow diagram in the reporting of trials is recommended by Moher, Schulz and Altman (2001, p. 1193). The structure has been adapted here to show participant flows through each of the studies that measured 3-D spatial visualisation skill during this programme of research. From the recruitment pool of 80 students, 26 volunteers consented and attended the first data collection session in April 2011. This corresponds to 12 diagnostic imaging students (20.3% response) and 14 radiotherapy students (66.7% response) and an overall response rate of 32.5% from both programmes.

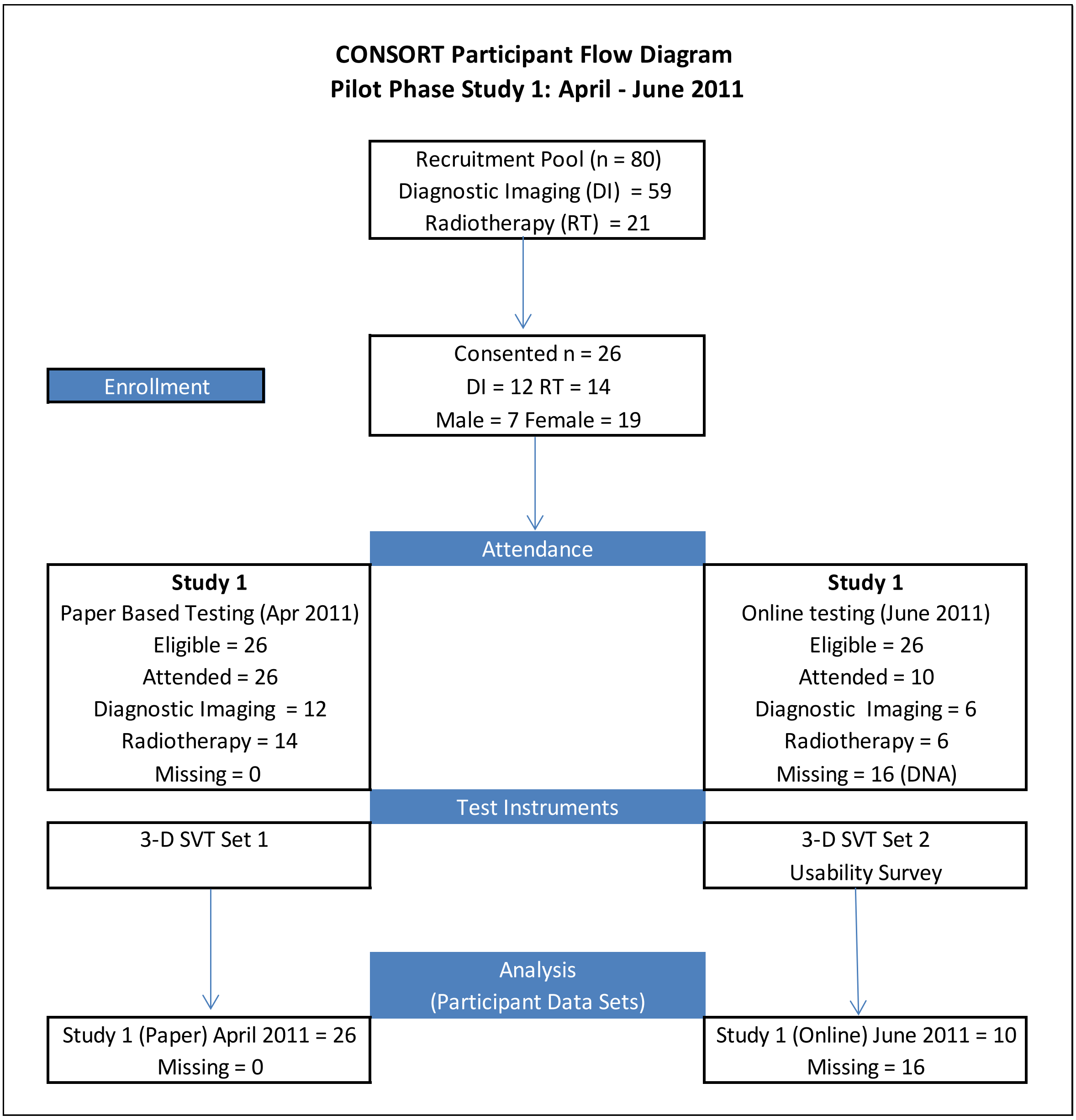


Figure 5.1: Summary of participant flow numbers and test instruments for study one

5.9.1 Participant demographics for gender and age

Of the 26 participants who consented to take part in the piloting of the 3-D SVT, seven (27%) were male and 19 (73%) were female. While the Society and College of Radiographers regularly publish data relating to the demographic profile of the radiographic workforce, they do not collect specific data relating to the profile of pre-registration learners. However comparative data for the 2014 – 15 academic year, the closest time period to this study available from the Higher Education Statistics Agency (HESA, 2020) showed that the proportion of males to females across all subjects allied to medicine (n =182,930) demonstrated that males accounted for 20.8%, (n = 38,135) and females 79.2% (n=144,795), a slightly higher female proportion compared to the females in the pilot study cohort. The latest available data for registered radiographers in the United Kingdom, published by the Health and Care Professions Council for male and female registrants, demonstrates that 8,317 (24%) are male and 25,965 (76%) are female (HCPC, 2018).

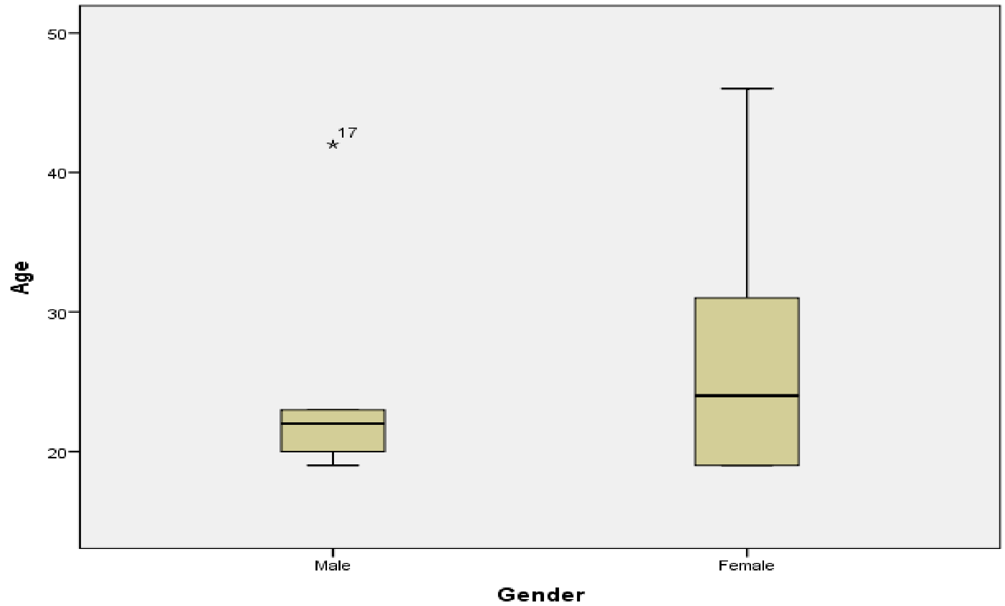


Figure 5.2: Boxplot for age distribution by gender showing similar median values but a smaller interquartile range for males

The mean age of the participants who responded to this question (n=25) was 26.1 years (SD = 8.5), with a range of 19 -46. Further analysis of the age profile of participants utilising a box plot for gender is shown in figure 5.2. This shows that while the median age (23) is similar for males and females, the interquartile range is smaller for males compared to females overall. These findings are compared with HESA age data for all first year undergraduate students in England based on in table 5.2. Overall the data would suggest that female radiotherapy students are older than their diagnostic counterparts and may be explained by a tendency for a higher number of females taking up study in radiography as a second career, compared to males.

Table 5.2: Comparative age data for all first year undergraduate students and pilot phase radiography students at the University of Portsmouth (UoP)

Age Range	All First year Undergraduate Students (England) 2014/15	Proportion	Radiography (UoP)	Proportion
≤ 20 years	721,545	63%	9	36%
21-24 years	298,220	26%	6	24%
25-29 years	54,720	5%	2	8%
≥30 years	6,530	6%	8	32%
Total	1,140,015		25	

5.9.2 Paper based spatial visualisation performance results

Analysis of the paper based test performance results, expressed as a percentage score for both subcomponents combined, shows an overall mean for all participants of 50.73% (n = 26,SD = 16.6, range = 8 – 81%) as shown in the histogram (figure 5.3) and the associated box plot (figure 5.4) below. Additional analysis to determine if there was a difference in performance in the paper based test between the two pathways demonstrated no statistically significant difference determined by one-way analysis of variance (ANOVA) ($F(1,24) = 0.41$, $p = .053$. Further analysis employing a means plot can be used as a graphical addition to ANOVA to test the equality of population means. The vertical bars show the degree and direction of variance of each subgroup mean compared with the overall population mean. The upper and lower decision limits are based on a significance level of .05, so any point falling outside these limits is likely to be significantly different from the overall mean. The plot (figure 5.5) shows an overall mean for all participants (n=26) of 50.73. Pathway 1, the radiotherapy participants, had a mean score of 48.8% (SD = 19.2, n = 14) while pathway 2, the diagnostic imaging students, achieved a mean score of 53% (SD = 13.3, n = 12).

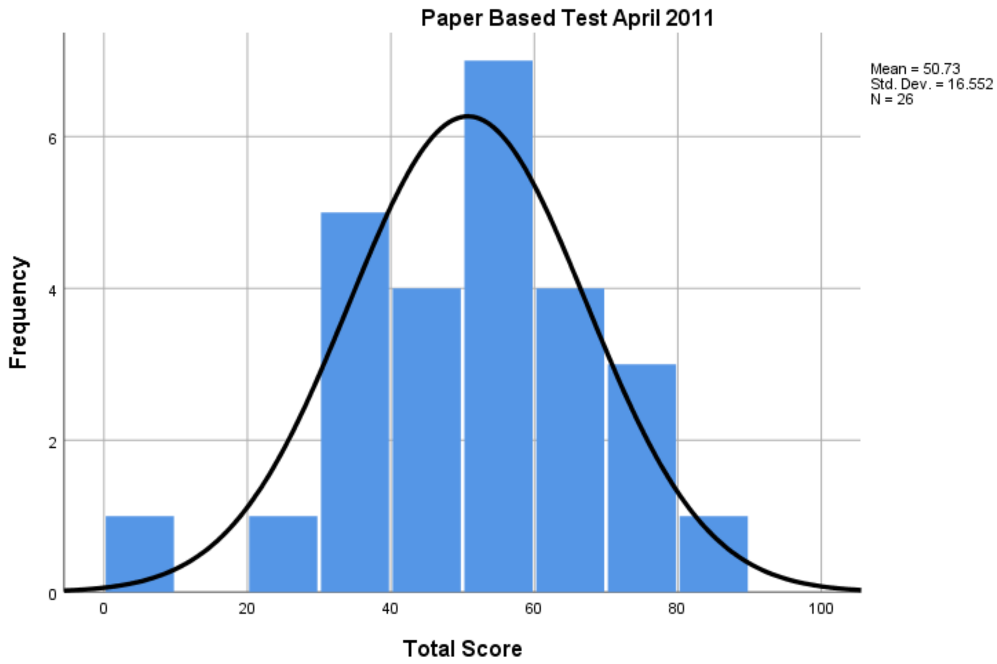


Figure 5.3: Distribution of total performance score for all participants

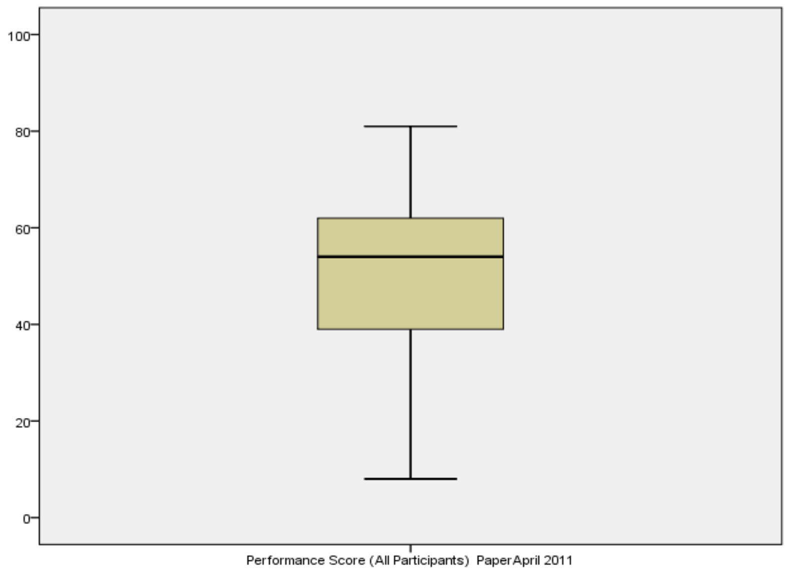


Figure 5.4: Boxplot for the performance score achieved by all participants in the first paper based test

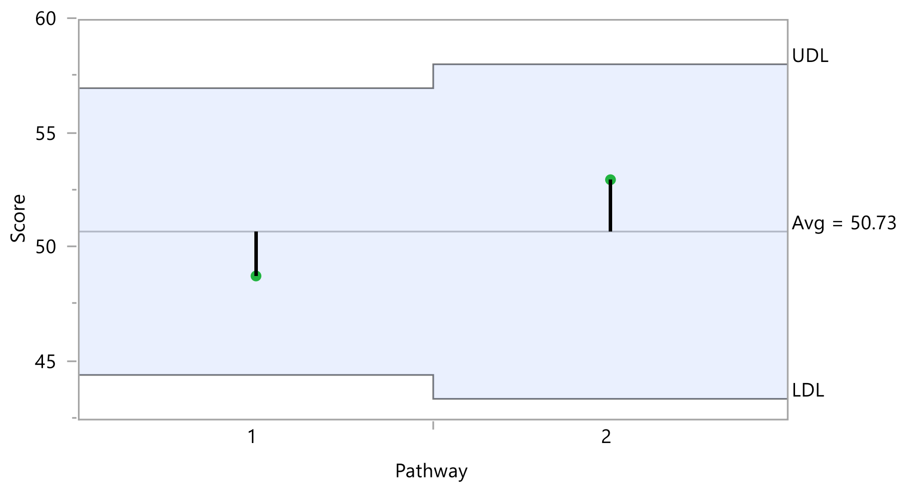


Figure 5.5: Analysis of means plot demonstrating a superior performance by diagnostic imaging students (pathway 2) in the paper-based test (April 2011)

5.9.3 Online spatial visualisation performance results

From the original 26 participants who completed the paper-based data collection event at time point 1 in April 2011, 10 participants (38.5%) attended the second part of the data collection for study one which took place in June 2011, 49 days later. This test employed a Microsoft PowerPoint based platform, accessed via the University of Portsmouth virtual learning environment, WebCT. The overall performance mean for all participants was 48.8% (SD = 18.3, n = 10) as shown in the histogram (figure 5.6) and the box plot shown in figure 5.7 demonstrates an overall performance range of 66% with a minimum score of 19% and a maximum of 82%. Participants from the radiotherapy group (pathway 1) achieved a mean score of 40.3 (SD = 14.5, n = 6), while their diagnostic imaging counterparts achieved a mean score of 61.5 (SD = 17.3, n = 4). The analysis of means plot (figure 5.8) shows a larger performance differential for the diagnostic imaging participants compared to the paper based test in April 2011, but a one way ANOVA demonstrated that there were no statistically significant differences between each pathway ($F(1,8) = 4.42$, $p = .06$. The lower number of participants from both pathways in the June test (n=10) compared to April (n=26) is acknowledged. The impact of this missing data will be discussed further in section 5.5.4 (p. 170) below.

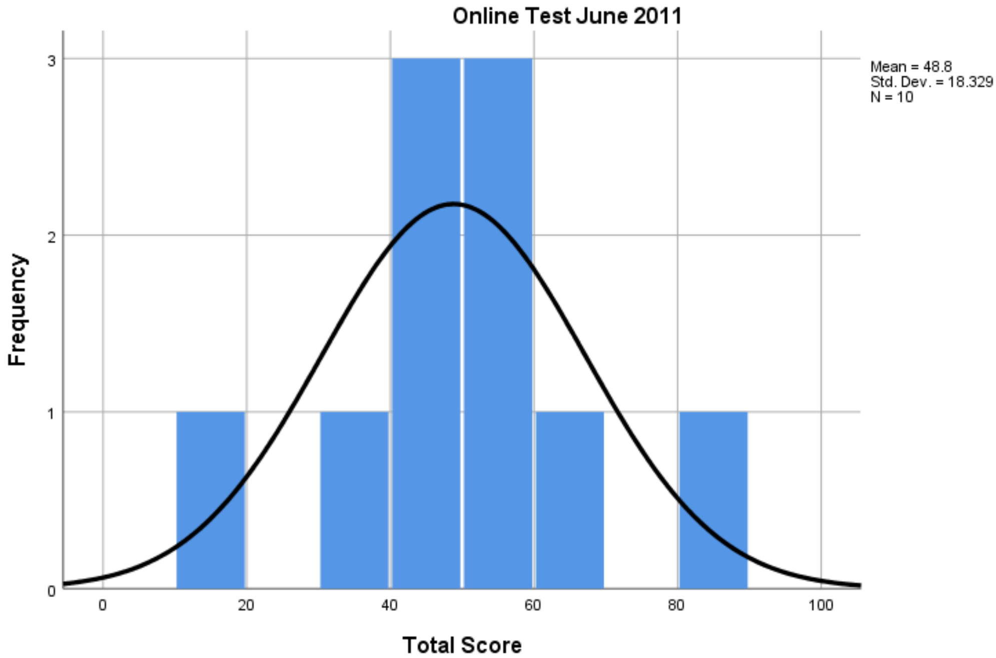


Figure 5.6: Histogram and normality plot showing distribution performance score for all participants in the online test in June 2011

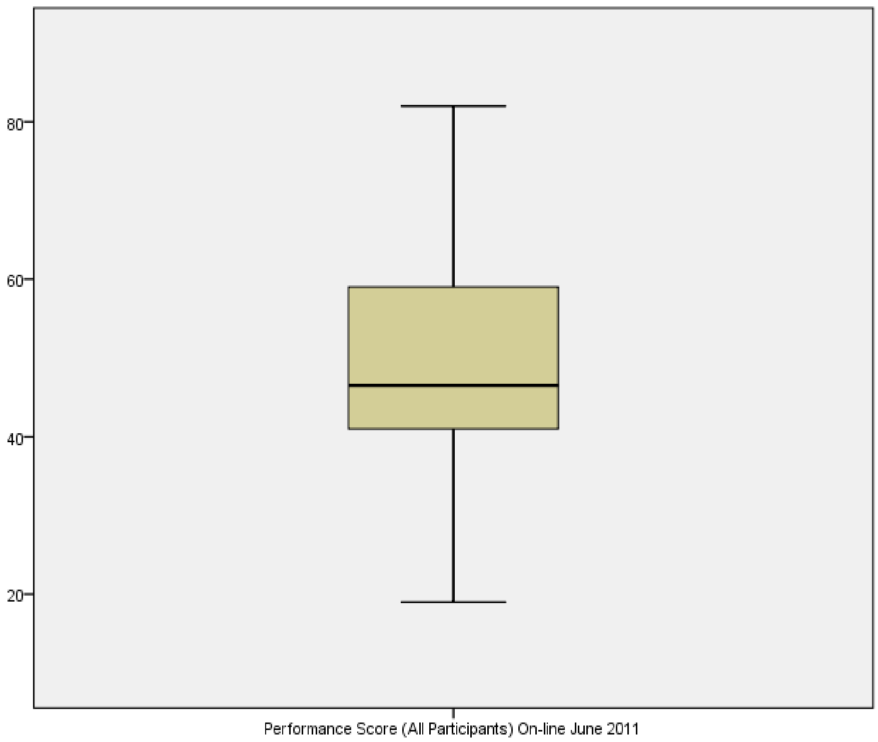


Figure 5.7: Boxplot for performance score for all participants from both pathways for the online test (June 2011)

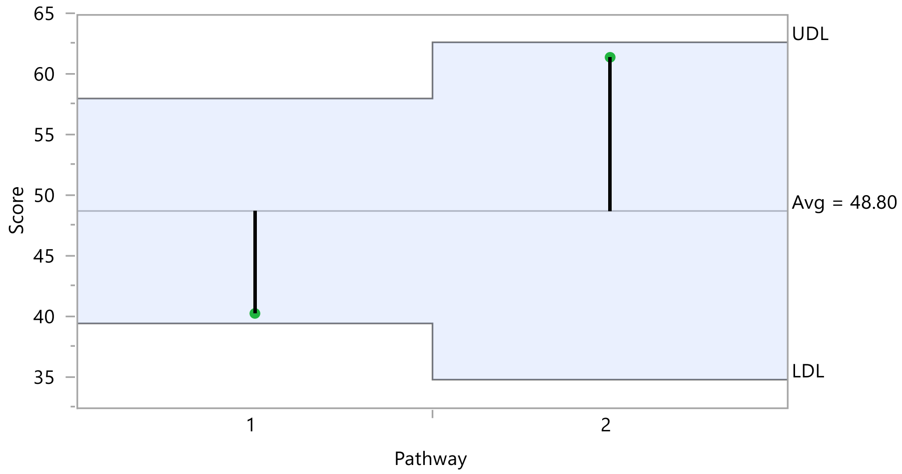


Figure 5.8: Analysis of means plot demonstrating a superior performance by diagnostic imaging students (pathway 2) for the online test (June 2011)

5.9.4 Test subcomponent and missing data analysis

The previous sections (5.5.3 & 5.5.4) have presented and analysed the results for overall performance in the 3-D SVT (with scores for both subcomponents combined) in its paper and online versions. While this has provided valuable understanding, spatial visualisation skill has been identified as having three subcomponents; mental rotation, spatial perception and spatial visualisation, hence the combination of two test instruments in this programme of research. So it was important to determine if there were any differences in performance between the three domains and both tests. Mental rotations require an object to object manipulation and transformation with respect to an environmental frame of reference, while the individuals` egocentric frame of reference remains unchanged. However, the SBST requires individuals to imagine object cross sections by reorienting their viewing position through a transformation of their egocentric frame of reference. This involves the mental imagination of objects and their cutting planes changing from an orthogonal view (looking directly at the paper or on-screen view) to a view of the cutting plane as if viewed in a mirror (refer to chapter 3.11, pp. 126 – 127, and figures 3.18 - 3.20).

Performance scores in the two subcomponents for the paper based test, conducted in April 2011 and expressed as a function of the number of correct answers is shown in table 5.3. From this it can be seen that males outperform females across both types of test objects, which appears to support the performance differential in favour of males which is widely reported in the literature (Linn & Petersen, 1985. P. 1487).

Table 5.3: Number of correct answer choices for mental rotation, solids test and total by gender

	Paper Based Test April 2011 Number of Correct Answers					
	Male			Female		
	MRT	SBST	Total	MRT	SBST	Total
n	7			19		
SD	2.1	2.1	3.4	2	3.4	4.6
Mean	5.3	10.9	16.1	4.2	7.9	12.2
Min	3	8	11	1	1	2
Max	8	14	21	7	14	20

Further examination of individual participant performance in the mental rotation and the solids cutting tests showed data missing at random at the item level in the paper based platform, as not all of the 26 participants attempted all test items within the specified time limits. In the online test, data was missing completely at random at the person level and missing at random at the item level. For the paper based tests, all participants attempted mental rotation items one to six, while five (19.2%) missed item seven and 25 (96.2%) missed the final item. In the SBST, all participants attempted the first 10 items with just five (19.2%) failing to answer the final question (item 15). By comparison, 13 (50%) of the participants who attended the paper based testing session undertook the online test. Analysis demonstrated mental rotation item level missingness for all participants, while all participants attempted items one to six, they all missed item 12, while all participants attempted all solids test items.

The impact of these missing values was predicted using multiple imputations (100 iterations) and compared to the original raw data. Chi square tests of the differences between the original raw dataset and the new imputation sets indicate no statistically significant difference between datasets across all time points as the following values demonstrate. For the online test in June 2011 $X^2(1, N = 36) = 0.08, p = 0.77$, while the paper based test conducted in April 2012 returned an $X^2(1, N = 39) = 0.019$, $p = 0.88$ and for the online test in April 2012 $X^2(1, N = 36) = 1.39$, $p = 0.23$. Across the entire data set available for the paper version there were 312 items (26 participants and 12 test items), from which a total of 117 (37.5%) items were not attempted. For the solids test there were 364 items, of which 13 (3.6%) were not attempted. To determine if this missing data would have an impact on overall performance score across the two subtests, additional analysis using a ratio scoring method as proposed by Geiser et al., (2006 p. 265). This alternative method calculates the performance score in each subtest based on the number of correct objects identified as a percentage of the number of objects attempted. The comparative scores for each scoring convention are shown in table 5.4. A paired-samples t-test was carried out to compare the recommended standard scoring method and ratio scoring. This demonstrated that there was a significant decrease in score for the MRT with standard scoring (M = 37.5, SD = 16.63) compared to ratio scoring (M = 61.08, SD = 27.03), t (25) = 8.59, p <.0005 (two-tailed). The mean increase was 23.58 with a 95% confidence interval ranging from 17.92 to 29.23. However there was no significant difference found with the same scoring method for the solids test with standard scoring (M = 61.27, SD = 23.56) compared to ratio scoring (M = 64.92, SD = 23.56), t (25) = 2.44, p < 0.22 (two-tailed). The mean increase was 3.654 with a 95% confidence interval ranging from 0.57 to 6.74.

Table 5.4: Comparison of spatial visualisation performance scores using recommended and ratio scoring

	Mental Rotation Test		**Santa Barbara Solids Test**		**Total Combined Score**	
	Standard Scoring (%)	**Ratio Scoring (%)**	**Standard Scoring (%)**	**Ratio Scoring (%)**	**Standard Scoring (%)**	**Ratio Scoring (%)**
n	26		26		26	
Mean	37.5	61.1	61.3	64.9	50.73	63.2
SD	16.6	27.03	23.6	24.57	16.6	23.1
Range	8-67	14-100	7-100	7-100	8-81	11-100

The reasons for non-completion of all test items could be twofold, either participant's ran out of time due to slow decision-making or they did not attempt items because of their perceived difficulty. This would suggest that participants may have found the mental visualisation of object rotations more challenging than object cross sections. While it is important to acknowledge that research and clinical practice should be considered separately, the non-completion of tasks by a learner in a time constrained test environment may indicate challenges with visualisation and hence clinical decision-making. While it is not possible to provide a more in depth analysis of incorrect answer choices for the MRT because the test instructions and answer key only provides information relating to correct answers, it is possible to determine the type of incorrect answer choice (known as the egocentric distractor or foil) in the solids test. This is the incorrect answer option that a participant may select if they are not able to change their viewing perspective (egocentric frame of reference) relative to the cutting plane and has been reported to be a possible indicator of less well developed spatial skill. This analysis is provided in detail for all available data across all test platforms for studies one and two in section 5.9, p. 182, below.

In a mental rotation study involving 501 (28%) male and 1264 (72%) female psychology students (n=1765), Peters (2005, p. 178) found that applying a time limit of three minutes for 12 items resulted in 145 males (29%) and 246 females (19.5%) failing to attempt the final three items. A follow up study involving 212 students drawn from the same subject pool examined the effect of doubling the time allowance. The results showed the number of items attempted in the first half of the test increased from eight items (67%) with standard timing of three minutes to 11 items in six minutes for males and seven items (56%) with standard timing and 11 in six minutes for females. While the results demonstrated that females benefit from additional time, the results are not significant since males also benefit from the additional time. So, whilst an increase in time allowances for future studies was considered, it was concluded that it would be inappropriate, since time is a luxury that cannot be afforded in the clinical situation.

5.10 Confidence with information technology

While the move to online processes reduce the risk of human transcription errors Middleton, et.al, (2009, p. 304) have identified that they do require high end IT skills. Radiography in general has migrated away from paper-based systems towards more integrated digital communications platforms and automated processes. Requests for diagnostic imaging are now made online; those images are acquired using digital image receptors and viewed on computer monitors, having been retrieved from a central repository. In radiotherapy, treatment related data including the treatment delivery plan, the size, shape and number of treatment beams; linear accelerator gantry and collimator positions are all stored on a central server. Recognising that learners in radiography come from diverse backgrounds and may have had varying exposure to and experience with information technology (IT), it was considered important to gauge level of confidence. In addition, engagement with computerised systems depends on an individual's age, and experience with and willingness to accept new technologies (Elias, Smith & Barney, 2012, pp 453 - 455). If a large proportion of respondents had indicated low confidence, then the proposed move to online testing may have put some students at a disadvantage. It can be seen from figure 5.9 that most respondents felt confident or very confident with their IT skills.

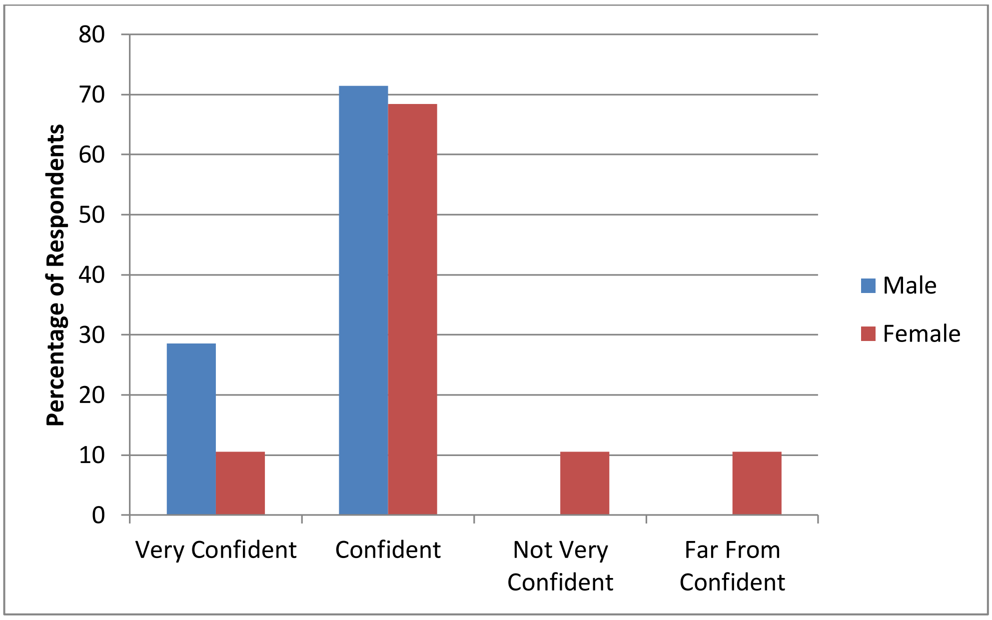


Figure 5.9: Participant self-report for confidence with information technology

5.11 Validity and reliability

To determine the reproducibility of the 3-D SVT subcomponents and the consistency of the results achieved in the paper based and online platforms, the performance of the 10 participants who completed both iterations was analysed by comparing the number of correct answer choices in the MRT and SBST subcomponents. The descriptive statistics are shown in table 5.5 and show a similar performance, both in relation to the number of questions attempted and the number of correct answers achieved in each subcomponent. To investigate this relationship further and to determine and the reliability of online testing further, the performance in each subcomponent was compared using Cronbach's alpha coefficient. For the MRT and the SBST paper and online subcomponents, Cronbach's alpha for both was .6. A comparison between the paper based MRT and SBST items demonstrated a Cronbach alpha of .52 and for the online items the value was .76. Coefficient values above 0.7 indicate good consistency as identified by Walker and Almond (2010, p. 86). While Tavakol & Dennick (2011, p.54) identified different reports about the acceptable values of alpha, ranging from 0.70 to 0.95, they refer to earlier work by Bland and Altman (1997, p.572) who suggested that for scales which are used as research tools to compare groups, alpha values between 0.7 to 0.8 can be regarded as satisfactory.

Table 5.5: Comparison of performance in paper and online 3-D SVT subcomponents

	Paper MRT Questions Attempted	Paper MRT Correct Answers	Online MRT Questions Attempted	Online MRT Correct Answers
Mean	8	6	8	5
SD	2.0	3.0	2.6	2.3
Min	6	2	4	3
Max	12	12	12	10

	Paper SBST Questions Attempted	Paper SBST Correct Answers	Online SBST Questions Attempted	Online SBST Correct Answers
Mean	14	10	15	8
SD	1.3	3.2	0.7	3.3
Min	10	5	13	2
Max	14	14	15	12

In relation to the development of their Mental Rotation Test, Vandenberg and Kuse (1978, pp. 601 – 602) reported Pearson Product – Moment correlations with the card rotation test of .62 and for Shepard & Metzler identical blocks .68. For other spatial tests, such as hidden figures and form boards these values are lower at .4 and .41 respectively. While the initial validity reported for the MRT was determined by correlations with what might be argued to be tests for general spatial skills, it has been widely adopted in visualisation studies since its introduction over four decades ago. When Cohen and Hegarty (2007, p. 181) reported on the development of the Santa Barbara Solids Test, in a study of 59 psychology students, they also employed the MRT and the Visualization of Views Test. They reported that the performance in both tests was highly correlated ($r = .47$) and using averaged score from both tests, which they referred to as the spatial score, reported a correlation of .5 ($p < .01$) with all types of test figures and cutting planes in the SBST. They also reported a split half Cronbach Alpha for internal consistency for the 29 test items in the SBST of 0.86, which they referred to as a satisfactory.

5.12 Introduction to study 2

The second study of the pilot phase sought to determine if any changes occurring in 3-D spatial visualisation performance over time could be detected by the paper-based and online test instruments.

5.12.1 Method and materials

The data collection plan for the second study in the pilot phase is summarised in table xx below. The paper based quantitative 3-D SVT used in study one was presented to participants in a question and answer booklet in the same format as that employed in study one. Test objects for the online test were scanned as JPEG images into QuestionMark Perception®, the quiz module of the University of Portsmouth virtual learning environment at the time. The 26 volunteers from the 2010-11 first year cohorts of the BSc (Hons) Diagnostic and Therapeutic Radiography programmes, who had previously completed the paper-based test in April 2011, were invited via a student intranet email, to participate in follow up testing 54 weeks after initial testing. Paper based testing took place in a university flat space classroom. The online test was scheduled to take place seven days later to fit with other timetabled activities and prior to the final clinical placement of the year, was conducted in a University of Portsmouth open access IT suite using standard specification desktop personal computers and monitors.

Answers for the paper-based test were written in an answer booklet of the same design used for the April 2011 paper-based test. The answer booklets were identified by individual participants University student identification numbers. Following manual marking and checking by the researcher they were independently checked by another member of the Radiography academic course team. Performance scores were then entered manually into the Microsoft Excel® spreadsheet employed for study one and collated with those results using the student identification number. Participants accessed the online test in WebCT® via their individual University username and password. The online test was marked automatically in QuestionMark Perception® as a percentage score and these results were manually entered into the same Microsoft Excel® spreadsheet for checking and comparative analysis.

5.13 Study 2 Results and analysis

Of the 26 participants who participated in study one, in 2011, 13 (50%) attended paper based follow up and 10 (38.5%) attended online testing in April 2012, at the end of their second year of study. The CONSORT participant flow diagram (figure 5.10) provides a summary of participants flows.

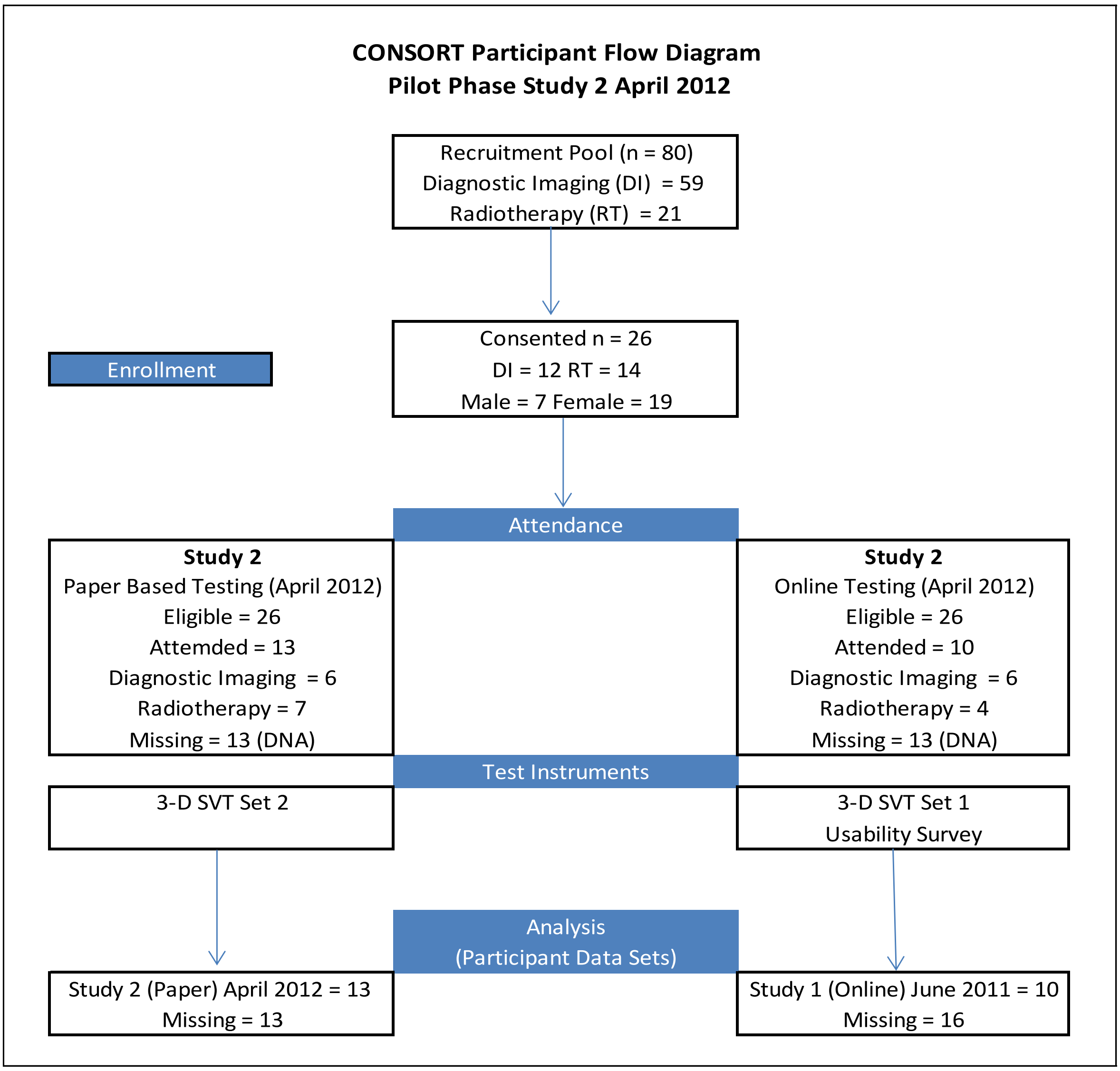


Figure 5.10: Summary of participant flow numbers and test instruments for study two

The paper-based test was administered first and produced an overall mean for both pathways combined of 50.69% (SD = 23.6 n = 13). The distribution of total performance scores and the associated normality plot are shown in figure 5.11 while the box plot is exhibited as figure 5.12. The analysis of means plot (figure 5.13) demonstrates that radiotherapy participants (pathway 1) achieved a mean score of 42.7% (SD = 22.2, n = 7) while the diagnostic imaging participants (pathway 2) had a mean score of 60% (SD =

23.4, n = 6), the minimum score was 19% and the maximum 82%. No statistically significant differences between each pathway were identified as determined by a one-way ANOVA (F (1, 11) = 1.87, *p* = .2.

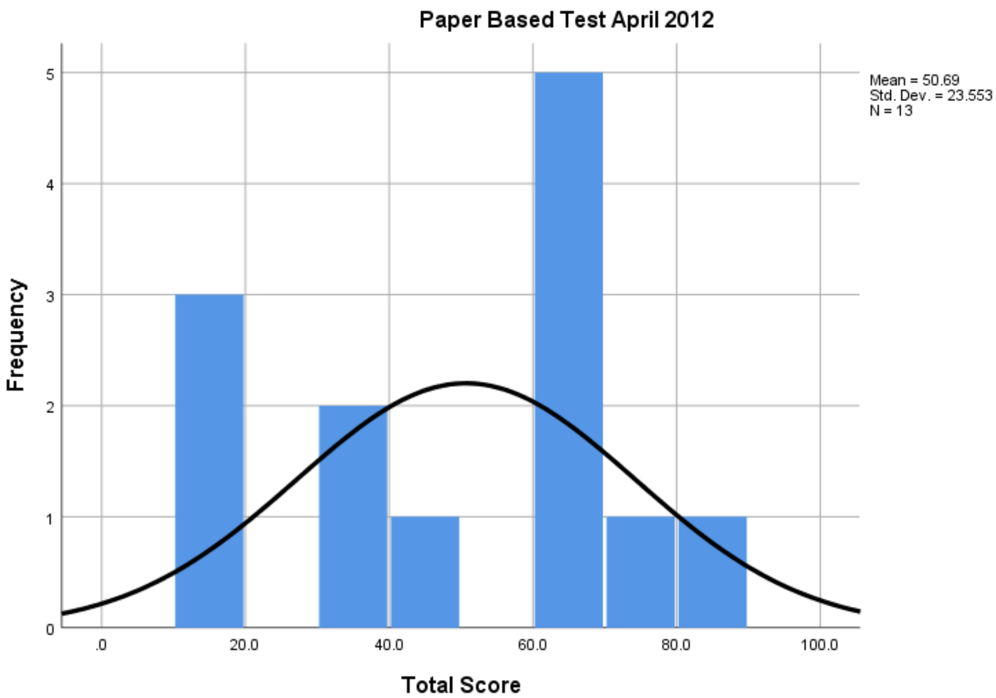


Figure 5.11: Histogram and normality plot showing distribution performance score for all participants in the paper test in April 2012

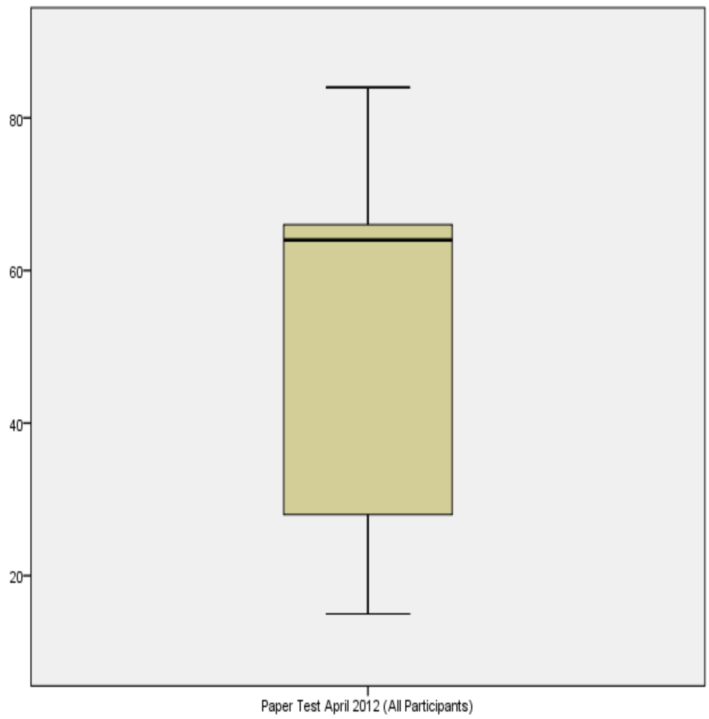


Figure 5.12: Boxplot for performance score for all participants from both pathways for the paper test in April 2012

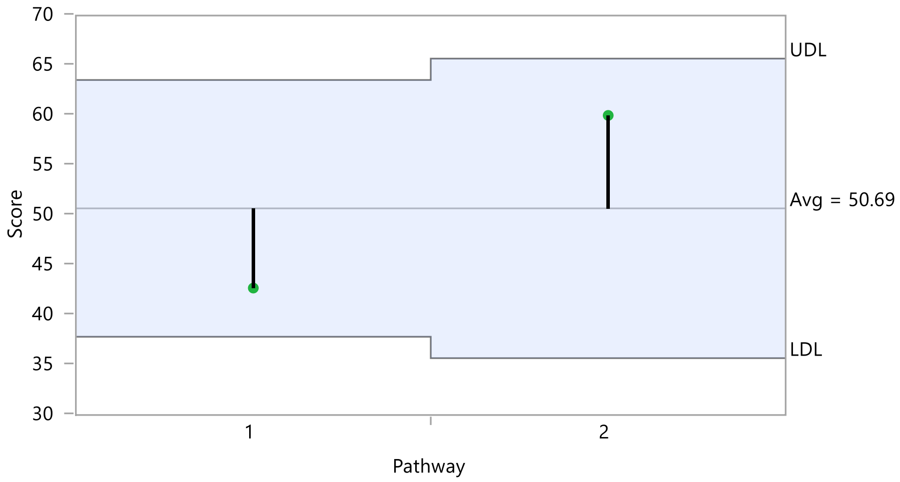


Figure 5.13: Analysis of means plot demonstrating the performance differential in favour of the diagnostic imaging pathway remains

The online test was administered seven days later demonstrated an overall mean of 48.3% (SD = 23, n = 10). The minimum score gained was 10% and the maximum 70%. Participants from the radiotherapy group had a mean score of 42.3% (SD = 29.3, n = 4) while the diagnostic imaging students had a mean score of 52.3% (SD = 19.7, n = 6). There were no statistically significant differences between each pathway and the 2012 online test means as determined by one-way ANOVA (F (1, 8) = 0.43, p = .5. The distribution of total scores for all participants with a normality plot is demonstrated in figure 5.14.

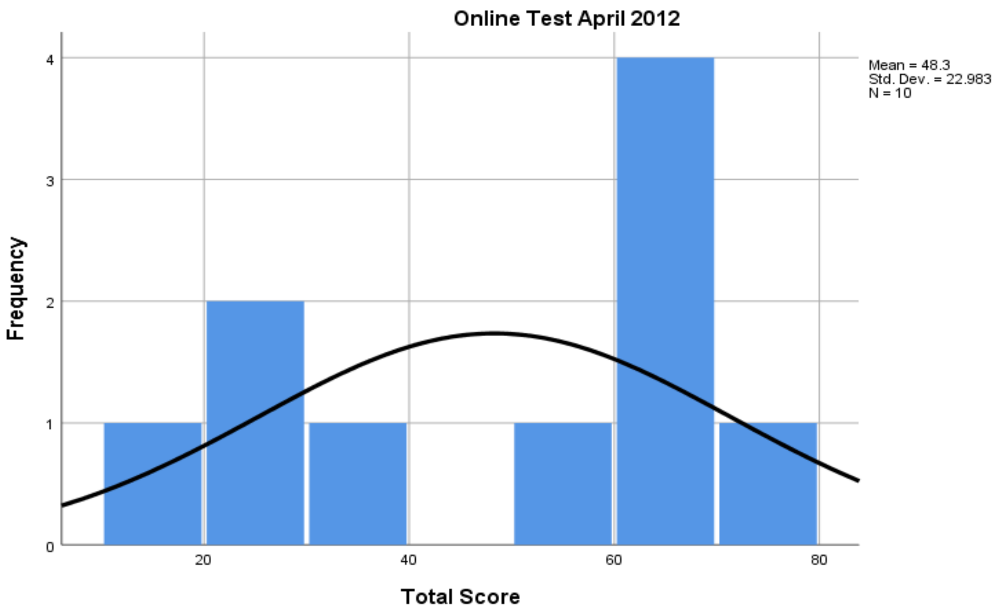


Figure 5.14: Histogram and normality plot showing distribution performance score for all participants in the online test in April 2012

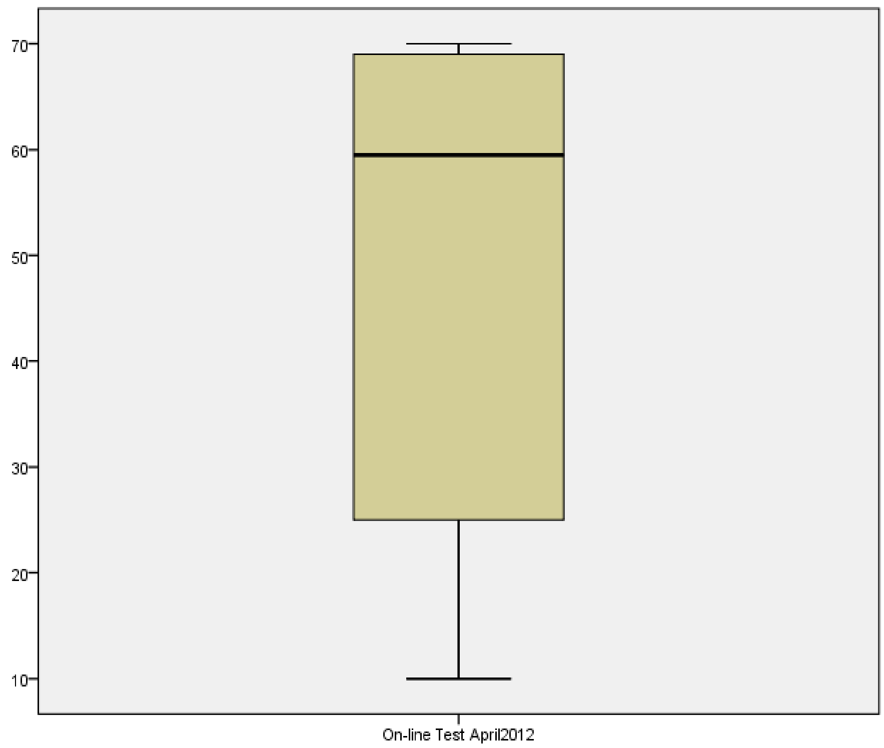


Figure 5.15: Boxplot for performance score for all participants from both pathways for the online test in April 2012

The analysis of means plot (figure 5.16) continues to demonstrate the performance differential in favour of the diagnostic imaging pathway but not as strongly as in the paper-based test.

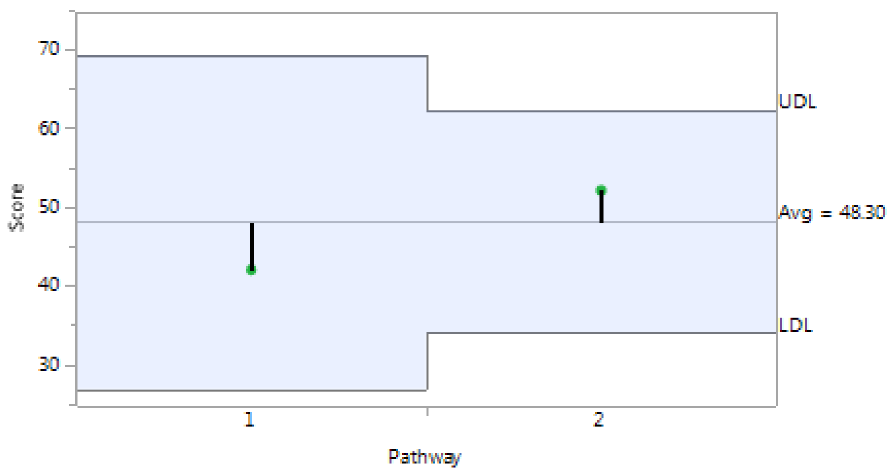


Figure 5.16: The analysis of means plot for online testing (April 2012) showing a continued performance differential for diagnostic imaging participants

5.13.1 Comparison of performance over time

The comparative performance scores in both platforms, for all participants and for all time points in studies one and two are shown in table 5.5 and demonstrates that the mean performance for the paper-based platform is similar for both time points, but the range is narrower and both minimum and maximum scores are higher in April 2012 compared to 2011. The online test in April 2012 produced a narrower range but the spread of scores shows a minimum of 10% and maximum of 70% which is lower than the performance demonstrated in the online test in April 2011 and the paper-based test in April 2012.

Table 5.6: Summary of spatial visualisation performance for all participants and all time points

	Paper 2011	**Online 2011**	**Paper 2012**	**Online 2012**
n	26	10	13	10
Mean (%)	50.7	48.8	50.7	48.3
SD	16.6	18.4	23.6	23
Range	73	63	66	60
Minimum (%)	8	19	19	10
Maximum (%)	81	82	85	70

A further breakdown of comparative performance scores, by programme pathway at all data collection time points is shown in table 5.6. This demonstrates a varying profile across all time points, but a performance advantage in favour of diagnostic imaging students remained throughout the study.

Table 5.7: Performance score comparison for both pathways and test platforms for all time points

	Paper 2011	**Online 2011**	**Paper 2012**	**Online 2012**
Pathway Group 1 (Radiotherapy)	N = 14	N = 6	N = 7	N = 4
Mean	48.8%	40.3%	42.7%	42.3%
SD	19.2	14.5	22.2	29.3
Min	8	8	19	10
Max	81	59	67	69

Pathway Group 2 (Diagnostic Imaging)	N = 12	N = 4	N = 6	N = 6
Mean	53%	61.5%	60%	52.3%
SD	13.3	17.3	23.4	19.7
Min	27	41	19	23
Max	73	82	85	70

No statistically significant differences were demonstrated for the participants from the radiotherapy pathway and each test mean as determined by one-way ANOVA (F (3, 27) = 0.32, p = .81. This was also the case for the diagnostic imaging pathway, one-way ANOVA (*F* (3, 24) = 0.43, *p* = .74. The analysis of means plots for both pathways at all time points are shown in figures 5.17 and 5.18 and confirms the performance differential in favour of the diagnostic imaging pathway with an overall mean across all time points of 55.6% compared to 45% for the radiotherapy pathway.

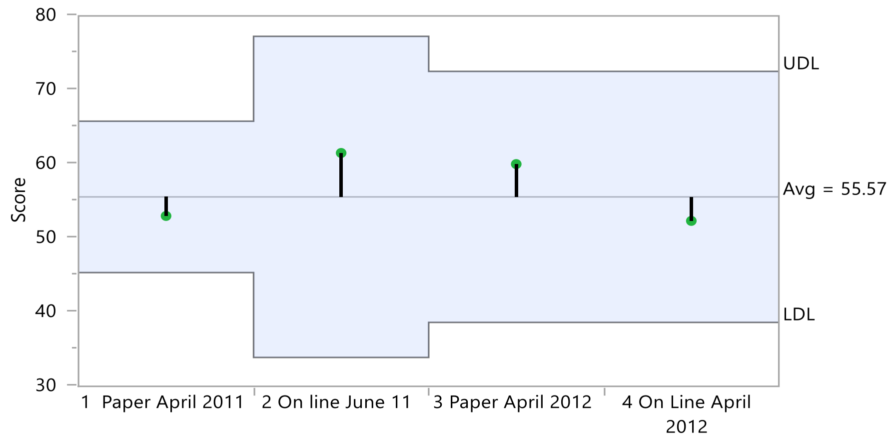


Figure 5.17 Analysis of means plot for the diagnostic imaging pathway at all time points

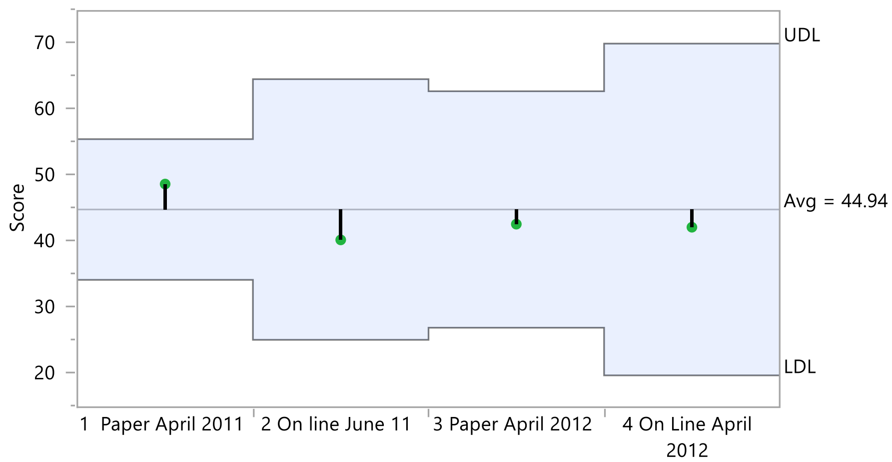


Figure 5.18 Analysis of means plot for the radiotherapy pathway at all time points

Comparing the performance scores from all tests for both pathways demonstrates that there were statistically significant differences between the diagnostic imaging participant (pathway 2) test means and those of the radiotherapy students (pathway 1), with diagnostic imaging students performing better than their radiotherapy counterparts at all time points (one-way ANOVA ($F(1,57) = 4.82$, $p = .03$). A post hoc Tukey HSD test confirmed this significance at $p < .05$. This confirms that, in this study cohort, diagnostic imaging students seem to have better 3-D spatial visualisation skill than the radiotherapy students (figure 5.19).

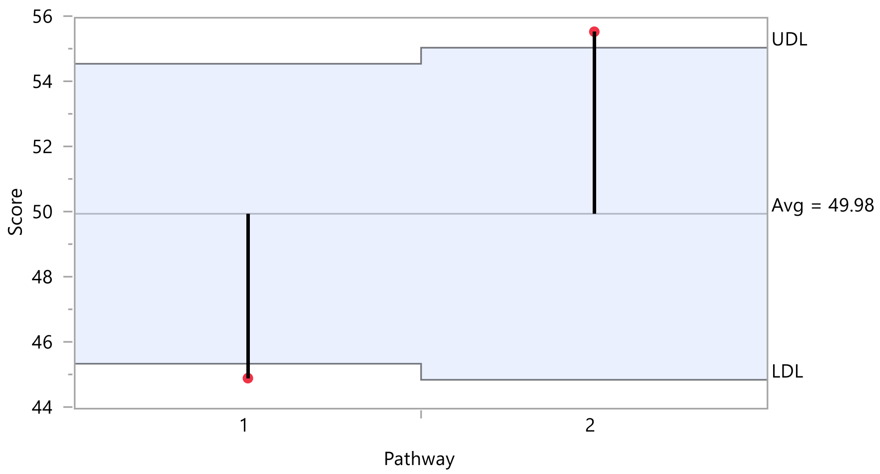


Figure 5.19 Analysis of means plot for performance in all tests, confirming the significant difference in performance shown by diagnostic imaging students (pathway 2)

5.14 Analysis of egocentric distractor choices

As reported in studies conducted by Cohen and Hegarty (2007, p. 183; 2012, p. 869), if a participant does not, or cannot change their view perspective when completing the SBST items, they may select an incorrect answer choice known as the egocentric foil. The authors also suggest that the likelihood of this happening is more frequent in those participants with lower spatial visualisation skill and may be used as an alternative method for screening for those individuals.

To determine the relationship between incorrect answers when the foil has been selected, which may offer an alternative indicator of less well developed spatial visualisation skill; the number of times the egocentric foil was selected was plotted against the total number of incorrect choices by participants in the paper based tests in April 2011 and April 2012. The resultant scatter plots shown in figures 5.20 and 5.21 below demonstrate the relationship between incorrect answer choices and the number of times the egocentric foil was selected. Regression analysis of the results shows a strong positive relationship between the numbers of foils selected and the number of incorrect answers (R^2 = .82 for the April 2011 test and R^2 = .94 for the April 2012 test). This relationship, if replicated in the experimental studies, could lead to the potential for the egocentric foil analysis to be employed as a supporting measure for identifying those learners with less well-developed spatial visualisation skill.

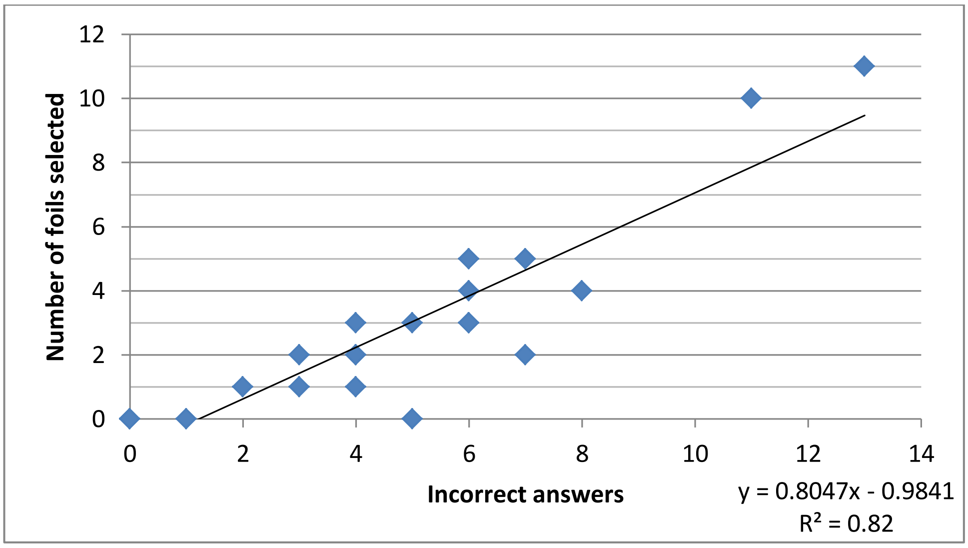


Figure 5.20: Scatter plot to demonstrate the relationship between incorrect answers and egocentric foil choices, showing a relatively strong positive correlation for the paper-based April 2011 SBST

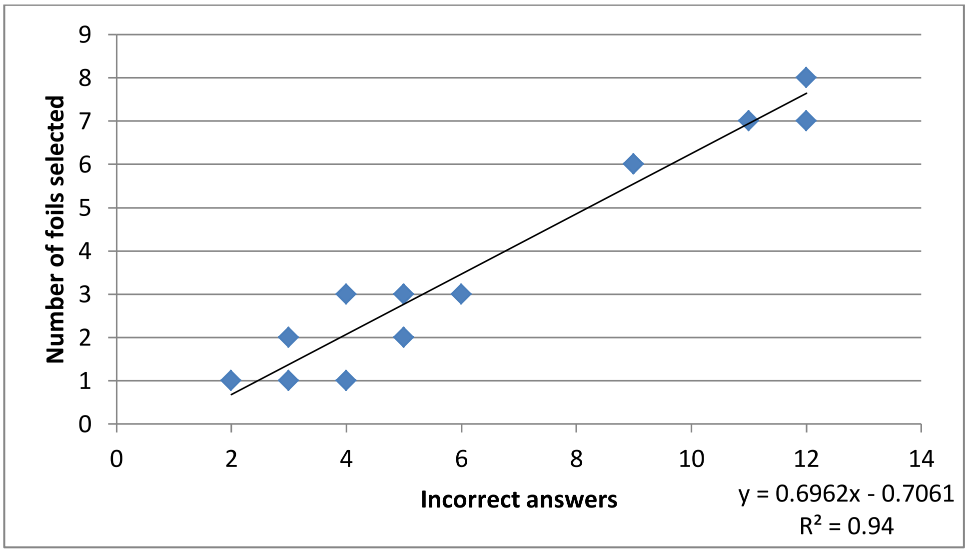


Figure 5.21: Scatter plot to demonstrate the relationship between incorrect answers and egocentric foil choices, showing a strong positive correlation for the paper based April 2012 SBST

Extracting the same data for the online test in April 2011 showed a much weaker relationship (R^2 = .04) as shown in figure 5.22 below.

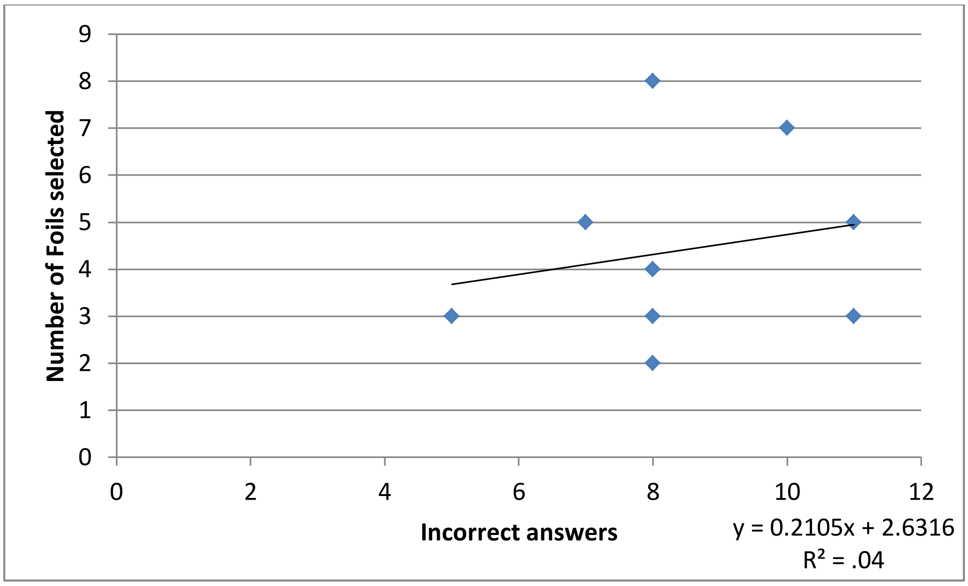


Figure 5.22: Scatter plot to demonstrate the relationship between incorrect answers and egocentric foil choices, showing a weak positive correlation for the June 2011 online SBST

Due to the way that QuestionMark Perception® calculated overall performance score, no breakdown of item order or the number of correct versus incorrect answers was available. As it was not possible to extract this data for the online test in April 2012, a comparative analysis for the online tests has not been possible and so no conclusions may be drawn in relation to online testing. However this will be carried forward and monitored during the experimental phase to determine possible trends. Further analysis of individual object complexity, type of cutting plane and egocentric distractor choice across all time points shows that across all object types (single, joined and embedded) cut with an oblique plane, the egocentric distractor was chosen in 67%, 75.6% and 69% of cases respectively as shown in figure 5.23. Given these findings, the increasing use of radiotherapy treatment beams at oblique, rather than cardinal angles and the ability to visualise beam paths and anatomical relationships is likely to be challenging for those learners with less well-developed spatial visualisation skills. From a diagnostic imaging perspective, if a patient cannot be placed in the recommended optimal position for a particular image projection angle, the patient position, imaging technique and X-ray tube angle may need to be modified to accommodate this change in patient position. This would require the application of a combined mental model of anatomical position and beam direction similar to that required in radiotherapy.

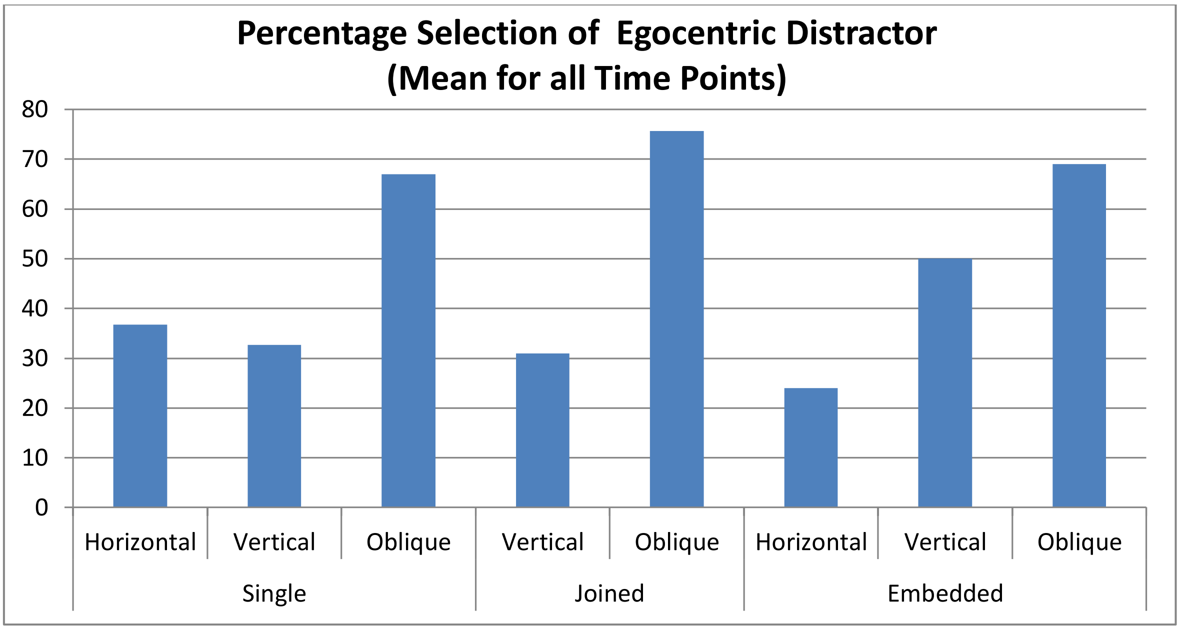


Figure 5.23: The relationship between object complexity, cutting plane type and proportion (%) of participants selecting the egocentric distractor (foil)

While an overall performance differential was observed between the radiotherapy and diagnostic imaging cohorts across all time points and on both platforms it was not statistically significant. There was, however, a difference in the pattern of incorrect answers for the SBST subcomponent. The incorrect answers are known as distractors and are categorised as alternate, combination and egocentric. If the identification of an incorrect answer is a purely random process, then there should be an equal number of each type of distractor selected. Of most interest is the egocentric distractor referred to as the foil from this point on), since this is the shape that participants might imagine if they fail to translate their view perspective relative to the cutting plane of the object (Cohen & Hegarty, 2007, p. 180). It is also the one that has an appearance which most closely resembles the correct answer. If the proportion of foils is higher than the proportions of the other incorrect choices then this may indicate a difficulty with the transformation of spatial representations of objects and therefore lower 3-D spatial visualisation skill.

From the 10 participants who attempted both paper and online test versions in April and June 2011, three gained a maximum score in the paper SBST. However, in the online version, one participant selected seven incorrect answers, of which five (71%) were foils, the others each selected eight incorrect answers with three (38%) and four (50%) being foils. Analysis of the selections of other participants showed that all but two had a higher number of foils in the online test. Overall, the paper based test produced 38 incorrect answers choices, of which 18 (47.4%) were foils whereas the online test produced a total of 84 incorrect answers with foils accounting for 45 (53.6%).

5.15 Discussion

As the experimental phase of this programme of research, specifically study 4 was conceived and designed as a controlled longitudinal study to determine if individual baseline 3-D spatial visualisation skills could be measured and to identify changes due to development over time. It was important, therefore, to understand whether the test instruments developed for study one could measure any change over time. The results of study two demonstrated a change in performance at an individual level for each data collection time point and test mode. A difference between the two radiography pathways has also been demonstrated, although the small number of participants from each pathway is acknowledged.

5.16 Introduction to study three

The primary aim of the final study conducted during the pilot phase of the programme of research was a qualitative survey which was designed to determine participant acceptance of the traditional paper-based tests compared with the online platforms. In addition, the findings of the survey would inform the design of the proposed online 3-D SVT platform for the experimental phase of this programme of research. To satisfy this aim, the following research question was formulated: does the acceptability (defined as suitable and appropriate) and utility (defined as fitness for purpose) of the online platform compare with the paper-based test?

5.16.1 Method and materials

Following the completion of the online Microsoft PowerPoint® version of the 3-D SVT conducted in June 2011, participants were invited to complete a paper based four part survey composed of three questions with closed answer choices of yes, no and neutral. The fourth question provided participants with the opportunity to provide free text comment to support their answer choices (appendix 7). For those participants completing the online QuestionMark Perception based test conducted in April 2012, a revised paper-based questionnaire with additional questions was developed. Participants were asked to indicate their preferences on a five-point Likert type scale ranging from strongly agree to strongly disagree and were also given the opportunity to provide free text comments to support their rating. Completion of the questionnaire took place in the same scheduled data collection sessions immediately after completion of the online 3-D SVT in June 2011 and April 2012. For those participants who completed the Microsoft PowerPoint® test in study one (n = 10), the questionnaire formed part of the answer booklet. For those participants (n = 10) completing the QuestionMark Perception version of the test, the questionnaire was available via a hyperlink link once the 3-DSVT was submitted.

5.17 Results and analysis

Descriptive analysis of the responses to the questionnaire showed that five students (50%) who completed the Microsoft PowerPoint test indicated that they preferred the PC based test, while four (40%) were neutral and one student (10%) preferred the paper-based test method. The 10 students completing the QuestionMark Perception, all agreed or strongly agreed that PC based instructions were clear. Nine agreeing or strongly agreeing that the PC based test objects were easy to see. In relation to the individual tests, six agreed that the PC based MRT was better than paper based version (two were neutral and one offered no opinion) while five agreed that the PC based SBST was better than paper based version (three were neutral and one offered no opinion).

From both iterations of the online test in 2011 and 2012, respondents provided a total of 16 free text comments. These were transcribed verbatim (see appendix 8) and analysed using a thematic approach with the aim of ascertaining the acceptability and utility of the online platform compared with the paper-based version. These themes could then be used to inform the design of the proposed online 3-D SVT platform for the experimental phase of this programme of research. Identification of themes followed the framework advocated by Braun and Clarke (2012, p. 58) and it was important at this stage for the researcher to remain mindful of not letting the aim drive the identification of themes. The first phase involved familiarisation with the comments by reading through them. This identified a first round of broad, surface level, themes and code words. A re-reading of the comments (phase two) generated a second list which identified subthemes and underlying feelings and experiences which could be linked to the overarching broad themes. The third phase reviewed all the themes and code words to ensure that each one could be distinguished from the others. Once this was completed each of the themes and associated words were tabulated and summarised (refer to appendix 8) and are discussed below. The thematic analysis identified four key themes relating to the computer and paper platforms, the images (summarised as image clarity, display, size and overall screen layout) and the clock. These themes will be explored further by examining the individual observations. Participants who recorded a preference for the online tests indicated that:

> *"Computer images were sharper" and "the test seemed easier on the PC than on paper – I didn't struggle as much with the PC test"* and
>
> *"It took a while for me to see the rotational ones on the computer but once I had the hang of it, it was much easier for me than the paper one"*

The following comment is interesting and may be related to the clarity and size of the images which some participants identified as a source of difficulty, with one participant reporting that they:

> *"Preferred mental rotation electronic and cutting test paper"*

While another indicated that:

> *"I`d like to retake the test using a PC but having a blank piece of paper to draw the images & draw how I think they would look rotated to help me choose my answer"*

This response may relate to an alternative solution strategy, reported by Hegarty and Waller (2004, p. 188), who indicated that individuals may solve visualisation problems by imagining the object being rotated or by imagining changing their perspective in relation to the object. Two participants indicated that they had difficulty in viewing the test objects online. The first reported that this was due to:

> *"Having trouble seeing black on white (been to opticians)"*

and the second identified that:

> *"…….images are……incomplete which my eye finds both distracting and confusing. It is a line going away & coming towards me".*

All students complete a health questionnaire and occupational health assessment at the commencement of their studies and a self-declaration of any changes to their health status annually thereafter. Therefore it was assumed that all participants would have normal or corrected to normal vision so the question was not asked as part of the demographic questionnaire. An alternative reason for the visual difficulties reported could have been due to a previous user adjusting the monitor resolution and then not resetting it to its default. In a pooled space, open access, computer suite this would be difficult to control for. This may also explain why some participants had difficulty with the size of some of the images. As one participant stated:

"Had to scroll down to see some of the images i.e. too big", "the images in the cutting exercise was too big I had to zoom in and out, which was distracting"

with another observed:

"Cutting plane test: having to scroll down to see the examples was disturbing"

The final source of reported difficulty was the visibility and location of the timing clock on the screen. This related to the QuestionMark Perception platform with the following observation from one participant:

"I did find the ticking clock disturbing as it was (I feel) pushing me to go faster it would have been better to have just the minutes and seconds in say increments of 30 or 15 seconds" while another indicated that:

"The fact that I could see the clock made me more stressed".

5.18 Discussion of acceptability and usability findings

Overall, the online test platforms were well received. Of the 10 participants who attempted the Microsoft PowerPoint version of the 3-D SVT as part of study one in April 2011. Only one student indicated that they preferred the paper based test, of the remaining students five indicated a preference for the Microsoft PowerPoint test and four identified no preference between paper and online tests. This prompted further developments and the test was migrated to the quiz module of the University virtual learning environment, QuestionMark Perception, for deployment in study two, conducted in April 2012. Part of the development included reviewing participant instructions for on-screen viewing and readability. All participants completing the April 2012 questionnaire (n = 10) agreed or strongly agreed that these instructions were clear. When asked whether they preferred the paper or online versions of the MRT and the SBST, eight (80%) indicated a preference for the online versions of both tests, while two offered no opinion. This finding was interesting given the thematic analysis of the free text observations which showed some concerns about image clarity, size and screen layout. Based on the overall acceptance of the online platform and taking into account participant observations, it concluded that further development for use in the experimental phase longitudinal study (study four) would be appropriate.

5.19 Summary of findings from the pilot phase studies

In addition to the acceptability and usability findings from study three, the findings relating to the analysis of spatial visualisation test performance and the impact of timing from the first two studies will also be summarised and discussed.

5.19.1 Analysis of performance

Study one demonstrated that the 3-D spatial visualisation skill of a cohort of pre-registration learners in radiography could be measured using a combination of mental rotation and cross-sectional solid object test items. While the online test in April 2011 showed a different pattern of incorrect answers for the SBST, the comparison of performance scores between the traditional paper based and on –line test platforms did not produce statistically significant differences. These findings would suggest that participants would not be disadvantaged by taking a paper or online test. The results from study two showed that both the paper and online test could detect change in performance over time.

5.19.2 Impact of timing

It is important to note that not all participants attempted all test items due to the impact of the time limits imposed. The possible influence of the timing constraints and the option of increasing the time allowance was discussed in section 5.5.4, p.173, but given that clinical decisions need to be made in a timely manner, the conclusion was that the time limit would remain unchanged. Overall, the online tests were well received by participants in both Microsoft PowerPoint and QuestionMark Perception presentation types as demonstrated by the results of the usability survey (study three). Combined with the statistical non-significance in performance score across all time points, using both online and paper testing, the conclusion drawn from the three studies was that there was scope for further development of the online test platform. This development would take account of the timing issues that were encountered and participant observations relating to the quality of the images. Developing an online test platform for 3-D spatial visualisation would also deliver other advantages including a reduction in the time required for the preparation of the test, removal of printing costs, automatic marking and download of results and randomisation of test object appearance.

5.20 Chapter summary

This chapter has reported the design and testing of online versions of a traditional paper based MRT and a cross sectional cutting planes test. Performance in the online platform was compared to that for the paper-based format. The results showed that there were no significant differences in participant performance with the online test. This would indicate that a move to an alternative platform would be unlikely to disadvantage any participant. The experimental phase of the research would therefore develop the online platform and deploy it to measure the baseline 3-D spatial visualisation skill of a cohort of students at the commencement of their radiotherapy education, compare their performance with a cohort of diagnostic imaging students and to track any development over time which may occur as a result of a combination of clinical practice and time spent in clinical simulation environments.

Chapter 6

The experimental phase studies

6.1 Introduction to the experimental phase studies

The experimental phase of this programme of research was designed to measure baseline 3-D spatial visualisation skill, to detect any change that may occur over time and to explore if a relationship existed between biological and environmental factors and test performance. Three studies were conceived and developed and a summary of their objectives, design and associated research questions is presented in table 6.1 below.

Table 6.1 Summary of experimental phase research objectives, design and questions

Study N°	Research Objectives	Design
Study 4	To determine if the baseline spatial visualisation skill of pre-registration learners in radiotherapy could be measured	Longitudinal, controlled study
Study 5	To determine if a relationship exists between baseline spatial visualisation skill and performance in a complex radiotherapy positioning task	Observational study
Study 6	To determine if a relationship between baseline spatial visualisation skill and previous spatial visualisation experience exists	Quantitative self-report survey
Research Questions		
Study 4	1. To what extent can the spatial visualisation skill of pre-registration radiotherapy students be measured? 2. Does spatial visualisation skill change during the programme of study?	
Study 5	1. To what extent does baseline visualisation skill have an impact on the performance of a complex positioning task using the 3-D virtual environment for radiotherapy training (VERT™) platform?	
Study 6	2. What factors may affect the development of spatial visualisation skill?	

Volunteers were recruited from the 2012-13 first year cohorts of diagnostic imaging students to act as a control group for the experimental group of radiotherapy students within the same institution. The justification for the inclusion of a control group was based on this group of students having access to a real digital X-ray suite in the Health Care Science Simulation Centre while the radiotherapy cohort (who would form the experimental group) would have access to the VERT™ platform.

6.2 Method and materials

The findings of the first study in the pilot phase reported in chapter 5.5, p.167 demonstrated that there were no significant differences in performance scores across the paper based and online tests. The findings of study three, reported in chapter 5.12, p.192, demonstrated that the acceptability of the online test platforms for mental rotation and cross section solids testing was non-inferior to the traditional paper based methods. Therefore the decision was taken to develop and employ the online test platform for the longitudinal study. Given the observations about the lack of clarity of the MRT images, a higher quality version of the original 20 item Vandenberg and Kuse MRT in Microsoft Word format was sourced.[4] The images for the mental rotation and solids test objects were scanned into the quiz module of the University of Portsmouth virtual learning environment (Moodle[5]). The module is capable of automatic randomisation of object type and order of appearance and was programmed to display 10 MRT objects with a three minute time limit and 14 SBST objects with a time limit of five minutes at each data collection point. The order of testing remained the same as that used in studies one and two of the pilot phase, namely the MRT was presented first, followed by the solids test. The automatic randomisation function also means that test objects can be displayed in a different sequence for each participant thereby reducing the risk of practice and order effects reported by Quaser-Pohl and Lehman (2002, p. 246) and Terlecki et al.,(2008, p. 998) respectively.

All testing would be conducted in University of Portsmouth information technology laboratories, using standard University specification desktop PC`s and monitors. Participants were permitted to select their PC and could adjust monitor screen resolution to meet individual optical and visual requirements. Monitor and seating height could be adjusted as required and the monitor viewing distance was left to individual choice based on comfort. Participants accessed the test instrument via the year one clinical learning module repository on Moodle using their secure usernames and passwords. Participants were requested to follow the on-screen instructions for the MRT section while they were read out by the test administrators. They were then asked to view the practice test

[4] http://spatiallearning.org/index.php/resources/testsainstruments

[5] Moodle replaced WebCT as the University of Portsmouth virtual learning environment for the 2012-13 academic year onwards

objects and when they had taken time to familiarise themselves with the structure, the answers were provided by the test administrators. There was no time limit for this section. At this point participants were invited to ask questions for clarification if required. When all participants indicated that they were happy with the instructions and requirements for the MRT they were instructed to start the test when they were ready. Timing would start automatically when participants navigated to the first question via a radio button link at the bottom of the practice objects screen. Navigation through the test permitted movement backwards and forwards through the test object as required during the allotted time. The test ended automatically at the end of the elapsed time.

6.3 Recruitment and sampling strategy

All students registered on the first year of the 2013-14 BSc (Hons) Radiography programmes were invited to attend a briefing session during the first teaching week in September 2013. The potential recruitment pool was made up of 43 diagnostic imaging students and 18 radiotherapy students. To provide an indication of the required sample size for the experimental study, an online sample size calculator (Creative Research Systems) was employed. Given that the recruitment pool population was 61 students, for a 95% confidence interval and a 5% margin of error and assuming a minimum response rate of 50%, the sample size would need to be 53 participants.

Information sheets and consent forms, (appendix 4), were distributed by the student consultative committee representatives for each programme and were returned to the designated drop box located in the academic office by individual students or brought to the first data collection session as was the case in the pilot phase. The first data collection session was scheduled for the beginning of October 2013 prior to the delivery of clinical preparation workshops. It collected 3-D spatial visualisation performance data and demographic information before students had any exposure to clinical preparation workshops. Follow up testing sessions were scheduled to take place in April 2014 at the end of the first year of study following a total of nine weeks clinical placement experience. The second data collection point was scheduled for October 2014 post advanced skills practical workshops and immediately prior to the first clinical placement experience of the second year of study. The final data collection was planned for March 2015 shortly after the end of the second clinical experience block (clinical placements one and two provided a total nine weeks experience). A summary of all data collection time points in relation to clinical placement timings is shown below in figure 6.1.

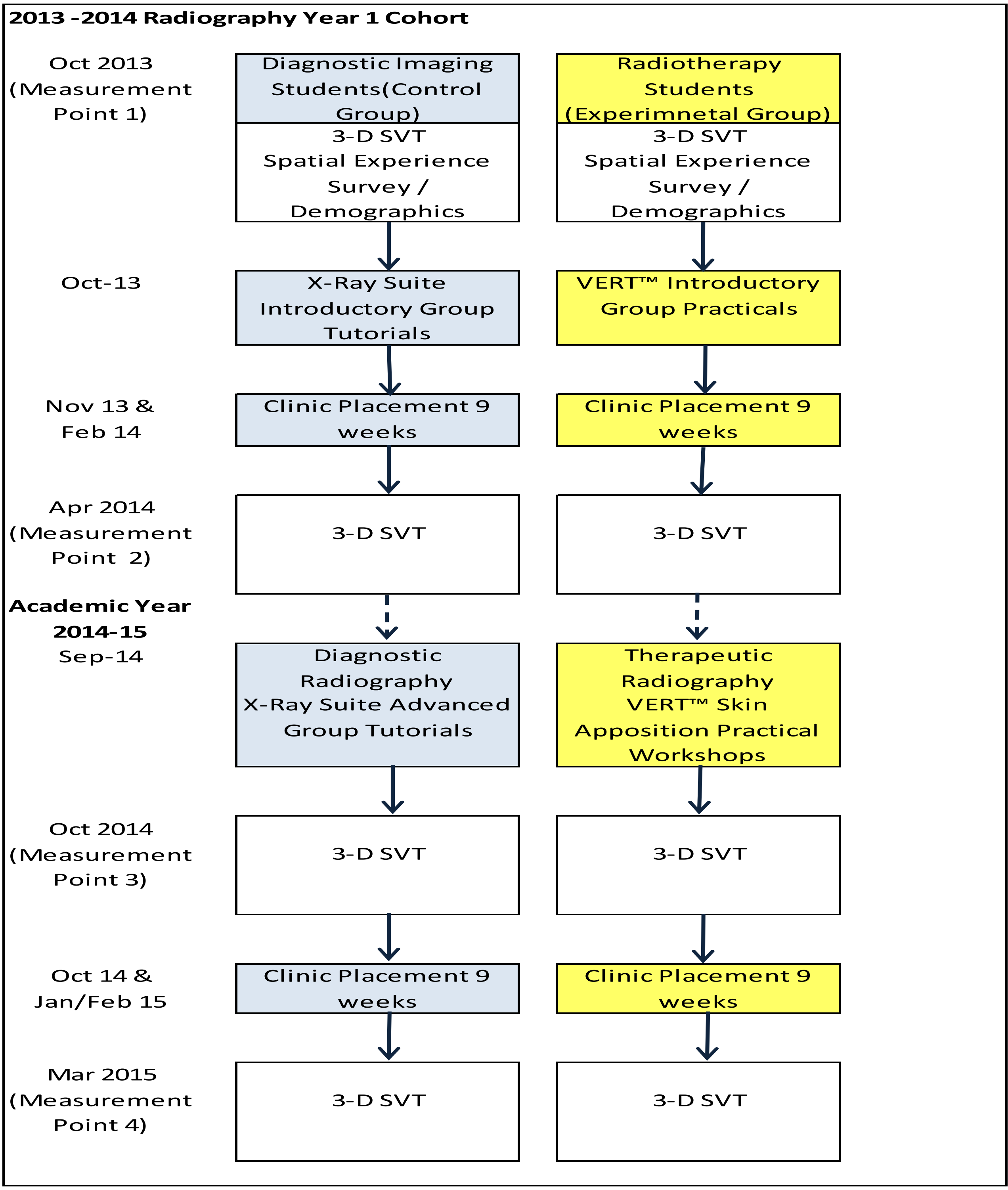


Figure 6.1: Schedule for data collection time points and their relationship to clinical practice placement

6.4 Data available for analysis

Population data for individual performance scores in the MRT and the SBST were automatically calculated as percentage values in Moodle and downloaded into a Microsoft Excel® spreadsheet. Prior to export to the Statistical Package for Social Sciences versions 22 - 24 (IBM Corp 2012-2018) and JMP Pro version 14.0 (Statistical Analysis System Institute, 2018) for descriptive and multivariate statistical analysis, the Excel spreadsheet was checked for accuracy the data was cleaned by removing all text and participant names and response codes. This was required since the spreadsheet identified each participant by name and allocated a unique Moodle identification number to them. This Moodle number changed for the same participant with each data collection time point that they attended. Therefore, following the first data collection point in October 2013, participant names were removed from the spreadsheet and replaced with a number, the first participant on the list was identified as number one; the second was two and so on. This number would be used to collate that participants data with that from subsequent data collection sessions. The list of participants and their unique identifying number were stored as a separate password protected document and used to code subsequent test result downloads prior to checking and export to SPSS for analysis. Blank cells arising as a result of a non-attempt of a test item or a no response to a demographic question were coded as “99” and given a non-submission data variable label which would be picked up in SPSS.

In addition, the incorrect answer choices in the solids test subcomponent of the 3-D SVT (discussed in chapter 5.10, p.187) were analysed to determine the number of times the egocentric distractor had been selected. The CONSORT participant flow diagram which summarises participant numbers from both radiography pathways for each of the three studies is presented in figure 6.2.

6.5 Study 4 Results

This section will begin by reporting the demographic profile of participants by age and gender and their flow through the four time points of this longitudinal study. It will then present the results from baseline testing conducted in October 2013. It will triangulate the findings by comparing them to results reported from previous spatial visualisation studies in radiography. It will then discuss the potential for identifying and grouping participants into low, intermediate and high performance bands in order to identify areas for additional visualisation support and present the findings from this exercise. It will then report the changes in test performance that occurred during the study. It will continue with an analysis of the pattern of incorrect answers in the SBST subcomponent before concluding with an analysis of the impact of missing data and a discussion relating to the overall findings.

6.5.1 Demographic profile and participant flow

The participant flow for each pathway and each data collection time point is summarised in the modified CONSORT flow chart exhibited in table 6.2.

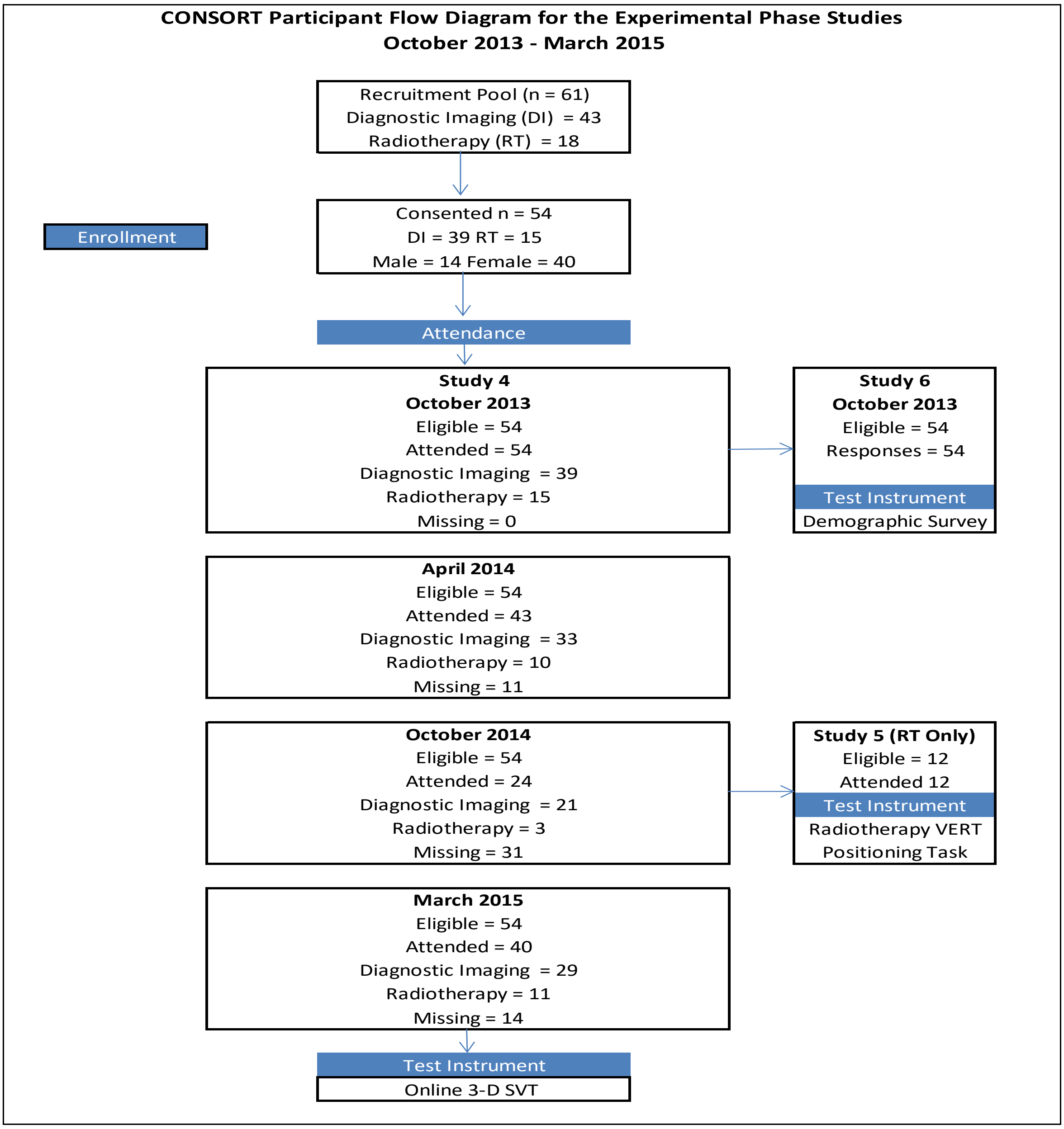


Figure 6.2: Summary of participant flow numbers and test instruments for the experimental phase studies 4 – 6

Analysis of participants by study pathway showed that 39 (90.7% response) were diagnostic imaging students and 15 (83.3%) were radiotherapy students. The breakdown of male and female participants and their age profile is summarised in table 6.2 below. From this, it can be seen that across both pathways, for the 52 participants who submitted data for age and gender, there were 13 (25%) male students. Further examination of the proportion of males in each cohort shows that 10 (25.6%) were studying diagnostic imaging while three (23%) were on the radiotherapy pathway (one female did not specify age and one participant; age 22, did not state gender).

A comparison of the age profile for both pathways showed the same mean age for male students was 26.3 (SD= 6.3) and a similar range, but the female radiotherapy students were, on average, younger than their diagnostic imaging counterparts. The mean age for radiotherapy female students was 19.2 (SD = 1.3, range 18 – 21) while the diagnostic imaging female cohort had a mean age of 22.4 (SD = 6.3, range 18 – 39). The higher number of participants from the diagnostic imaging pathway is a reflection of the level of commissioned training places which, at the time of the study, were determined by Health Education England. The number of commissions for radiotherapy training was lower due to the smaller number of radiotherapy departments within which students can gain their clinical experience compared to diagnostic imaging.

Table 6.2: Comparison of age profile by gender and study pathway

		Male	**Female**
All	n	13	39
	Mean	26.3	21.7
	SD	6.8	5.6
	Min	18	18
	Max	38	39
Diagnostic Imaging	n	10	29
	Mean	26.3	22.4
	SD	6.3	6.3
	Min	18	18
	Max	36	39
Radiotherapy	n	3	10
	Mean	26.3	19.2
	SD	10.1	1.3
	Min	20	18
	Max	38	21

6.5.2 Baseline 3-D spatial visualisation performance

A comparison of 3-D SVT performance scores for the diagnostic imaging (control) group and the radiotherapy (experimental) group is summarised in table 6.4 and shown as frequency histograms with normality plots in figures 6.3 and 6.4. From table 6.4, it can be seen that, unlike the performance observed in pilot studies one and two, where a performance differential in favour of diagnostic imaging students was observed throughout, at the commencement of this study, the radiotherapy study cohort began their programme of education with a slightly higher performance.

Table 6.3: Comparison of baseline performance for both pathways showing a performance advantage for the radiotherapy group

	October 2013	
	Diagnostic Imaging	**Radiotherapy**
n	39	15
Mean (%)	36.5	47.4
SD	15.9	18.5
Minimum (%)	0	21
Maximum (%)	75	79
Percentiles 25	25	29
50	37	46
75	46	58

Analysis and comparison of individual performance for each pathway in each subcomponent of mental rotation (representing patient position) and the solids tests (cross sectional perception, equivalent to visualising beam path) is shown in figure 6.5 For the control group of diagnostic imaging students and figure 6.6 for the experimental group of radiotherapy students.

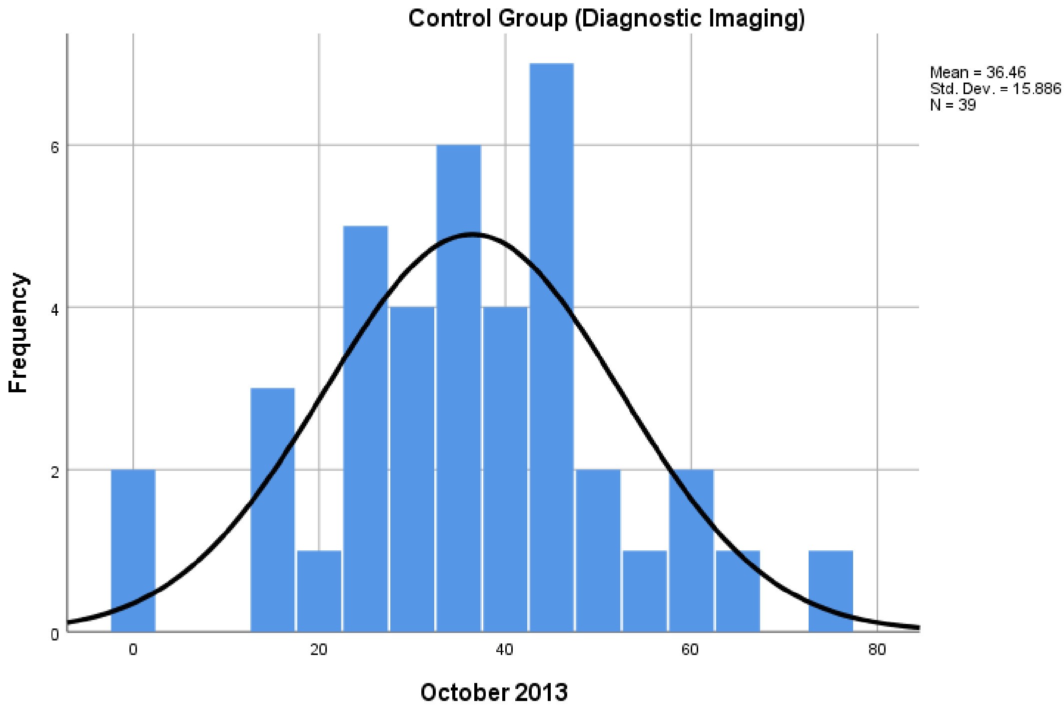


Figure 6.3: Histogram and normal distribution plot for control group performance

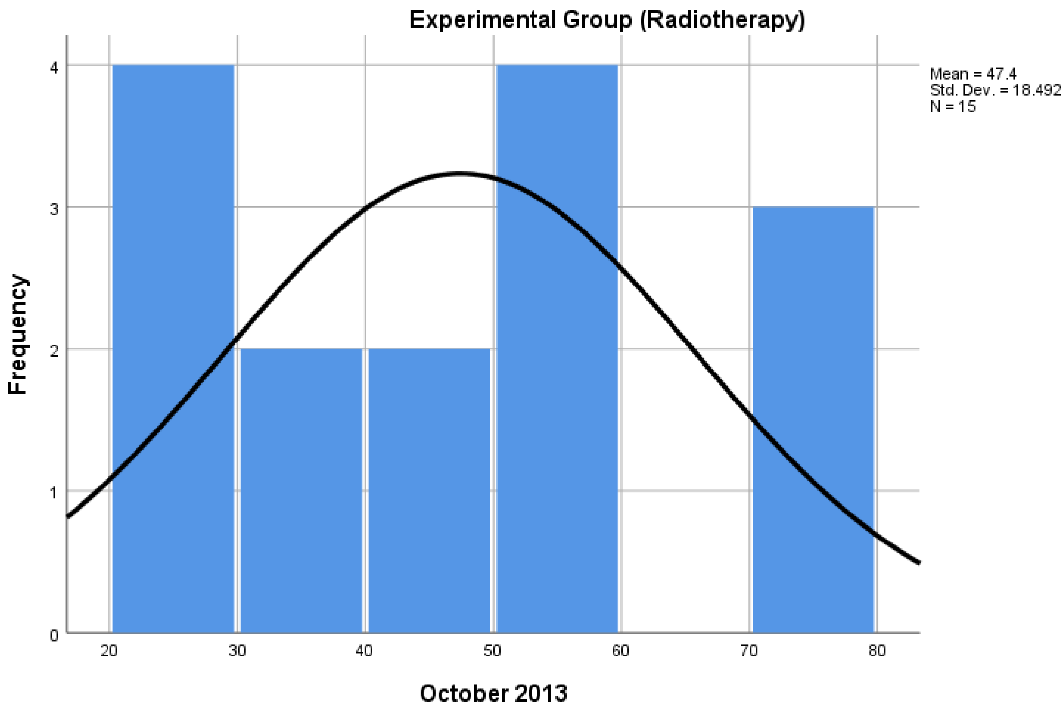


Figure 6.4: Histogram and normal distribution plot for experimental group performance

6.5.3 Baseline test subcomponent performance and comparison

Additional analysis of the subcomponent scores for the MRT and SBST between the two study groups is shown in figures 6.5 for mental rotation and figure 6.6 for SBST.

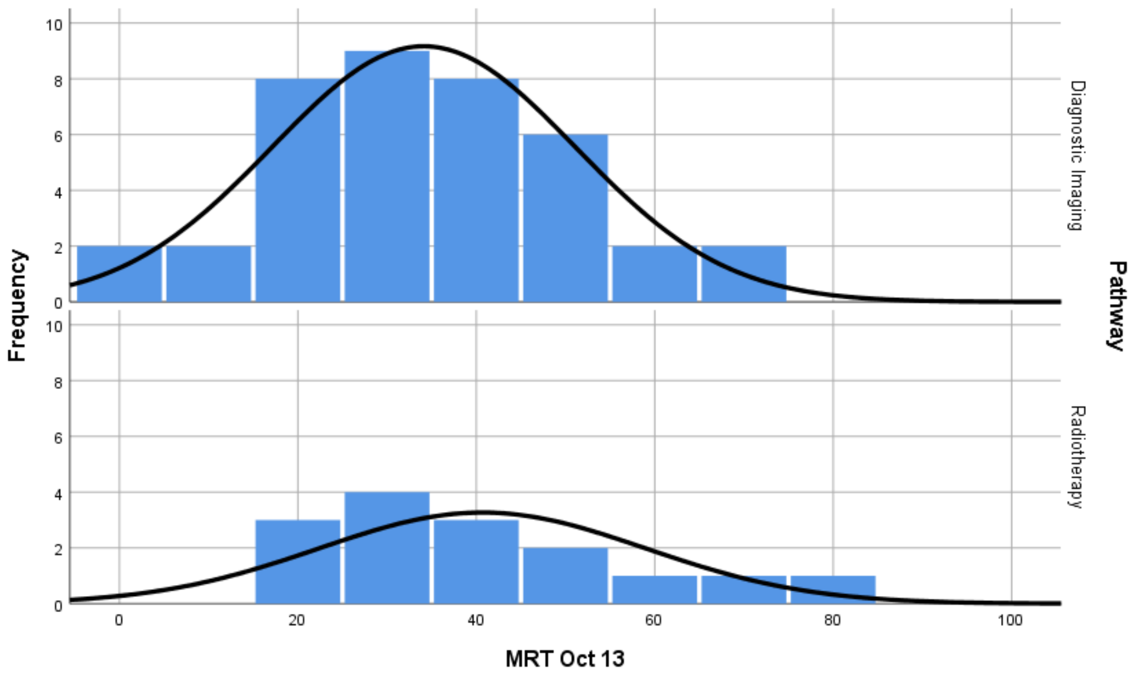


Figure 6.5: Distribution for mental rotation performance in October 2013 by programme pathway

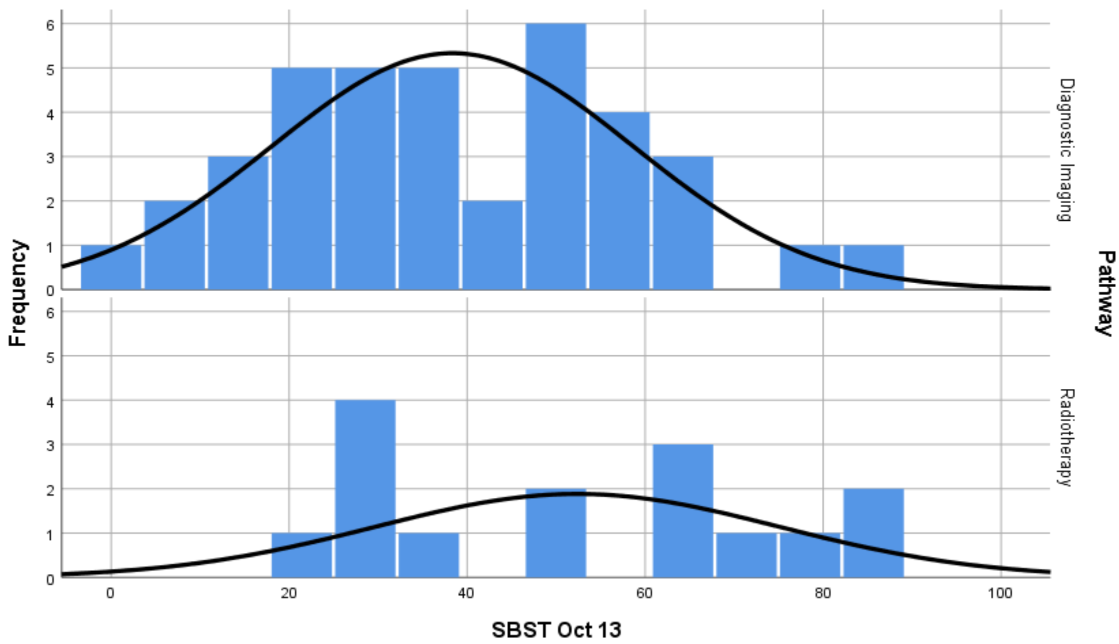


Figure 6.6: Distribution for performance in the October 2013 SBST by programme pathway

Visual checking of the histograms in figures 6.5 and 6.6 would indicate that, for both pathways, performance in the SBST subcomponent was better than in the MRT and that the performance of the radiotherapy participants is higher in both subcomponents compared to that of the diagnostic imaging group. This is confirmed in table 6.5 which compares the number of correct answers selected by both pathways in both subcomponents. Of note is the small performance difference in favour of radiotherapy female participants compared to both male and female diagnostic imaging participants.

Table 6.4: Comparison of the number of correct answer selections by subcomponent, gender and programme pathway

	Diagnostic imaging					
	MRT			**SBST**		
	Experimental Phase October 2013					
	(10 Items)			**(14 Items)**		
	Male	**Female**	**Combined**	**Male**	**Female**	**Combined**
n	10	29	39	10	29	39
Mean	2.6	3.7	3.4	4.3	5.6	5.4
SD	1.5	1.7	1.7	3.2	2.6	2.8
Min	0	0	0	0	2	0
Max	5	7	7	9	12	12

	Radiotherapy					
	MRT			**SBST**		
	Experimental Phase October 2013					
	(10 Items)			**(14 Items)**		
	Male	Female	Combined	Male	Female	Combined
n	4	11	15	4	11	15
Mean	4.8	3.8	4.1	9	6.7	7.3
SD	2.8	4.1	1.8	4.1	2.8	3.2
Min	2	2	2	3	4	3
Max	8	7	8	12	12	12

6.5.4 Triangulation of baseline results with other studies

The results of the spatial visualisation studies from study one of the pilot phase and the baseline measurements from study four of the experimental phase of this programme of research provided an insight into the 3-D spatial visualisation skills of volunteers from the 2011-12 and 2013-14 year one cohorts of diagnostic imaging and radiotherapy students in one HEI. The 3-D SVT developed for these studies employed a representative sample of test objects selected from the Vandenberg and Kuse MRT and the SBST and derived an overall performance score. This section will focus on comparing performance in the mental rotation subcomponent for the cohorts studied in this programme of research (table 6.6) with the findings of similar previous studies conducted by Appleyard and Coleman (2010) and Green and Appleyard (2011) and summarised in table 6.7.

Table 6.5: Number of correct answers for the MRT subcomponent (radiotherapy cohorts)

	MRT Pilot Phase (Study 1 Paper based, 12 Items)			MRT Experimental Phase (Study 4 Oct 2013, 10 Items)		
	Male	**Female**	**Combined**	**Male**	**Female**	**Combined**
n	3	11	14	4	11	15
Mean	5.0	4.0	4.2	4.8	3.8	4.0
SD	2.0	2.0	2.1	2.8	1.5	1.8
Min	3	1	1	2	2	2
Max	7	7	7	8	7	8

The studies conducted by Appleyard and Coleman and Green and Appleyard with first and second year students, were designed to determine the effect of the three visualisation modes available in VERT™ on performance in a radiotherapy positioning task and the impact of spatial visualisation on this performance. Spatial visualisation skill was measured with the 24 item Vandenberg and Kuse MRT. It can be seen that, in these studies, spatial visualisation performance of males and females combined is proportionally higher in five of the groups when compared to the performance of the participants in this programme of research. The exception was Green and Appleyard's 2-D group. A possible explanation for this higher performance may be the inclusion of second year students whose spatial visualisation may have improved during their studies.

Table 6.6: Comparative summary for number of correct answers for MRT performance reported by Appleyard & Coleman 2010 and Green & Appleyard 2011

<table>
<tr><th colspan="7">Appleyard & Coleman 2010 (24 Item MRT)</th></tr>
<tr><th colspan="4"></th><th colspan="3">VERT Operating Mode</th></tr>
<tr><th></th><th>Male</th><th>Female</th><th>Combined</th><th>3-D Tracking on</th><th>3-D Tracking Off</th><th>2-D</th></tr>
<tr><td>n</td><td>28</td><td>75</td><td>103</td><td>M:11
F:25</td><td>M:8
F:27</td><td>M:9
F:23</td></tr>
<tr><td>Mean</td><td colspan="3" rowspan="4">Not Reported</td><td>10.8</td><td>11.5</td><td>9.0</td></tr>
<tr><td>SD</td><td>4.3</td><td>5.1</td><td>5.6</td></tr>
<tr><td>Min</td><td colspan="3" rowspan="2">Not reported</td></tr>
<tr><td>Max</td></tr>
<tr><th colspan="7">Green & Appleyard 2011 (24 item MRT)</th></tr>
<tr><th colspan="4"></th><th colspan="3">VERT Operating Mode</th></tr>
<tr><th></th><th>Male</th><th>Female</th><th>Combined</th><th>3-D Tracking on</th><th>3-D Tracking Off</th><th>2-D</th></tr>
<tr><td>n</td><td>11</td><td>33</td><td>44</td><td>M:5
F:8</td><td>M:3
F:12</td><td>M:3
F:13</td></tr>
<tr><td>Mean</td><td colspan="3" rowspan="4">Not Reported</td><td>11.3</td><td>11.5</td><td>7</td></tr>
<tr><td>SD</td><td>4.7</td><td>5.8</td><td>5.1</td></tr>
<tr><td>Min</td><td colspan="3" rowspan="2">Not Reported</td></tr>
<tr><td>Max</td></tr>
</table>

The performance of the diagnostic imaging students who participated in the pilot and experimental phases of this programme of research is summarised in table 6.8 and shows a mean score for correct answers ranging from 2.6 (from 10 items), equivalent to a percentage score of 26% to 5.5 (from 12 items), equivalent to a percentage score of 45.8%. Studies reporting the measurement of mental rotation skills in diagnostic imaging are limited. However a study conducted by Duce et al. (2016, p.1162) which involved 33 novice ultrasonographers employed the 24 item Vandenberg & Kuse MRT as part of a bank of five “spatial ability” tests. The cohort was composed of 18 males (54.5%) and 15 females (45.5%) with a combined mean age of 21.6 years (SD = 5.2) and their combined MRT performance score mean was reported as 10.0 (SD = 6.3), an equivalent percentage score of 41.7% which is similar to the highest performance recorded for this programme of research.

Table 6.7: Number of correct answers for the MRT subcomponent (diagnostic imaging cohorts)

	MRT Pilot Phase (Study 1 Paper based, 12 Items)			MRT Experimental Phase (Study 4 Oct 2013, 10 Items)		
	Male	**Female**	**Combined**	**Male**	**Female**	**Combined**
n	4	8	12	10	29	39
Mean	5.5	4.5	4.8	2.6	3.7	3.4
SD	2.4	1.8	1.9	1.5	1.7	1.7
Min	3	2	2	0	0	0
Max	8	7	8	5	7	7

Triangulation of findings for performance in the SBST subcomponent of the 3-D SVT will focus on the number of times the egocentric distractor (foil) was selected and will be covered in section 6.58, p. 218.

6.5.5 Spatial visualisation grouping by overall test performance

One of the recommendations from the DoH (England) VERT™ evaluation project was the assessment of students' inherent spatial ability to assist identification of individuals who are likely to benefit most from experience in VERT™ (Appleyard & Coleman, 2010, p. 33). The results of study four (section 6.5.2, p. 205) demonstrated that student's baseline 3-D spatial visualisation skill can be determined at the start of their radiography education. With regard to individual benefit, the research evidence base to date, has shown that its use can result in the increased understanding of radiotherapy concepts and confidence in their application for the majority of students as discussed in chapter 2.8.1, p. 72. Research aim three sought to determine the longer term potential of VERT™ in the development of 3-D spatial visualisation skill. So are there groups of students who would derive greater benefit from a more individualised and focused approach to concept visualisation? The identification of the relative level of 3-D visualisation skill for each student would aid the development of visualisation activities which matched the three subcomponents of mental rotation, spatial perception and spatial visualisation. This would result in the development of additional, bespoke activities, more closely aligned to individual development needs. It would also move the use of VERT™ beyond the current principle of one size fits all approach to tutorials and workshops which are based predominantly on generic learning outcomes.

The principle of grouping individuals by their performance score requires a clear understanding of the boundaries between visualisation skill levels. The literature reporting the measurement of spatial visualisation, however, lacks a clear consensus relating to what is considered to be high, average or low skill as indicated by test performance. In a study using a combination of paper folding, cube comparison, form board and card rotation tests with 60 psychology students, Kozhevnikov, Hegarty, and Mayer (2002, p. 55) used a composite score from all tests to categorise, what they termed, "spatial ability performance". They identified individuals as having high ability if their score was in the top 25% of the distribution and low spatial visualisation skill if it lay within the bottom 25% average. The remainder who lay in the middle 50% were classed as intermediate. It should be noted that there was no report of the demographic profile of the cohort in relation to gender or age. In a later study of 59 psychology students, Cohen and Hegarty (2007, p. 183) defined low spatial "ability" performance as a score lying in the lower third of the distribution for results gained in the Vandenberg and Kuse MRT. Participants who gained a score in the upper third would be classified as having high spatial ability. However an important consideration in this interpretation is that an unspecified scoring convention was employed in which a maximum score of 80 could be achieved. During the course of this programme of research, Duce et al., (2016, p. 1164), recorded the performance of 33 trainee sonographers across the Vandenberg and Kuse MRT and three tests from the perceptual reasoning subset of the Wechsler Adult Intelligence Scale (version IV). Performance in each test was analysed separately by raw score, with individuals being grouped by standard deviations into low (< - 1 SD), intermediate (± 1 SD) and high (> + 1 SD) skill for each test rather than overall visualisation performance.

Based on these different grouping methods and recognising that spatial visualisation is not a unitary construct, the decision taken for this programme of research was to use the convention employed by Kozhevnikov, Hegarty, and Mayer. This was based on their use of a composite score and driven by the need to gain an understanding of performance across all the components of spatial visualisation skill. Therefore scores equal to, or above, the 75-percentile level of the cohort distribution would indicate highly developed 3-D spatial visualisation skill while those at, or below, the 25-percentile level would be classified as having less well-developed skill. The application of these groupings to individual performance scores achieved at baseline in October 2013 is demonstrated in table 6.8 and summarised in relation to each pathway in table 6.9.It can be seen from table 6.6, that while there are proportionally fewer low performers in the radiotherapy group (n = 2, 13.3%) compared to the diagnostic imaging group (n = 11, 28.2%), there are also more high performers in the radiotherapy group (n=3, 20%) compared to the diagnostic imaging group (n= 2, 5.1%). However it is acknowledged that the small number of participants in each of the skill groupings is recognised and acknowledged. To gain an understanding of the implications of these findings to the general population of learners in both diagnostic imaging and radiotherapy would require a larger collaborative study.

Table 6.8: 3-D spatial visualisation skill banding for each participant at the start of their programme of study

	Control Group			**Experimental Group**	
ID	**Baseline Score Banding Oct 13**	**ID**	**Baseline Score Banding Oct 13**	**ID**	**Baseline Score Banding Oct 13**
1	Low	26	Inter	5	Inter
2	Low	27	Inter	8	Inter
3	Inter	28	Inter	13	High
4	Low	29	Inter	16	Inter
6	Inter	31	Inter	21	Inter
7	Inter	33	Inter	30	Inter
9	Low	34	High	32	High
10	Inter	36	Low	35	Inter
11	Inter	37	Inter	38	Low
12	Inter	41	Low	39	Inter
14	Low	42	Inter	40	Low
15	Inter	43	Low	45	Inter
17	Inter	44	Inter	46	High
18	Inter	47	Inter	51	Inter
19	Inter	48	High	53	Inter
20	Low	49	Inter		
22	Inter	50	Inter		
23	Inter	52	Inter		
24	Low	54	Inter		
25	Low				

Table 6.9: Comparison of 3-D spatial visualisation skill grouping for diagnostic imaging and radiotherapy students in October 2013

Control Group Baseline Oct 13			Experimental Group Baseline Oct 13		
Group	n = 39	%	Group	n = 15	%
Low	11	28	Low	2	13
Inter	26	67	Inter	10	67
High	2	5	High	3	20

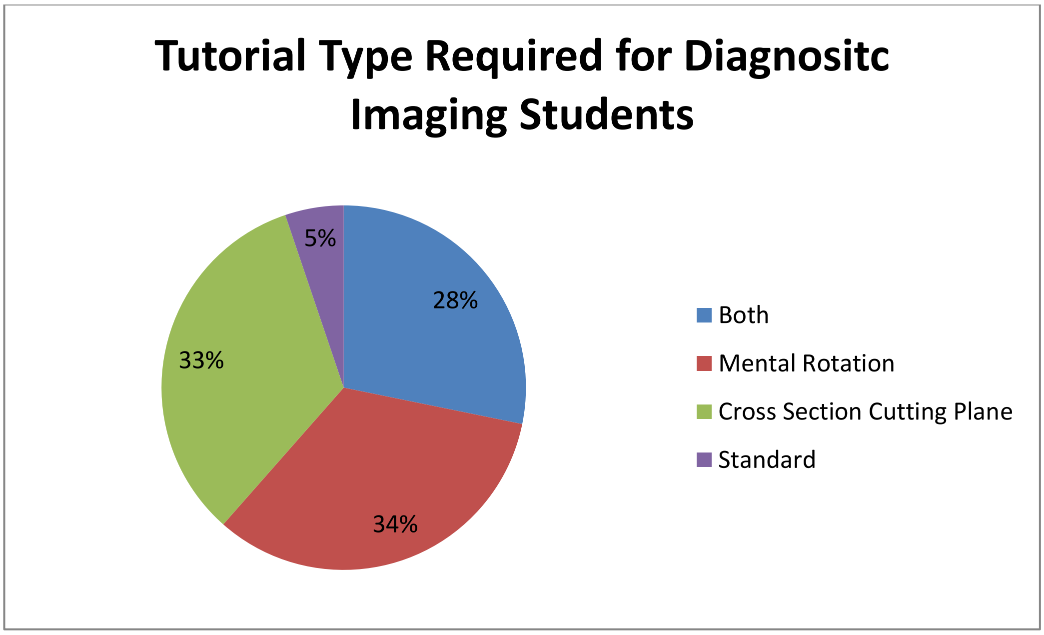
The grouping of individual learners by overall performance score into high, intermediate and low categories may provide a general indicator of where one learner sits in relation to others in the same cohort. It would also provide an additional indicator of the degree of any growth trajectory beyond an analysis of percentage gain in performance score. However any interpretation based on grouping by overall performance score alone should be applied with caution since it may contribute to the risk of increased stereotype threat for some students. Individuals with lower 3-D visualisation skills at baseline may perceive their performance and grouping as "I am not good at this therefore there is no point in trying". If this is the case then there is a risk that they may lose confidence and become demotivated.

6.5.6 Grouping by performance in visualisation subcomponent tests

To reduce the risk of demotivation and stereotype threat and to provide a deeper understanding of individual development needs an alternative grouping is proposed. Grouping by performance in the subcomponent tests of mental rotation and perception and visualisation (cross sections) would more clearly identify which components of the patient positioning and beam alignment processes would benefit from additional support. As the VERT™ platform can model entire patient pathways from identifying anatomical structure outlines on CT data sets to 3-D structure modelling and treatment delivery with radiation dose overlays, it offers the opportunity for a deeper focus on all aspects of the external beam radiotherapy pathway. While the use of the platform has a predominantly radiotherapy focus, its organ and beam's eye view modeling could also support the visualisation of patient positioning and X-ray tube alignment for diagnostic imaging students. Therefore the subcomponent performance scores were analysed to determine if they could provide an indication of specific areas of visualisation which may benefit from additional, focused, tutorial support. These results, for each participant, are presented in table 6.10 and as a summary of activities required for each pathway in figure 6.7. These show that, in the control group of diagnostic imaging students, 11 (28%) of the 39 participants would benefit from exercises in both mental rotation and beam's eye view cross sectional activities, while 13 (33%) would benefit from mental rotation or beam's eye view activities. In the experimental group of radiotherapy students, 2 out of 15 (13%) participants could benefit from activities in both components while 5 (33%) would benefit from one or the other. It is interesting to note from figure 6.7 that, across both pathways, the proportion of participants requiring support with either mental rotation (34%) or cross section cutting plane activities (33%) is similar. This would suggest that the standard practical workshops and clinical tutorials would benefit from additional focussed content to support the visualisation development needs for first year students.

Table 6.10: Recommended tutorial support required following baseline testing

ID	Tutorial Support Needed at Baseline (Diagnostic Imaging)	ID	Tutorial Support Needed at Baseline (Diagnostic Imaging)	ID	Tutorial Support Needed at Baseline (Radiotherapy)
1	Both	26	Cross Section	5	Mental Rotation
2	Both	27	Mental Rotation	8	Mental Rotation
3	Cross Section	28	Cross Section	13	Standard
4	Both	29	Cross Section	16	Cross Section
6	Cross Section	31	Cross Section	21	Mental Rotation
7	Mental Rotation	33	Mental Rotation	30	Cross Section
9	Both	34	Standard	32	Standard
10	Mental Rotation	36	Both	35	Cross Section
11	Mental Rotation	37	Mental Rotation	38	Both
12	Mental Rotation	41	Both	39	Cross Section
14	Both	42	Mental Rotation	40	Both
15	Cross Section	43	Both	45	Cross Section
17	Mental Rotation	44	Cross Section	46	Standard
18	Cross Section	47	Cross Section	51	Mental Rotation
19	Mental Rotation	48	Standard	53	Mental Rotation
20	Both	49	Mental Rotation		
22	Cross Section	50	Mental Rotation		
23	Cross Section	52	Mental Rotation		
24	Both	54	Cross Section		
25	Both				



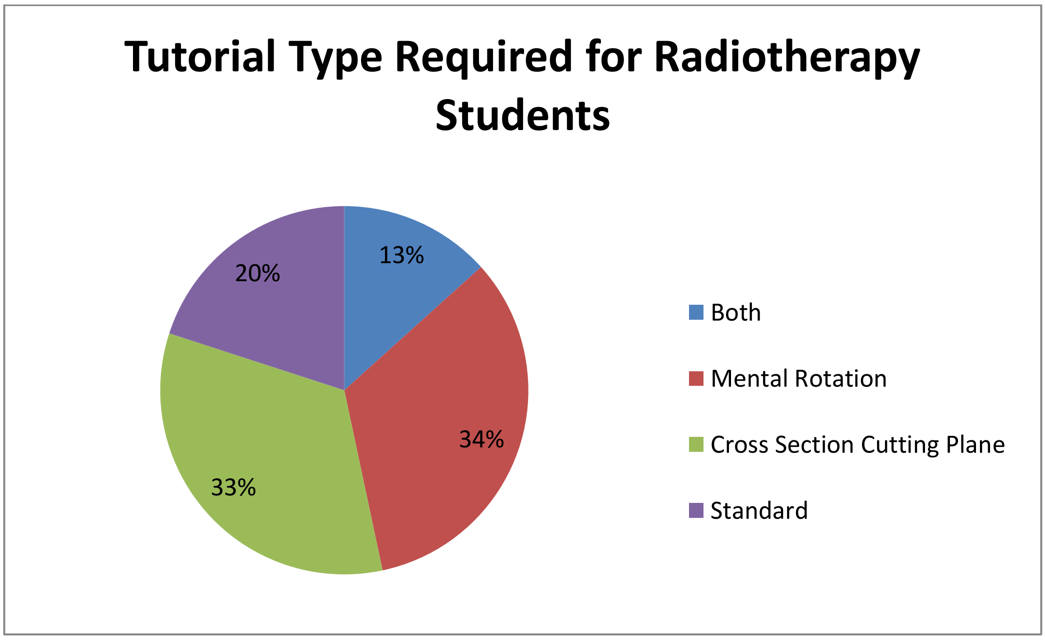


Figure 6.7: Suggested tutorial support required for at the commencement of their radiography education

6.5.7 Analysis of performance change over time

The review of spatial visualisation literature discussed identified that performance in spatial visualisation tasks may change over time. This was observed in study two of the pilot phase, reported in chapter 5.9.1, p.185. Due to the disparity in the number of participants in each group at each time point, 20 students (37%) attended all data collection time points, each pair Student's t test values were calculated for each group. For the diagnostic imaging group, this showed a statistically significant performance difference between the April 14 and March 15 test (p = 0.026) and the October 13 and March 15 test (p = 0.036), but not confirmed by Tukey-Kramer HSD. For the radiotherapy group the each pair Student's t test and all pairs Tukey-Kramer show no statistically significant difference. Both groups demonstrated an improvement in performance at the end of the study compared with baseline, as demonstrated in table 6.11. The number and percentage of participants in each skill banding at the start and end of the study is shown in table 6.12.

Table 6.11: Comparison of descriptive statistics for performance across for all time points for diagnostic imaging (DI) and radiotherapy (RT) pathways

Time Point	October 2013		April 2014		October 2014		March 2015	
Pathway	DI	RT	DI	RT	DI	RT	DI	RT
n	39	15	33	10	21	3	29	11
Mean	36.5	47.4	35.5	51.3	41.6	59. 7	46.2	49.2
SD	15.9	18.5	19.2	15.9	20.2	21.2	20.7	20.0
Minimum	0	21	4	29	8	37	4	4
Maximum	75	79	71	71	88	79	92	79
Percentiles								
25	25	29	19	33	26	37	33	37
50	37	46	37	52	42	63	42	54
75	46	58	50	67	53	.	63	63

It should be noted that the measurement time point in October 2014 coincided with the final academic week prior to the commencement of the first clinical practice placement week of year two for both study groups. During the data collection session some of the radiotherapy participants alerted the researcher to an anomaly in the presentation of the solids cutting test which resulted in the presentation of some objects being duplicated and therefore being viewed more than once. The error was caused following an upgrade to Moodle and Moodle Quiz prior to the start of the academic year but was not evident when the researcher tested the module prior to deployment. The fault did not affect the control group. Following rectification of the fault, the module was reopened for a period of one week (which coincided with the first clinical placement week) to enable radiotherapy students to complete the test remotely. Just three students attempted the test in its rerun, gaining scores of 37%, 63% and 79% respectively. Given the low number of participants and disparity in the scores gained, no further analysis or comparison of performance could be made for this time point.

Table 6.12: Performance grouping at the start and end of study four

Control Group (Diagnostic Imaging)					
ID	**Baseline Group October 2013**	**End of Study Group March 2014**	**ID**	**Baseline Group October 2013**	**End of Study Group March 2014**
1	Low	Low	26	Int	Int
2	Low	Low	27	Int	High
3	Low	Low	28	Int	Int
6	Int	Int	31	Int	Int
7	Int	High	36	Low	Int
11	Int	High	37	Int	Int
12	Int	High	43	Low	Low
15	Int	Int	44	Int	High
17	Int	Int	47	Int	Int
18	Int	Int	48	High	High
19	Int	High	49	Int	Int
20	Low	Int	50	Int	High
22	Int	Int	52	Int	Int
24	Low	Int	54	Int	Int
25	Low	Int			
Experimental Group (Radiotherapy)					
ID	**Baseline Group October 2013**	**End of Study Group March 2014**	**ID**	**Baseline Group October 2013**	**End of Study Group March 2014**
5	Int	Int	38	Low	Int
13	High	High	40	Low	Int
16	Int	Int	45	Int	Int
21	Int	Int	46	High	Low
30	Int	High	51	Int	Int
32	High	High			

Table 6.13: Number of participants changing spatial visualisation group between October 2013 and March 2015

Change from Baseline to End of Study		
SVS Banding	**Diagnostic Imaging (Control Group)**	**Radiotherapy (Experimental Group)**
Same	18 (62%)	7 (64%)
Higher	11 (38%)	3 (27%)
Lower	0	1 (9%)

The mean performance scores for both groups showed improvement by March 2015 compared to baseline performance in October 2013 although the increase observed for the diagnostic imaging group was greater than that seen in the radiotherapy group. However an examination of table 6.13 shows that, even with an increase in mean performance score, 18 (62.1%) of diagnostic imaging students remained in the same stratification band at the end of the study. The proportion of radiotherapy students remaining in the same band was similar at 63.6%. Proportionally more diagnostic imaging students moved up in banding level compared to radiotherapy. Of the 11 (37.9%) who changed, four students (36.4%) moved from low to intermediate and seven (63.6%) moved from intermediate to high. In the radiotherapy cohort, three students changed banding level, with two (66.7%) improving from low to intermediate and one (33.3%) moving from the low to intermediate band. These results are demonstrated in figures 6.8 and 6.9 below. The two figures show that while the intermediate and high groups demonstrated some improvement, there were still diagnostic imaging participants who remained in the lowest band (n= 4, 13.8%). While the reason for this is unexplained, they were successful in their academic and clinical studies, so it is possible that they did not understand the requirements of the test.

Examination of the radiotherapy cohort performance showed that the overall maximum mark remained unchanged. However the maximum and minimum marks for the intermediate band had increased by March 2015 compared to baseline. It should be noted that one student in the radiotherapy group moved from a high banding at baseline to low in March 2015. This was attributed to a non-submission of the solids test subcomponent, the cause could not be determined.

6.5.8 Analysis of test subcomponent performance

Visual checking of the histograms, figures 6.8 and 6.9, shows that the performance for both diagnostic imaging and radiotherapy in the SBST subcomponent was better than their performance in the MRT. It can also be seen that while performance in the MRT is similar for both groups, the radiotherapy groups' performance in the SBST is marginally better.

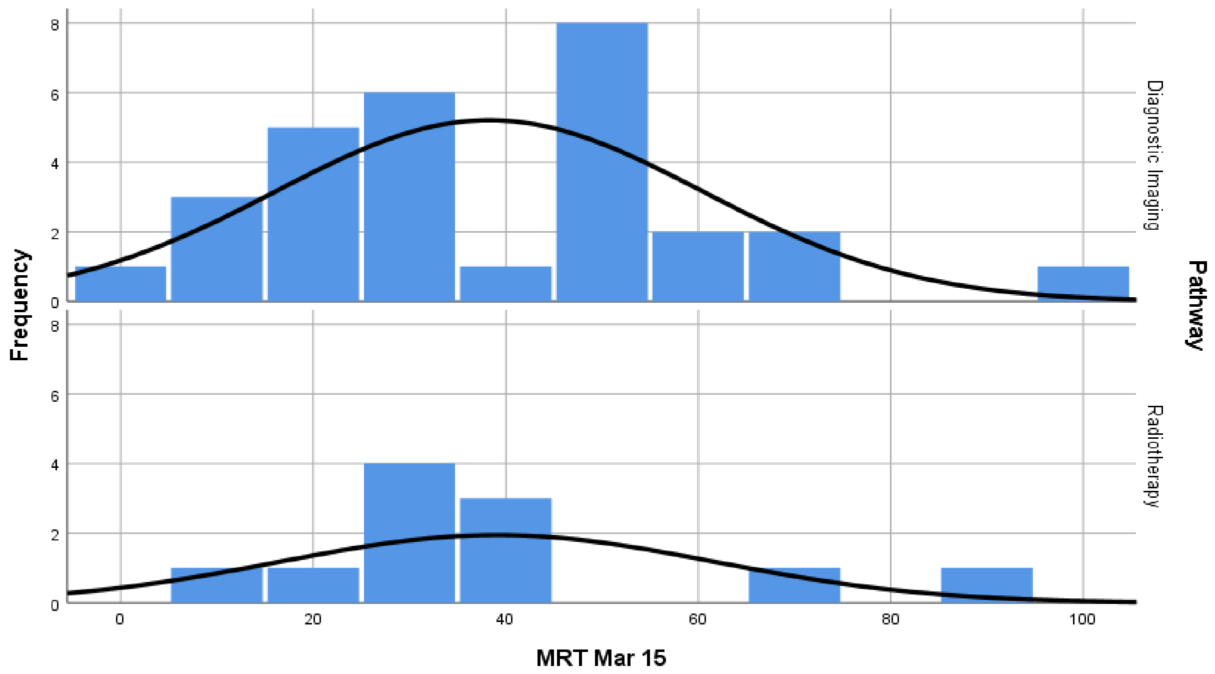


Figure 6.8 Mental rotation test subcomponent score distribution by pathway for March 2015

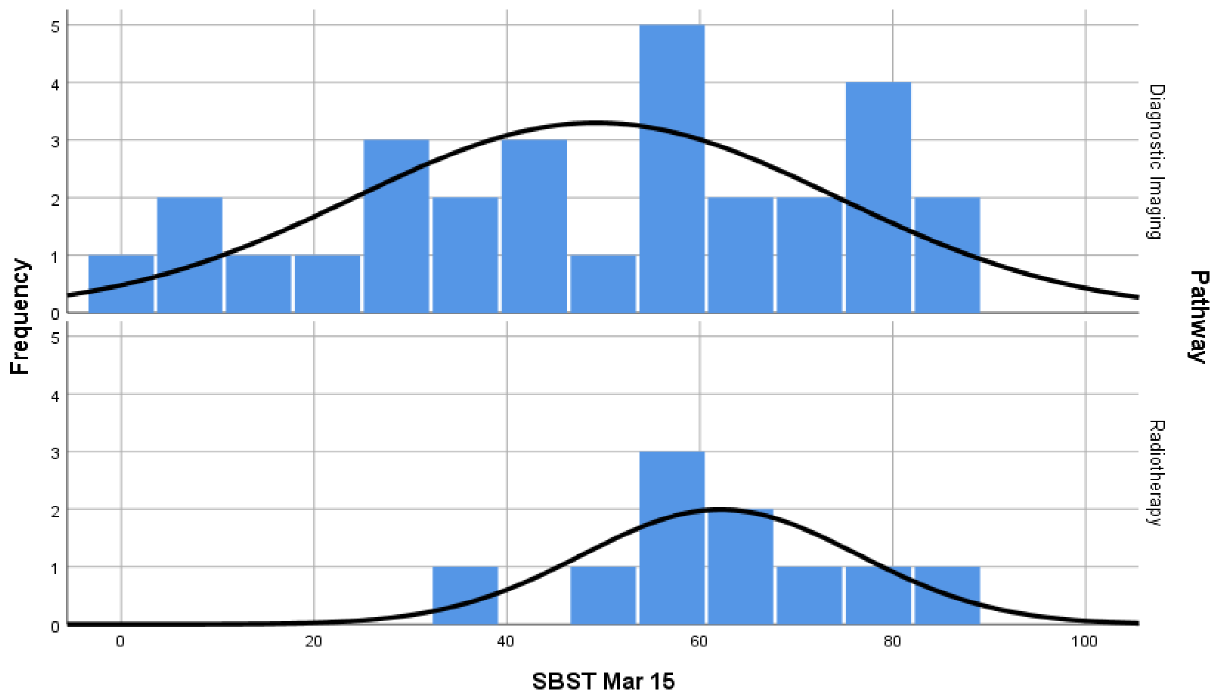


Figure 6.9: SBST subcomponent score distribution by pathway for March 2015

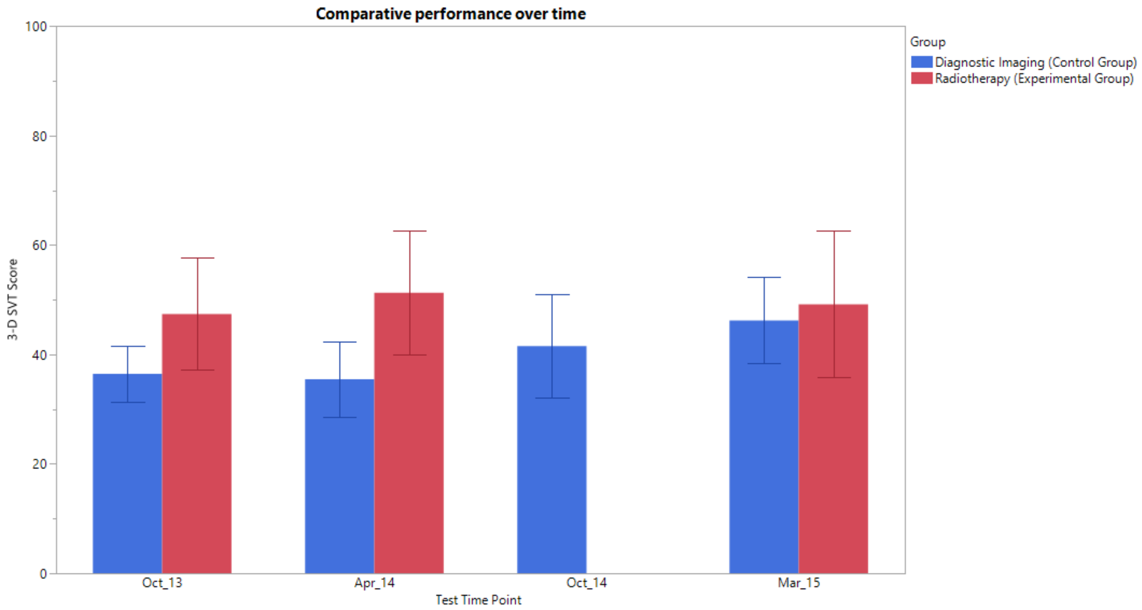


Figure 6.10: Boxplot comparing performance over all time points by programme pathway

The boxplot (figure 6.10) demonstrates that the radiotherapy group, on average, performed better than the diagnostic group across all testing phases. The spread of scores indicates that while the mean performance for the diagnostic imaging group was lower, there is no statistical difference between these two populations. The test results for the radiotherapy participants in October 2014 were removed from the analysis because only three students attempted the test in its rerun (following a software fault at first attempt), gaining scores of 37%, 63% and 79% respectively.

The growth trajectory for both groups is shown graphically in figure 6.11 and demonstrates a steeper improvement in performance is seen for the diagnostic imaging (control) group compared to that of the radiotherapy group. This may be due, in part, to diagnostic students having access to a real X-ray room on campus, which reinforces the hands on patient positioning skills developed in the clinical setting. Another possibility is that, in addition to the above, at the time of this study diagnostic imaging students had an introduction to image interpretation academic module during their second year. For radiotherapy students, their experience in pattern recognition and X-ray image review for IGRT was gained solely via ad-hoc opportunities in the radiotherapy clinical environment.

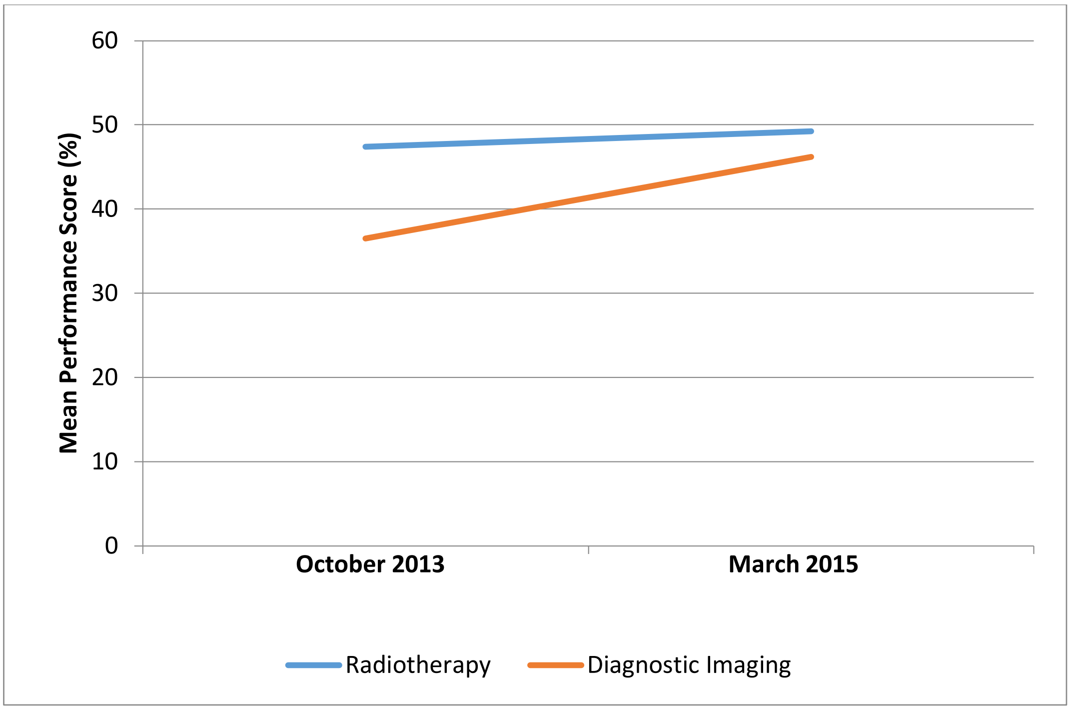


Figure 6.11: Performance growth from October 2013 to March 2015 for diagnostic imaging and radiotherapy students

Further analysis of the total mean performance score for the control group of diagnostic imaging participants at the beginning and the end of the study (figure 6.12) shows a total mean score of 40.6%. A comparison of the performance in the final test in March 2015 with baseline in October 2015 indicates a significant statistical improvement from baseline to the end of the study.

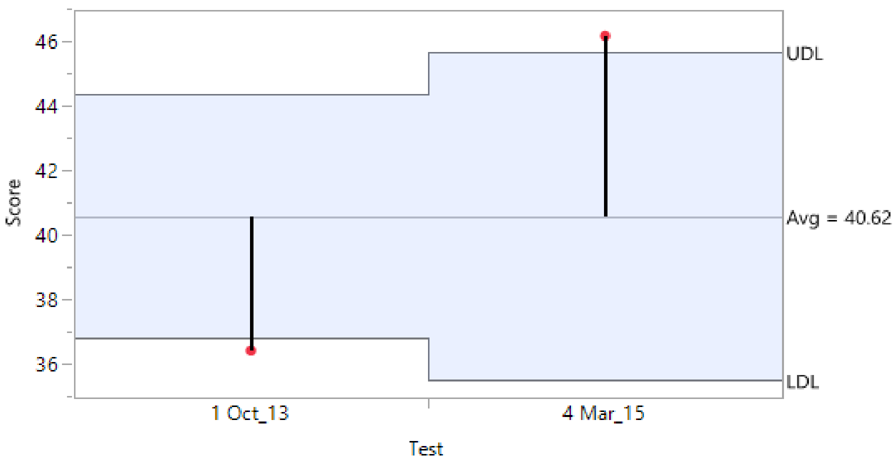


Figure 6.12: Analysis of means plot for diagnostic imaging students

In contrast, figure 6.13 shows that on average, across both time points, the experimental group of radiotherapy students scored 48.2%. While they performed better in March 2015 compared to baseline, this increase was not statistically significant.

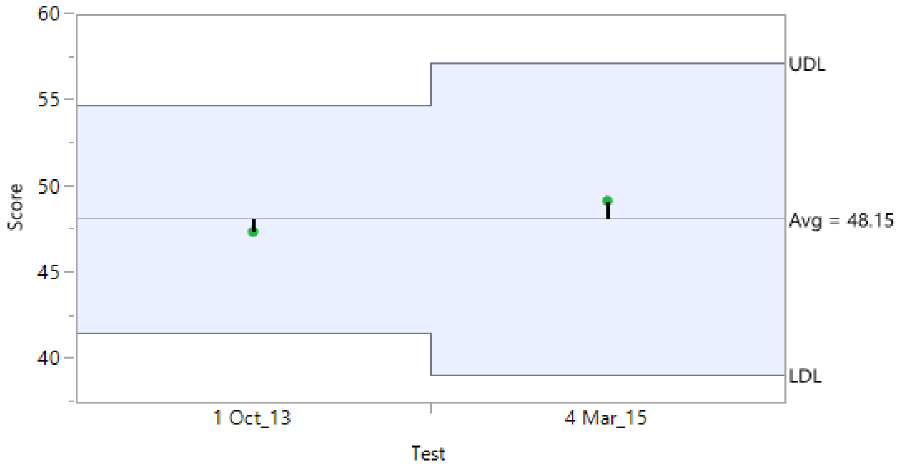


Figure 6.13: Analysis of means plot for radiotherapy students

Analysis of scores for the mental rotation and solids test subcomponents, shown in box plot format (figure 6.14) demonstrates that both groups performed better in the solids tests compared to the mental rotation tests.

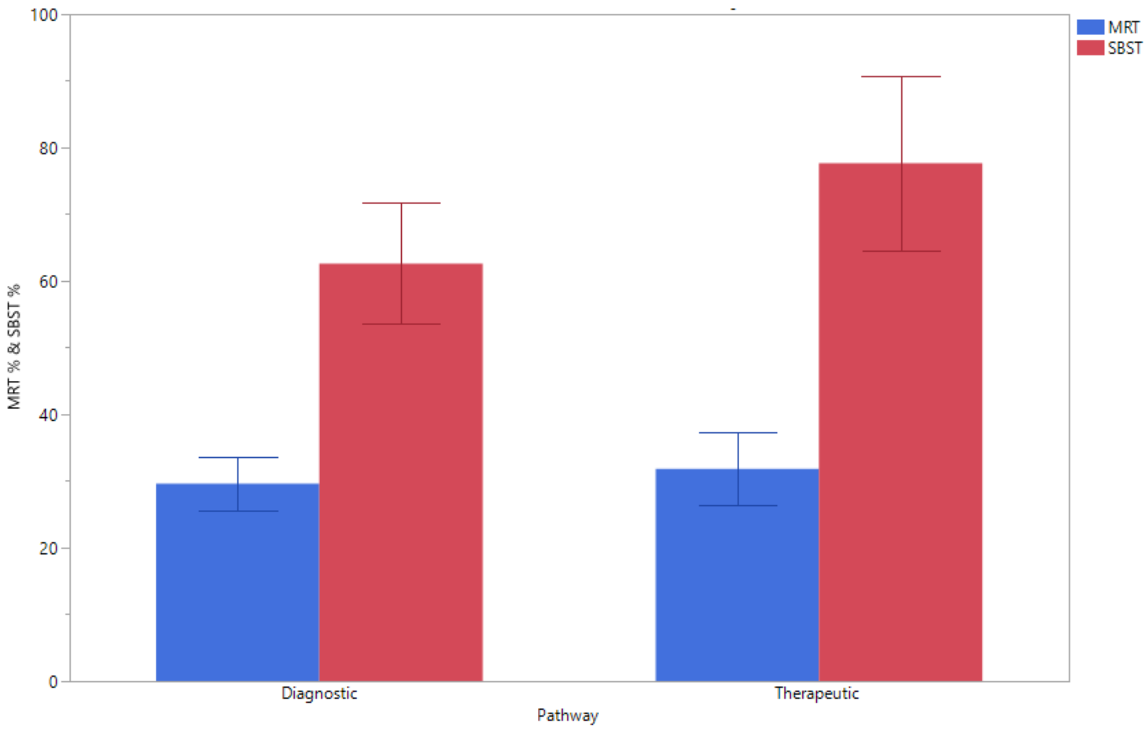


Figure 6.14: Box plot for subcomponent test performance score by study programme pathway

6.5.9 Santa Barbara Solids Test item analysis

Across all the data collection time points in study four; there were a total of 158 SBST data sets available for analysis. This analysis showed from a total of 1064 incorrect answers 585 (55%) were the egocentric distractor (foil). Triangulation of these results with those from previous studies that employed the SBST in its paper format showed that in a cohort of 60 Psychology students conducted by Cohen and Hegarty (2007, p. 182) the mean score of correct answers was 56% and 50% of the incorrect answers were of the egocentric distractor type. In a follow up study which recruited 223 participants from a range of programmes, the authors reported that the egocentric distractor accounted for 69% of the incorrect answers selected (Cohen & Hegarty, 2012, p. 872.). A more recent study by Bailey et al., (2018, p. 345), was designed to determine if performance differed in paper based and computerised versions of the test. The study involved a total of 244 undergraduate students, also recruited from a range of programmes with 118 (48.4%) completing the paper based test and the remaining 126 (51.6%) completing the computerised test. The results showed that the egocentric distractor was selected 53.3% of the time in the paper test compared to 40.8% for the online test. The analysis of the data set of 39 participants who completed the paper based test in studies one and two of the pilot phase of this programme of research showed that from a total of 201 incorrect answers, 117 (58%) were egocentric distractors.

Further assessment of the results from study four demonstrated that those in the low band selected more foils (48.1%) than the intermediate (46.2%) or high groups (10%). Analysis of the proportion of foils selected out of the total number of wrong answers shows a negative correlation at baseline score for the percentage of foils selected in relation to the total number of wrong answers. This is a statistically significant result determined by one-way ANOVA ($F(1, 47) = 5.45$, $p = .02$ which indicates that those participants who score higher also selected fewer foils. The radiotherapy group, on average, selected fewer foils (36%) compared to the diagnostic imaging group (45.2%) but this difference is not significant as determined by one-way ANOVA ($F(1, 47) = 1.48$, $p = .2$. Exploration of the data by spatial visualisation groupings of low, intermediate and high indicates that those in the higher band in the radiotherapy group did not select any foils while those in the low bracket selected foils 72.5% of the time. Conversely, those diagnostic imaging participants in the intermediate group actually selected more foils (47.6%) than those in the low control group (43.3%). Because the egocentric distractor is one of three incorrect choices, its selection by chance would produce values closer to one third making these findings are interesting. The results indicate that, in general, those students who achieved higher performance scores selected fewer foils which could be expected. As such it is possible that the selection of the egocentric distractor may be an additional indicator of lower 3-D spatial visualisation skill. Further analysis of these data shows that, in some cases, when participants selected incorrect answers, they only selected the foils. This consistency in selecting the incorrect answer and only the foil may indicate that 3-D spatial visualization is not as strong in those individuals.

6.5.10 Analysis of missing data

The flow of participants through each of the study four data collection time points is summarised in relation to missing data at the person level in table 6.14. From this it can be seen that, except for October 2014, the participation rate across both pathways ranged from 66.7% to 84.6%. The data collection event in October 2014 took place during the final campus week before the first clinical practice placement.

Table 6.14: Summary of participant flow by programme pathway for each data collection time point

	Diagnostic imaging (Control Group)	Radiotherapy (Experimental Group)	Total
	n (%)	n (%)	n (%)
Oct-13	39 (100%)	15 (100%)	54 (100%)
Apr-14	33 (84.6%)	10 (66.7%)	43 (79.6%)
Oct-14	21 (51.8%)	3 (20.0%)	24 (44.4%)
Mar-15	29 (74.4%)	11 (73.3%)	40 (74.1%)

Missing values were predicted using multiple imputations (100 iterations) and compared to individual total performance score data for both subcomponents combined from October 2013. The differences between the original dataset and the new imputation sets indicate no statistically significant difference between these two datasets across all tests. A Chi Square test found that, for April 2014, $X^2(1, N = 97) = 0.000003$, $p = .99$, while for October 2014 the result was $X^2(1, N = 77) = 0.026$, $p = .87$ and finally, in March 2015, $X^2(1, N = 94) = 0.14$, $p = .7$. These findings would suggest that the performance score results have not been impacted by missing values.

6.5.11 Discussion

This study has shown that at the commencement of a programme of study in diagnostic imaging or radiotherapy, students had different levels of 3-D spatial visualisation skill. Unlike the results of the quantitative pilot studies (studies one and two) where diagnostic imaging students outperformed radiotherapy students, the finding was reversed for the 2012-13 year one cohort reported here (although the small number of participants in the 2011 study is acknowledged). Radiotherapy students from the 2012-13 year one cohort had a higher performance score mean than their diagnostic imaging colleagues at baseline. This performance advantage was maintained throughout the 18-month duration of the study. While diagnostic imaging students had a lower performance mean at the outset, their growth trajectory was steeper and at the end of the study had shown a statistically significant improvement. The mean performance at the end of the study in March 2015, however, did not reach the level of the radiotherapy students mean performance at the same point. There were a small number of students who remained in the low category throughout, if this were to be observed in future studies, a think aloud exercise with participants explaining their thought processes while they are viewing test objects may provide additional information relating to their solution strategies . Gaining an insight in this way may further support the development of bespoke training activities in addition to focused tutorials.

Previous studies designed to measure spatial visualisation skill with the Vandenberg & Kuse MRT in radiography are limited and have focused on measurement at a single point in time. In a study conducted as part of the VERT™ evaluation report for the Department of Health (England), Appleyard and Coleman (2010, p. 24) employed the 24 item test in a randomised study of 103 students (male n = 28 [27.2%] and female n = 75 [72.8%]) to determine the relationship between spatial visualisation and performance in linear accelerator positioning skills. They report a mean performance score of 43.5%. This result is marginally higher than that of 40.7% at baseline gained in the mental rotation component of the 3-D SVT by the radiotherapy students in this study. There was no report of whether all students were recruited from the same or different year groups so the higher score may be as a result of more experienced students taking the test. In a study of similar design, Green & Appleyard (2011, p. 178) recruited 23 first year (52.3%) and 21 second year students (47.7%). The study cohort was composed of 11 male (25%) and 33 female (75%) students who achieved a mean performance score of 41.4%. Once again second year students scoring higher than the first years may have influenced this score.

In relation to the stratification of individual learners into discrete spatial visualisation skill bands, a study of 18 male (54.5%) and 15 female (45.4%) trainees in ultrasound conducted by Duce et al., (2016, p.1164) stratified performance in the 24 item MRT by standard deviation. They suggested that a score of less than one standard deviation of the mean would indicate low spatial visualisation skill, whereas a score in the range of plus or minus one standard deviation would indicate intermediate skill and a score of greater than one standard deviation would indicate high skill. Using this banding, 10 participants (30.3%) were classified in the low category, 18 (54.5%) were in the intermediate category and five (15.2%) were in the high category. The study also identified that the ability to detect individual differences in visualisation performance at baseline can support the identification of those learners with less well developed spatial visualisation.

In conclusion, it is proposed that analysis of performance in each subcomponent of the 3-D SVT can assist in the development of specific tutorials and practical workshops. It is envisaged that an attention to individual learners specific mental rotation, visualisation and perception development needs will lead to academic and clinical teams developing a more effective role for VERT™. In addition the findings from this programme of research appear to support previous findings that indicate that those with lower spatial visualisation skill have more difficulty in determining the difference between the egocentric distractor and the correct answer, suggesting that they find changing their perspective view challenging. Therefore, going forward, it is proposed that the focus should centre on an analysis of patterns for incorrect rather than correct answer choices.

6.6 Study 5: The radiotherapy skin apposition positioning task

6.6.1 Introduction

Skin apposition techniques are designed to treat tumours on or just below the surface of the skin with the radiation beam (field) coverage being determined by a shaped applicator attached to the linear accelerator head. The applicator must be positioned so that it is parallel to the skin surface so that the central axis of the radiation beam is perpendicular to the skin surface (Cherry & Duxbury, 2009, p.278). These techniques can be challenging for learners in radiotherapy since they require an operator to think and mentally visualise the positional relationships between skin surfaces and the end of the applicator. This means being able to imagine and apply movements of the linear accelerator in three perpendicular axes (X – horizontal, Y - longitudinal and Z - vertical). It also requires compensation for uneven skin contours by employing rotations around these axes (known as roll, pitch and yaw respectively). This is because uneven skin contours across the profile of the proposed treatment field may lead to an uneven dose distribution across the beam. Curvature of the skin surface will result in the skin at the periphery of the treatment field being further away from the applicator in comparison to the distance between the skin and the applicator at the centre of the beam (McKenzie & Thwaites, 2007, pp. 709-710). If this is the case then the angle of incidence at the centre of the field should be arranged such that the angles of obliquity and therefore the distances at the edges of the field are approximately equal.

In the clinical setting, patient positioning is achieved by adopting a stepwise process which begins with adjustment of the linear accelerator treatment couch height. This will bring the patient closer to the end of the applicator and will make visualisation of the relationships referred to above easier. The next step will align the applicator with positioning marks on the surface of the patient's skin by adjusting the longitudinal and lateral position treatment couch. From this point, further adjustments to the gantry angle, treatment head and couch rotations, supported by small adjustments to the couch position in the X, Y and Z planes will provide optimal positioning of the applicator in relation to the proposed treatment field. If adequate apposition cannot be achieved by adjustment of the linear accelerator, there is an additional opportunity to make small positional adjustments of the patient. In VERT™ all the movements of the linear accelerator can be replicated apart from adjusting the patient position on the couch, which is fixed. This is compensated for by making additional movements of the linear accelerator. Previous studies employing the skin apposition positioning task using VERT™ have been conducted by Appleyard and Coleman (2010) and Green and Appleyard (2011) with combined groups of first and second year students. In addition, Flinton (2015) reported on a comparison between performance of the task on a real linear accelerator and in VERT™ using first, second and third year cohorts (refer to chapter2.8, p. 61). Therefore the aim of this study was to determine if a relationship existed between baseline 3-D spatial visualisation skill and performance of the skin apposition task in VERT™ in a single cohort of second year pre-registration radiotherapy students.

6.6.2 Method and materials

An introductory Microsoft PowerPoint™ tutorial which provided an overview of the concept and principles of the skin apposition technique was delivered at the beginning of clinical preparation workshops for all second year radiotherapy students. This was followed by practical workshops using VERT™ which took place over three timetabled sessions during the six weeks prior to the first clinical practice placement of the year. The virtual patient database within VERT™ contains 13 different applicator positioning tasks of varying difficulty. Performance metrics within the software measure the closeness of fit between the applicator surface and the virtual patient skin surface and can be displayed on screen as shown below in figure 6.15.

Figure 6.15: A VERT™ skin apposition task showing the virtual patient, position of the intended radiation field (black lines), beam defining applicator and associated performance metrics. (Screenshot from UoP VERT™ platform with permission of Vertual Ltd, 2016)

Those students who had participated in baseline 3-D spatial visualisation testing in October 2013 would have their performance in the skin apposition task measured and compared with their baseline spatial visualisation skill test performance. The positioning task was carried out using the VERT™ platform with a virtual patient of mid-range difficulty as determined by the researcher in collaboration with another experienced clinical and academic radiotherapy radiographer. Each student was allocated to a 15-minute session during which they used the VERT™ OEM hand pendant to manipulate the virtual linear accelerator to achieve what they considered a clinically acceptable set-up. Each participant was offered the choice of 2-D or 3-D visualisation mode. They also had the option of manipulating the view themselves or have it done by the researcher under instruction. Measurement factors, summarised in table 6.15, were used to determine overall performance and accuracy.

Table 6.15: Measurement factors and performance indicators for skin apposition task

	Measurement Factors (Recorded Value)	Performance Indicator
1	Mean distance from applicator corner to skin surface	Closeness of fit as a mean of the distances from the four applicator corners to skin surface
2	Accuracy of skin apposition	The standard deviation of the mean distance
3	Number of equipment manipulations	The total number of adjustments to couch, gantry and collimator rotation
4	Time to complete procedure (seconds)	Timing to begin with first equipment adjustment; timing to end when participant indicates a satisfactory setup
5	Number of collisions between the virtual applicator and patient	Indicator of safety awareness and proximity of applicator to virtual patient skin surface

6.6.3 Results and analysis

Of the original 15 participants who had participated in baseline testing, 12 (80%) were eligible for positioning task performance assessment (the other three had interrupted their studies or withdrawn from the programme). Of these, 10 students successfully completed the positioning task within the 15-minute time allocation. Analysis of the time taken to complete the task and the number of equipment adjustments (moves) required found that the time ranged from 5 minutes: 15 seconds to 13 minutes: 43 seconds (mean = 9 minutes: 38 seconds), the mean number of moves required was 48.7 (SD = 15.7, range 27 – 77). Further examination of individual performance showed that the participant who made the fewest number of moves took 9 minutes: 42 seconds (equivalent to one move every 21.6 seconds), while the participant with the highest number of moves completed the task in 12minutes: 20 seconds (one move per 9.6 seconds). Both participants had intermediate baseline 3-D SVT performance, attaining scores of 50 and 46 respectively and identified as likely to benefit from additional support in mental rotation visualisation tasks. Comparing these findings with the performance of those participants in the high group (n = 3) at baseline, shows completion times ranged from 5 minutes: 15 seconds (with 44 moves) to 11 minutes: 44 seconds (57 moves), the equivalent of an average of one move per 10 seconds. For the two participants in the low group, identified at baseline as likely to benefit from support in both mental rotation and cross sectional tasks, one took 7 minutes: 45 seconds to make 46 moves and the other required 57 moves which took 13 minutes: 43 seconds, an average of one move every 12 seconds.

Assessment of the distance from the corners of the applicator to the skin surface showed a mean of 14 mm (SD = 9.1, range 5.2 – 29.7) and the accuracy of skin apposition expressed as the standard of the mean corner to skin distance (as defined by Green & Appleyard, 2011, p. 179), ranged from 0.9 to 13 as shown in table 6.16. Further analysis of the difference between the longest and shortest distances from each of the applicator corners to the skin surface for each set-up also produced a mean of 14mm.

Table 6.16: Relationship between skin apposition performance outcome measures and spatial visualisation skill

ID	Baseline SVS	Baseline SVS Banding	Task Completed	Time Taken (m:s)	Total Moves	Mean Applicator Distance (mm)	SD
5	58	Intermediate	Yes	12:42	59	6.2	1.8
8	46	Intermediate	No		24		
13	79	High	Yes	6:01	30	18.9	0.9
16	29	Intermediate	No		20		
21	46	Intermediate	Yes	12:20	77	6.3	5.7
30	58	Intermediate	Yes	8:20	33	5.2	3.6
32	75	High	Yes	11:44	57	22.3	4.6
38	21	Low	Yes	13:43	57	29.7	8.6
40	25	Low	Yes	7:45	46	8.2	1.9
45	29	Intermediate	Yes	8:48	57	9.6	12.6
46	71	High	Yes	5:15	44	9.1	11.6
51	50	Intermediate	Yes	9:42	27	25.5	13

The wide range of total equipment movements may be indicative of the level of confidence that participants had, either with using the VERT™ platform or the principles of the technique itself. The participant who performed the set-up in the lowest number of moves (27) achieved a mean distance between applicator corners to skin surface of 25.5 mm. Whereas the participant with the highest number of equipment adjustments (77) achieved a mean distance between applicator corners and skin of 5.2mm, the lowest achieved by any participant and indicating a close fit between applicator and the surface of the virtual patient. Also of note are the observations from the examination of performance of the two students who were unable to complete the task. Both were identified as having intermediate spatial visualisation skills at baseline. One asked for assistance after 2 minutes, 45 seconds, then requested an adjustment of the view perspective after 3 minutes 53 seconds and finally requested abandonment at 11 minutes 54 secs. During this time, 24 equipment adjustments were made. The second participant abandoned the task at 9 minutes, 43 seconds having made 20 adjustments. Both participants indicated that they were unable to visualise the specific concept of skin apposition techniques and apply the principles in practice.

Further analysis of the relationships between completion times, the number of equipment adjustments, set up performance, set up accuracy and baseline 3-D SVT score all demonstrate weak positive relationships as seen in the scatter plots exhibited in figure 6.16 to figure 6.19 and the summary of R^2 values for these metrics in table 6.17.

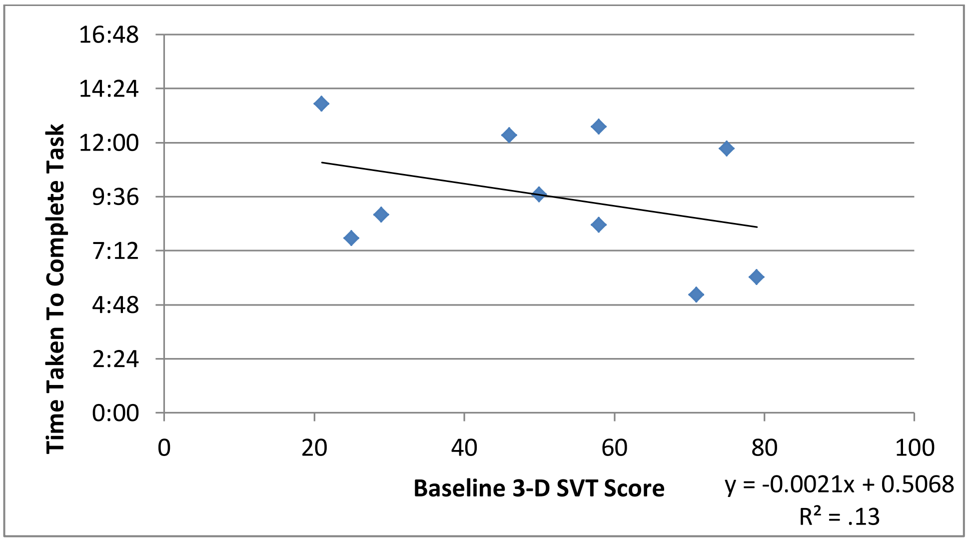


Figure 6.16: Scatter plot showing relationship between 3-D SVT score and task completion time

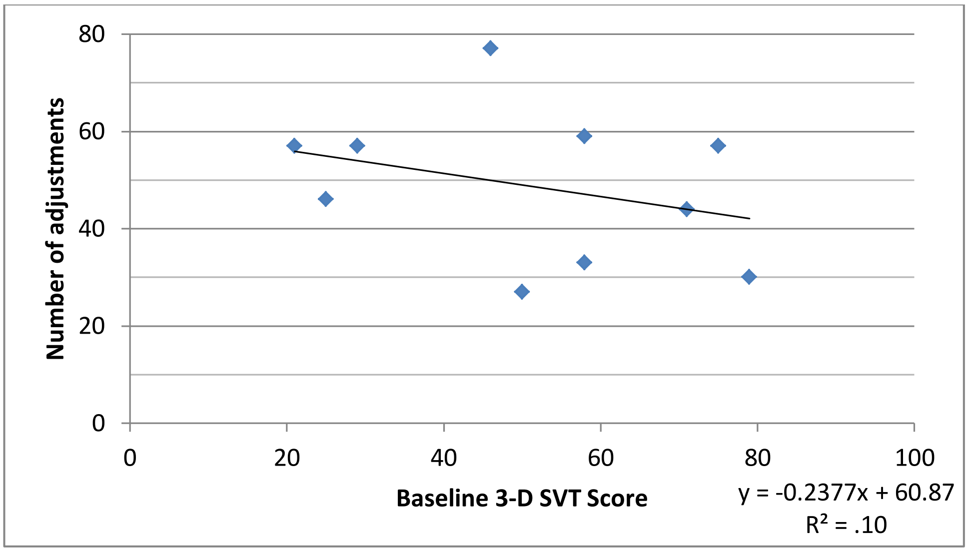


Figure 6.17: Relationship between 3-D SVT score and number of equipment adjustments

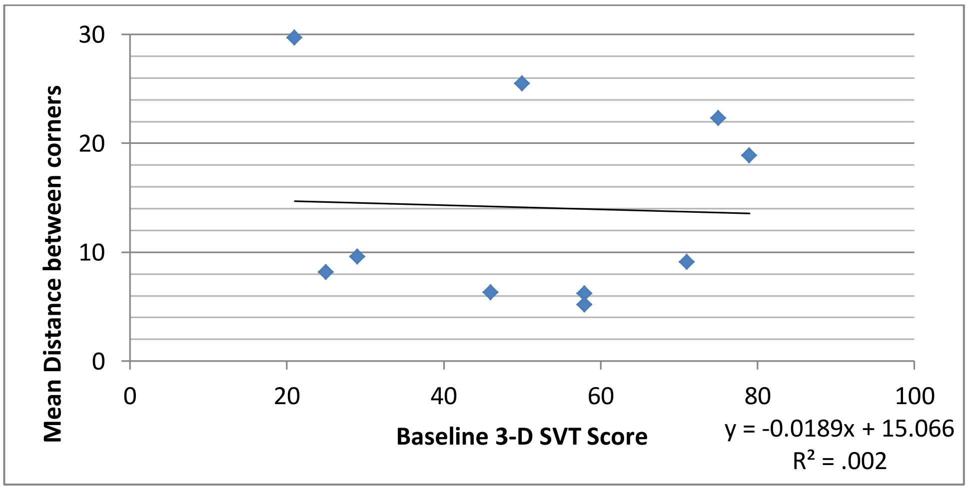


Figure 6.18: Scatter plot showing relationship between 3-D SVT score and mean distance between applicator corners (set-up performance)

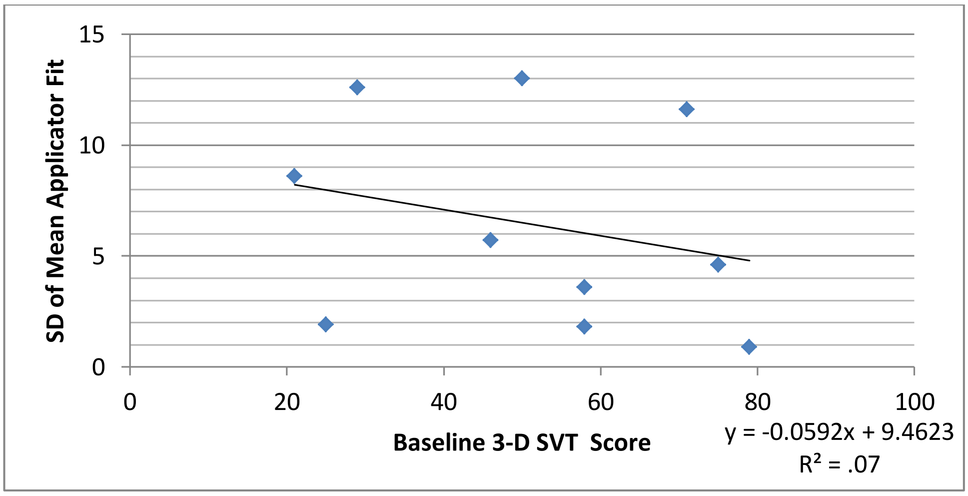


Figure 6.19: Scatter plot showing relationship between 3-D SVT performance and set up accuracy

Table 6.17: Summary of R^2 values for the skin apposition technique performance metrics

Relationships between baseline 3-D SVT and performance measures	R^2 Value
Time taken to complete task	.13
Number of equipment adjustments (moves)	.10
Set-up score (mean distance between applicator corners and skin surface)	.002
Set-up accuracy (SD of mean distance between corners and skin surface)	.07

The variation in performance observed in this study is likely to be related to the different practices with the technique observed by students across their clinical placement sites. If the applicator is positioned so that it is in contact with the skin then the distance between the skin and the radiation source will be 95cm. As outlined above, due to uneven skin contours this is not always possible, so departments will have clinical protocols which permit a gap of 5cms between the centre of the applicator and the skin surface. This results in a skin to radiation source of 100cms. Participants were instructed to apply the technique that they were most familiar with, so those attempting to achieve the latter practice employing a gap were likely to have greater distances from the applicator to skin surface.

6.6.4 Discussion

The aim of the study was to determine if a relationship existed between baseline 3-D spatial visualisation skill and performance of a complex radiotherapy positioning task using the VERT™ platform as a surrogate for the real clinical situation. It was expected that those students with higher spatial visualisation skill would complete the task with fewer equipment adjustments and greater accuracy. As such, the findings of the study were inconclusive and a number of factors may be responsible for this. The degree of positioning accuracy is determined by the performance metrics of the mean and standard deviation of the distance between each of the four corners of the applicator and the surface of the virtual patient. To achieve a close fit, coordinated equipment adjustments need to be made in order to move the virtual patient as close to the applicator as possible. All adjustments in VERT™ were made using the hand pendant, whereas in the clinic they would be made using a combination of the hand pendant and the movement controls on the treatment couch itself.

In addition, the VERT™ platform has integrated collision detection software which will automatically stop all equipment movements if one component (for example, the treatment couch) were to encroach into the safety zone of another (for example the gantry). Students are made aware of the need for safety and safety mindfulness during the first practical workshop of year one and that any collision between the linear accelerator and its accessory equipment is considered as dangerous and unsafe practice. In any assessment setting, be it formative or summative and conducted in the clinic or VERT™, a collision would be considered to be dangerous and unsafe practice and therefore classed as a technical failure. It is possible that the high number of equipment adjustments recorded for some participants was driven by the need to avoid such collisions. An additional factor could have been relative unfamiliarity (lack of experience) with the technique during clinical placement practice. Lack of confidence with the hand pendant controls was also considered as an influencing factor although its impact was considered to be low. This was based on the fact that participants were allowed to select the linear accelerator model that they were most familiar with from their clinical experience. While the time taken to perform the procedure was analysed and reported, it was not integrated into the factor analysis because the overriding end points for the technique were an accurate set-up and maintaining patient safety (i.e. no collisions between the patient and the equipment) rather than speed.

In previous studies of similar design, Appleyard and Coleman (2010, p. 50) reported a moderately positive correlation between spatial ability and set-up score of $r = .494$, $p <.01$ in a cohort of 103 students. In a study of 44 first and second year students, Green and Appleyard (2011, p. 181) reported a moderate positive correlation with a Pearson R value of .343, $p= .23$. Both studies used weighted outcome factors and compared student performance with that of experienced staff. However, it was not clear whether these staff were experienced in the clinical application of the technique and / or experienced in using VERT™. While not a criticism, it could be suggested that the weighting of each outcome factor in terms of its relative importance and what an experienced member of staff considers to be an accurate applicator set up could be viewed as rather more subjective than objective.

The performance metrics embedded within the VERT™ software are helpful when giving an indication of the accuracy of applicator set-ups and can support immediate contingent feedback to individuals on their performance in practical workshops. Nevertheless, given the range of acceptable technique variations and treatment distances described above, it would be helpful to revisit these metrics to determine whether it would be possible to develop additional metrics to identify a permissible fly zone which would take account of these differences.

6.7 Study 6: Impact of previous spatial visualisation activities and predictive factors

6.7.1 Introduction

Success in spatial visualisation tasks has been linked to a number of factors in the literature. For example, the male performance advantage over females in mental rotation tasks has been reported by Linn and Peterson (1985, p. 1479) and the interrelationships between age and spatial experiences were identified by Salthouse, Babcock, Skovronek, Mitchell & Palmon (1990, p. 128). Later work by Techentin, Voyer and Voyer (2014, p. 398) also pointed to the possibility of negative age effects in relation to working memory capacity, visualisation and mental rotation.

The spatial visualisation literature review (chapter 3.4, p. 79) also identified biological factors such as brain organisation (Zacks, 2008, p.2) and impact of dominant hand use (Casey, 1996, p.246). While Rust and Golombok (2009, p. 12) identified that 50% of the variation in the general intelligence quotient is due to inherited characteristics. If this is the case then the remaining 50% must be due to environmental influences, a theme identified in earlier work by Plomin and Petrill (1997, p. 60) who proposed a link between, what they referred to, as genetic disposition and environmental influences. These influences or factors include physical activity (Jansen & Pietsch, 2010, p. 60), choice of toys, games and other activities with a high spatial content (Terlecki & Newcombe, 2005, p. 436) and engagement with computer gaming (De Lisi & Cammarano, 1996, pp. 356-7). The aim of this study, therefore, was to identify factors which may have an influence on the baseline visualisation skill of diagnostic imaging and radiotherapy learners at the commencement of their radiography education.

6.7.2 Method and materials

For this study a demographic questionnaire containing a total of 11 sections was developed and designed to collect information related to participant age, gender, dominant hand use and engagement with activities of a spatial nature (refer to appendix 10). The determination of whether an individual is left or right handed can be defined as a function of the task they are performing (Llaurens, Raymond & Faurie, 2009, p. 882) and most will demonstrate a strong preference for one hand over the other (as in the case of the researcher, who when playing cricket, will bat right-handed but bowl left-handed, (as identified in chapter 4.14, pp 149 - 150.). Referred to as handedness, this preferential bias to act with the right or left hand (Andersen & Siebner, 2018, p. 123) can be quantified using laterality scales. The scale selected for use in this study was the Edinburgh Handedness Inventory (Oldfield, 1971, p. 110). The inventory (see appendix 10, question three) takes account of different preferences for eight different everyday tasks and yields a hand preference (or laterality index) ranging from strongly-right handed to strongly left-handed. Individuals identify preferences on a five-point Likert type scale of always right (score five), mostly right (score four), no preference (score three), mostly left (score two) and always left (score one) from which a preference score can be derived. So someone who is right handed for all tasks would score 40, no overall preference would have a score of 24 and if they were left handed for all tasks they would score eight.

Information related to spatial experience and activities was gathered from a spatial activities questionnaire. Questions were based on and adapted from activity surveys described by Newcombe, Bandura and Taylor (1983, p. 380) and Terlecki and Newcombe (2005, p. 434). These surveys identified computer and videogame usage frequency and preferences (appendix 10, questions four to eight) and activities that a population of North American high school and college students might be expected to engage in. Using this list as a starting point, activities of a similar nature were grouped together and where required adapted to match UK style activities. In addition, questions relating to involvement in sports, hobbies and types of toys and games played with when younger were included (appendix 10, questions nine to eleven). The survey included closed questions requiring a yes or no response and open ended and multiple-choice Likert-type frequency questions in order to ascribe quantitative values to qualitative opinion data and, to make it amenable to statistical analysis. All participants who attended the first spatial visualisation testing session in October 2013 accessed the survey via a hyperlink in Moodle Quiz, once they had submitted their answers for the SBST subcomponent of the 3-D SVT.

6.8 Results and analysis for demographic profiles and influencing factors

The findings related to possible influences on 3-D spatial visualisation skill arising from biological and environmental factors will be reported and presented using all available data sets from both phases of this programme of research. A total of 80 first year diagnostic imaging and radiotherapy volunteers were recruited, 26 (32.5%) for the pilot phase and 54 (67.5%) for the experimental phase. It should be noted that while 3-D spatial visualisation skill measurement for the two phases took place at a different time point in the academic year, the data being analysed is unlikely to have been influenced by clinical placement experience. While the first pilot phase testing took place in April (towards the end of the academic year), the first experimental phase testing occurred in October (at the beginning of the academic year), the only difference between the test instruments was that the pilot phase employed a paper based platform initially and the experimental phase employed an online version, but using the same design of test objects. The results reported below will examine the relationships between the factors of age, gender, handedness, activities and games and the number of correct answers achieved in the 3-D SVT overall and in the subcomponent for mental rotation and solid cross sections.

Demographic data was automatically downloaded from Moodle Quiz into a Microsoft Excel spreadsheet and was collated with individual participant 3-D SVT test performance scores using the unique participant identification number in the same way as described for the pitot phase studies described in chapter 5.4.3, p.166. Following cleaning by removing Moodle quiz participant identification numbers and response codes, extra spaces and empty cells and converting text into numbers for coding, data was exported to SPSS version24 for analysis.

6.8.1 Participant age profile

The overall profile for participants who provided data for age (n = 78) shows a mean of 23.8 years (SD = 7.1, range 18 – 46). A histogram demonstrating the frequency distribution and normality plot is presented as figure 6.20 and shows a skewing towards the lower age range.

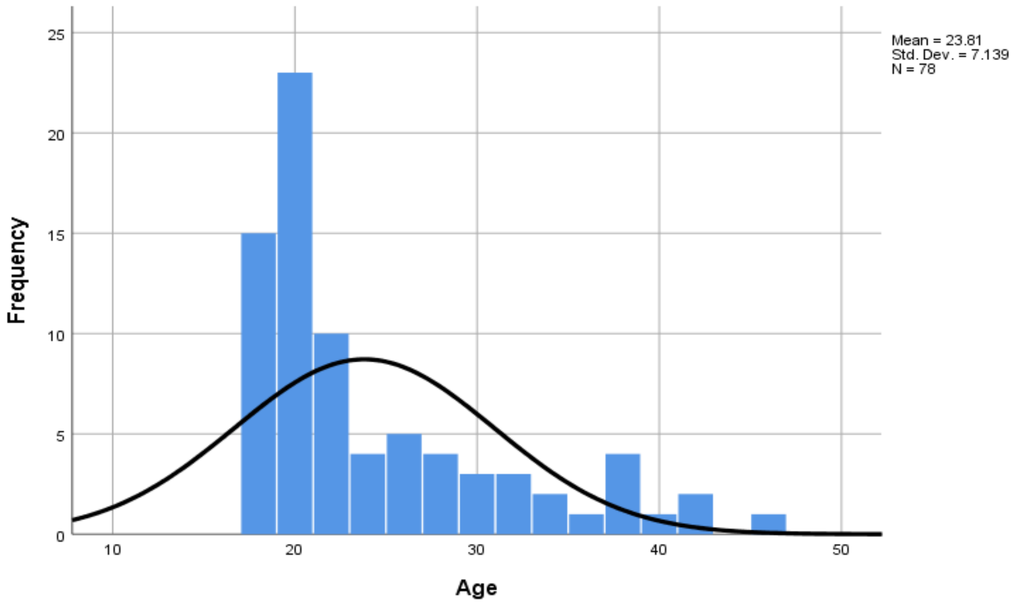


Figure 6.20: Age distribution and normality plot for all participants in study one (April 2011) and study four (October 2013)

Further examination of the data shows that male participants (n = 21, 26.9%) were older than their female counterparts (n = 57, 73.1%) by an average of 2.2 years but the range was similar as demonstrated in table 6.18.

Table 6.18: Summary of descriptive statistics for age and gender (both study phases)

	Male	Female
n	21	57*
Mean	25.4	23.2
SD	7.0	7.2
Min	18	18
Max	42	46

*One female participant in the pilot phase and one in the experimental phase who did not provide age data

6.8.2 Impact of gender and age profile on 3-D SVT performance

While the literature (for example, Techentin, Voyer & Voyer, 2014, p. 398), identifies age as a moderator for decline in working memory capacity and spatial visualisation tasks such as mental rotation in older people, the classification for what is considered as old is variable. In relation to computer navigation tasks, Pak, Czaja, Sharit, Rogers and Fisk (2008, p. 3047), refer to older age as 60–91, with middle-age being 40–59 and young age as 18–39. But with reference to the use of technology in the workplace, Morris and Venkatesh (2000, p. 455), apply the term older to individuals at just 40 years of age. While the numbers of participants over the age of 39 reported here are small (n= 3, 3.8%), the relationship between age and the number of correct answers achieved was examined by performing a chi-square test of independence. The relation between these variables was not significant X^2 (420, n = 78) = 481, p .20. The influence of gender on performance in each subcomponent test and total number of correct answers gained is summarised in table 6.19.

Table 6.19: Comparison of number of correct answers achieved for males and females

Gender	**Male**			**Female**		
n	21			59		
Test	MRT	SBST	Total	MRT	SBST	Total
Mean	3.9	7.4	11.3	3.9	6.6	10.5
SD	2.3	4.3	6	1.7	3.0	4.0
Min Max	0 8	0 14	0 21	0 7	0 14	2 20

It can be seen that the mean number of correct answers in the mental rotation subcomponent of the 3-D SVT is the same for both males and females and the range is similar, although it is acknowledged that this is a low score in relation to the total number of test objects in this subcomponent (12 in the pilot phase and 10 in the experimental phase).

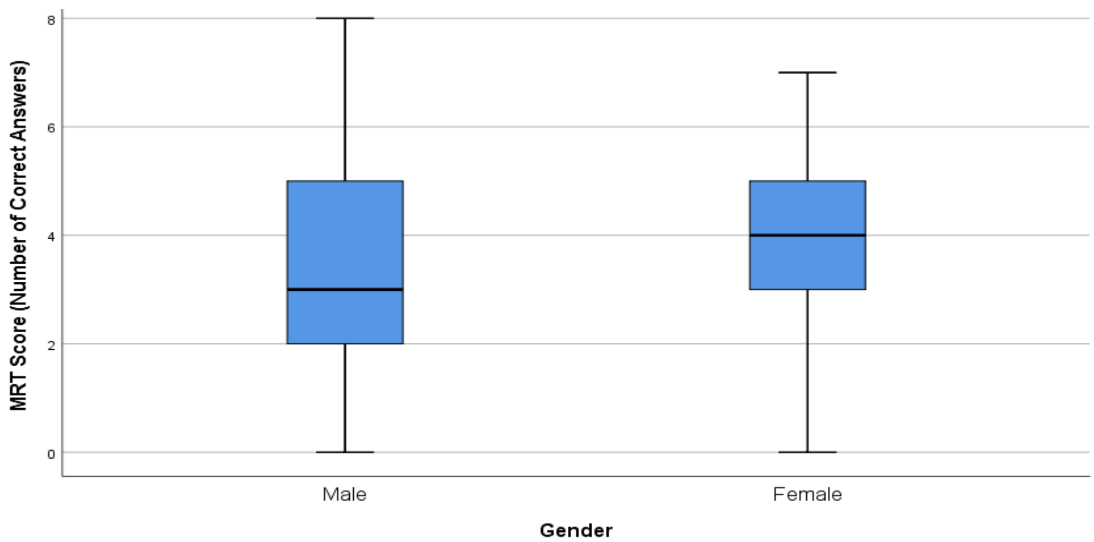


Figure 6.21: Box plot to show number of mental rotation test correct answers gained by males and females (April 2011 and October 2013)

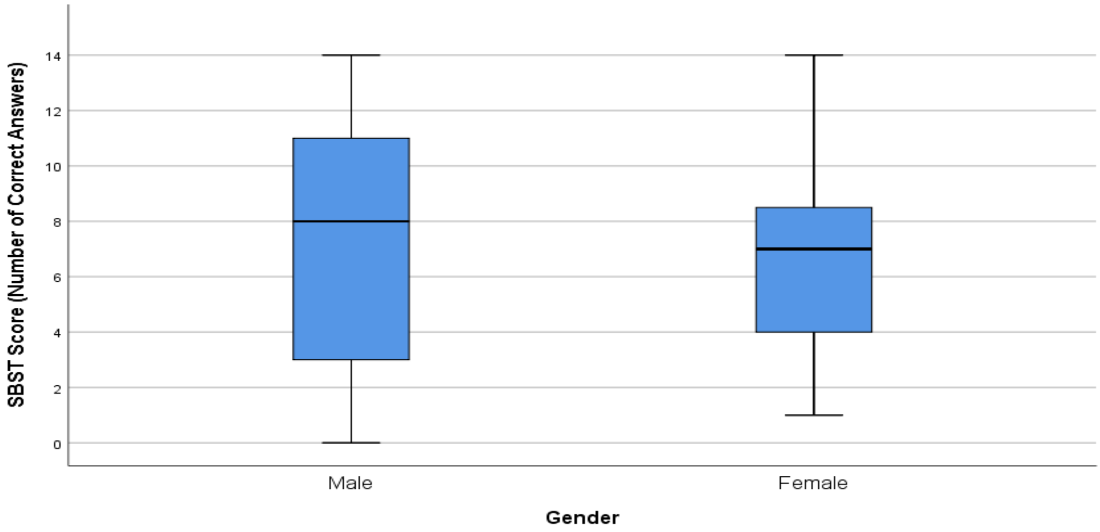


Figure 6.22: Box plot to show number of solids test correct answers gained by males and females (April 2011 and October 2013)

Examination of the mean scores and the range shows a similar performance in the MRT for both pathways, while the median value for females is higher than for males (figure 6.21). For the SBST, the mean and median performance is higher for males but the range is similar (figure 6.22). No ceiling effect (that is, no participant answered all questions correctly) was observed for the MRT; however one male and two females identified all items correctly in the solids test. This lack of male advantage is an interesting finding given the widely reported male advantage in mental rotation tasks and similar to that observed by Flinton (2015, p. 134) when comparing male and female performance in a VERT™ positioning task. In that study females outperformed males, which led Flinton to conclude that gender differences may be removed by training. As participants from the experimental group were tested prior to any practical experience, the impact of training was unlikely to be the explanation in this study.

6.8.3 Influence of handedness

The determination of whether an individual is left or right handed can be defined as a function of the task they are performing (Llaurens, Raymond & Faurie, 2009, p. 882). In this study, two factors of hand use were examined, those of preferred hand for writing and overall preference for task performance. The results for preferred writing hand are presented in table 6.20. The total number of participants stating that they are left handed (n=15, 19%) is higher than the figure stated for the general population with McManus (2009, p. 37) reporting that 10% are left handed.

Table 6.20: Participant demographics for pathway, gender and dominant writing hand

Study Phase	Gender	Right	Left
Pilot	Male	7	0
	Female	16	3
Experimental *	Male	9	4
	Female	32	8
Total		64 (81%)	15(19%)

(* one participant did not specify)

The analysis of overall hand preference across all tasks for the experimental group as shown in figure 6.23 demonstrated that 43 (79.5%) participants indicated that they had a preference for being always or mostly right handed. One participant (2%) demonstrated no overall preference and 10 (18.5%) participants were mostly left, no participants indicated an always left preference which is more in line with the preference shown by the general population. The overall 3-D SVT performance score for those with a right-handed preference showed a mean of 40% (SD = 13.6, minimum 25, maximum 58) while those with a mostly left preference achieved a mean of 42% (SD = 13.6, minimum zero, maximum 79).

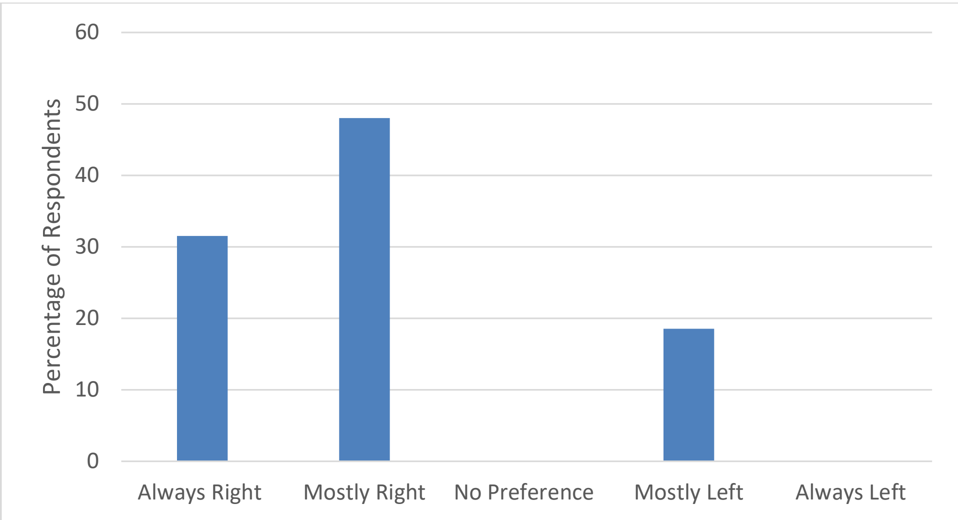


Figure 6.23: Percentage experimental group participants expressing a preference for right or left hand

6.8.4 Profile of activities for experimental phase participants

Participants in experimental phase study four were asked to provide data relating to the types of activities that they had previously engaged in, or were still engaged in at the time of the data collection for baseline spatial visualisation skill measurement in October 2013. In addition they were asked to indicate which toys and games they had played with when younger. There were a total of 53 data sets available for analysis and the numbers of participants for each category are reported in table 6.21 and 6.22 below. The proportion of respondents for each category is shown in figures 6.24 and 6.25.

Table 6.21: Summary for number of participants engaging in activities of a spatial nature

Recreational Sports & Activities					
Team sport	Individual sport	Drawing in 3D perspective	Mechanical/Technical drawing	Arts and Crafts	Juggling/ Baton twirling
49	48	28	21	42	8

Table 6.22: Summary of the types of toys participants reported playing with as children

Toys							
Action figures	Arts / Crafts	Construction toys	Model building	Puzzles	Dolls / Puppets	Electronic hand held games	Board games
22	28	31	16	37	35	27	42

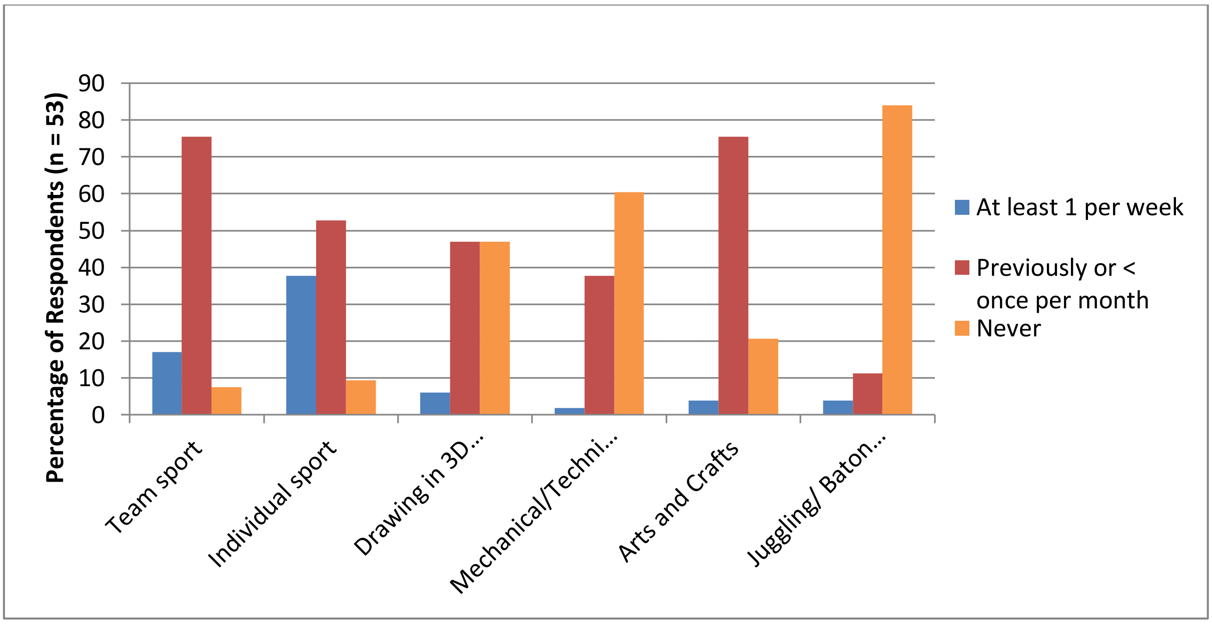


Figure 6.24: Frequency of engagement with recreational sports and activities for the experimental study phase participants

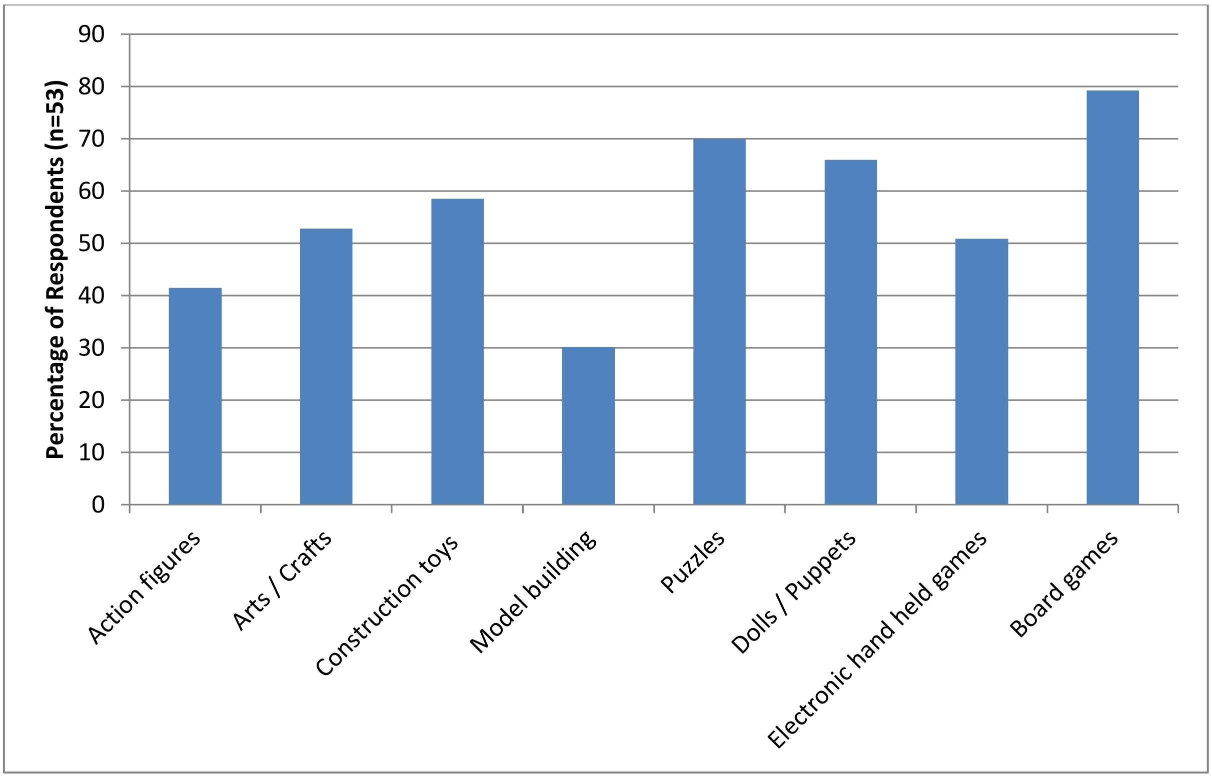


Figure 6.25: Proportion of experimental phase participants reporting play with types of toys and games when children

The relationship between gender, activities and total performance score was examined by performing a chi-square test of independence. The relation between activities and performance was not significant as demonstrated in table 6.23.

Table 6.23: Chi-Square analysis for influence of gender and activities on3-D SVT performance

Spatial Activity Factors	Chi-Square Value	P value
Gender	1.36	0.51
Team Sport	4.30	0.12
Individual Sport	4.74	0.09
Drawing in 3D perspective	3.47	0.18
Mechanical/Technical Drawing	1.07	0.59
Arts and Crafts	2.70	0.26

6.8.5 Influence of computer gaming

Experimental phase participants were asked to indicate whether they played computer games and if so, the frequency and type(s) of games played. Table 6.24 shows the number of identifying that they did play (n = 20, 37%) and figures 6.26 and 6.27 show the frequency and duration of computer gaming across both pathways in the experimental study phase. These values are somewhat lower than those figures reported by Müller, et al., (2015, p. 568). In a study cohort of 12, 938 European college students (Male = 6,097, 47%, Female = 6,841, 53%, mean age = 15.8, SD = 0.7) which indicated that 7,828 (60.5 %) of the sample reported playing, with a higher proportion of males, n = 5,036 (84.7 %) compared to females n = 2,792 (42.8 %). Participants were also asked to indicate the length of time they had been computer gaming with 17 (80%) indicating that they had been playing for more than 5 years. The results show that those students who played computer games generally performed better than those who did not. The results also indicated a significant relationship between length of computer gaming and score, however this test appeared to show that students who had played games for longer scores lower which appears to be counterintuitive. Analysis of the specific results show that one student who played games infrequently scored highly (73%) while another frequent gamer scored low (14%). This suggests that this is a statistical inference rather than a real-world inference and that a larger data set would be required in order to test this result in more depth.

Table 6.24: Participant demographics for programme pathway, gender and computer gaming

Pathway	Gender	Computer	Non-Gamer (including not
Diagnostic	Male	7	3
	Female*	8	21
Radiotherapy	Male	2	2
	Female	4	7
Total		20 (37%)	34 (63%)

(*2 participants indicated that they were not gamers but answered the questions relating to frequency and game type. These participants were excluded from the analysis)

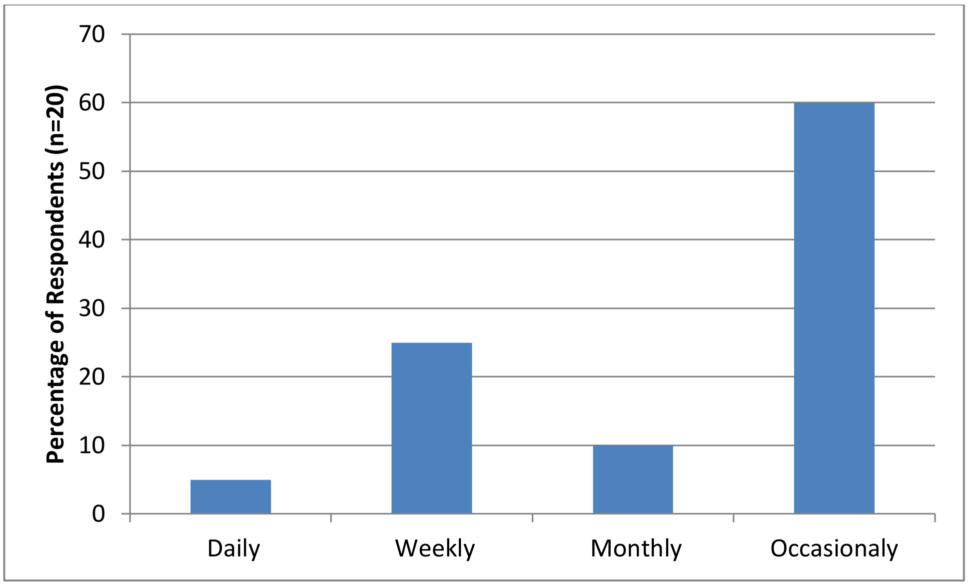


Figure 6.26: Frequency of computer gaming activity for the experimental phase participants

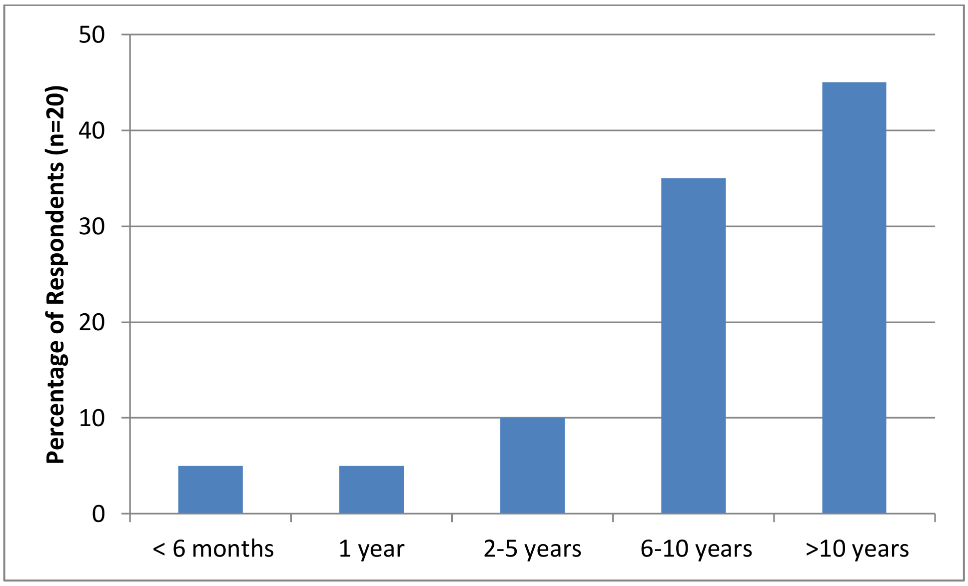


Figure 6.27: Duration of computer gaming activity for the experimental phase participants

Table 6.25: Summary of statistical significance for computer gaming factors

Spatial Activity Factors	Chi	P value
Engagement with computer gaming	5.06	.02*
Gaming Frequency	0.04	.85
Gaming Length	4.24	.04*

* significant values

6.8.6 Discussion

This study sought to determine if biological and environmental factors would have an influence on 3-D spatial visualisation skills. Biological factors include age, gender and preferred hand use for a range of everyday tasks, while environmental factors cover engagement with activities, toys and games which have spatial components. Examples include team and individual sports, 3-D perspective drawing, playing with construction toys and computer gaming. The findings showed that while computer gaming may be an influencing factor in relation to baseline 3-D spatial visualisation skill test performance at the commencement of radiography education, the remaining factors demonstrate a weaker relationship. In a meta-analysis of the literature pertaining to the relationships between motor expertise and performance in tasks of a spatial nature, Voyer and Jansen (2018, p. 120) report that the effect sizes for athletes, participants in ball sports, runners, dancers and cyclists are close to zero. This analysis would confirm the findings of this study.

6.9 Chapter summary

The chapter began by discussing the development and deployment of an online 3-D SVT instrument to measure baseline visualisation skill and to track any development over time (study four). The longitudinal study demonstrated that learners can be classified as having low, intermediate and high spatial visualisation skills based on their overall baseline test performance score. Additional analysis of the visualisation subcomponents of mental rotation and visual perception can identify specific areas which may benefit from additional visualisation exercises and tutorials. This in turn would provide a more focused and more effective role for VERT™, which moves beyond the current learning outcomes based on a one size fits all approach to practical clinical workshops and tutorials. Study four also found that spatial visualisation skill gains in both study groups were observed over time. This gain was significant for the control group of diagnostic imaging students but not significant for the experimental group of radiotherapy students although they did start with higher performance scores at baseline. Possible reasons for this were considered and relate in the main to these students not being able to gain hands-on practice with real radiotherapy equipment on campus in the way that their diagnostic colleagues can in the X-ray suite.

Study five explored the relationship between the radiotherapy cohort baseline spatial visualisation score and performance of a simulated positioning task in the radiotherapy virtual environment. The findings were inconclusive but the published evidence base for performance in this task is limited. It is recommended that the potential and feasibility of developing additional performance metrics, in addition to those currently available in the software, should be explored further. If this is possible then a larger collaborative study should be considered.

Study six sought to determine if age, biological factors such as gender and handedness and environmental influences such as previous experience with spatial activities could predict spatial visualisation performance score at the start of radiography education. The findings demonstrated that in line with previous studies reported in the literature; those who engage in computer gaming activities perform significantly better than those who do not. Other spatial activities do not have a significant relationship as a predictor of performance.

While there might be value in exploring the relative importance of the biological and environmental factors in more detail when considering training in a virtual environment, the explicit purpose of VERT™ is to support and facilitate the necessary visualisation skills to support clinical practice, rather than how to perform in a 3-D environment. As such, while understanding which factors might influence a student's visualisation skill is interesting, a student's baseline score, and the subsequent tailoring of tutorials to their needs is likely to be more informative and can ensure best practice in academic and clinical education through identification of effective use of clinical preparation time in VERT™. As the main aim of this phase of the programme of research was to determine if their baseline performance score could identify learners with high, intermediate and low skill, it is suggested that there is no added value in collecting spatial activity information.

Chapter 7

Discussion, conclusions and recommendations

7.1 Introduction to the discussion chapter

The Virtual Environment for Radiotherapy Training (VERT™) was introduced into radiotherapy education and clinical practice in England during 2008 to improve the clinical experience, confidence and understanding of year one students. Since that time with the regular introduction of additional functions and software upgrades its role has expanded to cover aspects of radiotherapy physics, cone beam CT and patient information. Worldwide, the platform is used for supporting learner understanding, staff development and patient education using a wide range of scenarios as reported in chapter 2.8, pp. 50 - 73, the role of the platform in the development of 3-D spatial visualisation skills of learners has been less widely investigated.

This chapter will begin by summarising the background to the programme of research and the specific aims and objectives that it set out to achieve and the research questions that it sought to answer. It will place the aims and research questions at the centre of the discussion related to the measurement of the spatial visualisation skills of pre-registration learners in radiotherapy. It will draw together the key themes relating to the importance of well-developed 3-D spatial visualisation skill in modern radiotherapy practice reported in chapter two, how it may be measured and whether it may be improved through practice, also discussed in chapter two. It will continue with an examination of the strengths and limitations of this programme of research and will present the conclusions drawn from the results of the six studies that were conducted during the course of this programme of research. These will be set within the context of the strengths and limitations of the chosen methodologies. Finally, the chapter will identify the implications for future radiotherapy educational practice and make recommendations for the direction of future research within radiotherapy and other areas of healthcare science and medical fields where -3D spatial visualisation skills are required.

7.2 Summary of findings in relation to the research aims, objectives and questions

Modern radiotherapy practice using advanced and complex techniques such as IMRT, VMAT and SBRT, can achieve a steep dose gradient between the tumour target volume (which receives high dose) and surrounding normal tissue and organs. The mental model building and visualisation of these relationships requires radiotherapy radiographers to have highly developed 3-D spatial visualisation skills. This is because the 3-D mental visualisation of internal anatomy is a fundamental component of the accurate positioning of a tumour target volume in relation to normal anatomy and the planned direction of radiotherapy beam paths. The Virtual Environment for Radiotherapy Training was conceived and developed to support the visualisation of these tumour target volume, radiation dose and anatomical relationships in these complex techniques. The platform's introduction across clinical radiotherapy centres and education providers through 2008 was followed by a Department of Health (England) evaluation project. The report by Appleyard and Coleman (2010, p. 7) identified that, while the impact of VERT™ was largely positive. However one of the recommendations was that inherent spatial ability of radiotherapy students should be assessed to assist identification of individuals who would benefit most from VERT™ experience (ibid, p. 26). In order to determine this inherent spatial ability, it was important to first define what inherent spatial ability is. From this starting point, it would be possible to identify the components of spatial ability and how they might be measured. The spatial visualisation literature, reported in chapter 3.7 & 3.8, pp. 91 – 100, identified a range of tools ranging from 2-D tests employed as part of general intelligence testing to complex 3-D objects used in the testing of specific spatial visualisation components. If appropriate measurement tools, representative of the patient positioning workflow in radiotherapy could be found, it would be possible to determine baseline spatial visualisation skill at the commencement of radiotherapy education and training. In so doing the concept of banding individuals, based on this, into high, intermediate and low visualisation skill would support the identification of those who would benefit from the visualisation tasks available in VERT™

Therefore the overarching aims of this programme of research were:

1. To gain an understanding of the spatial visualisation skill of pre-registration learners in radiotherapy in one United Kingdom Higher Education Institute;
2. To determine whether it was possible to stratify pre-registration radiotherapy learners in terms of their baseline spatial visualisation skill;
3. To determine the longer term potential of VERT™ in relation to the development of 3-D spatial visualisation skill.

The specific objectives based on these aims were formulated and are summarised in table 7.1 below, together with the associated findings.

Table 7.1: Summary of research objectives and findings

Objective (Study number)	Findings
1. Conduct a systematic review of the literature to identify and define the components of spatial visualisation skill required for radiotherapy practice	Linn and Peterson (1985, p. 1482) identified the components of mental rotation, visualisation and perception (chapter 3.3, p. 76): Mental rotation can be likened to correctly positioning patient while visualisation and perception are related to building a mental model of the radiation beam path and its relationships with normal anatomy and the tumour target volume.
2. To identify and test valid and reliable 3-D spatial visualisation skill measurement tools via a critical evaluation of the spatial visualisation testing literature (Study 1)	The critical evaluation identified the Vandenberg & Kuse MRT & the SBST as validated tools for SVS subcomponents (chapter 3.11, p. 76); These tools were pilot tested using paper and online presentation platforms. The Cronbach alpha for the online platform was .76. Walker and Almond (2010, p. 86) have identified alpha values above 0.7 as indicating good consistency, (chapter 5.7 Validity and reliability, p. 178).
3 Develop an appropriate test instrument for use in radiotherapy;	The development of a 3-D SVT consisting of Vandenberg & Kuse MRT objects and SBST cross sectional items
4 Compare performance in paper and online versions of the test instrument	This was conducted as study one in the pilot phase of this programme of research. The findings reported in chapter 5.9, p. 179 showed comparable performance scores across the paper and online test platforms.
5. To determine if the baseline spatial visualisation skill of pre-registration learners in radiotherapy could be measured (Study 4)	Yes. Results also show that students can be grouped according to performance score (chapter 6.5.6, p. 214).
6. To determine if spatial visualisation changes over time (Studies 2 & 4)	The results of study 4 demonstrate that it does (chapter 5.8.1, p. 180 for study 2 & chapter 6.5.6, pp. 210 – 217 for study 4).
7. To determine if a relationship exists between baseline spatial visualisation skill and performance in a complex radiotherapy positioning task (Study 5)	The findings of this study were inconclusive, would benefit from the development of additional performance metrics (chapter 6.6.3, pp. 224-228).
8. To determine if a relationship between baseline spatial visualisation skill and previous spatial visualisation experience exists (Study 6)	The only statistically significant factor is computer gaming (chapter 6.8.5, pp. 244 - 246).
9. To make recommendations for future educational practice	Grouping individuals by their performance scores into, high intermediate and low SVS and the follow up analysis of MRT and SBST subcomponent scores can facilitate development of focused SVS syllabus content. Analysis of SBST incorrect answers can act as an additional screening tool for low SVS; Use the 3-D SVT in other healthcare fields involving practice in complex environments.

In the past decade there has been a shift away from paper based testing towards computerised, electronic test formats. Although, as Bailey, Neigel, Dhanani and Sims (2018, p. 340) identify, in general, research has not yet established equivalence between computerised and paper testing. The pilot phase of this programme of research, reported in chapter five, tested a new 3-D SVT in paper and online platforms. The online test was well received by students and did not produce significantly different overall performance results in comparison to the paper version. There was, however, a difference in the pattern of incorrect answers for the SBST subcomponent. Overall, the paper based test produced 38 incorrect answers choices, of which 18 (47.4%) were foils whereas the online test produced a total of 84 incorrect answers with foils accounting for 45 (53.6%).

Taken together, these findings warranted further exploration and with further development, an online test platform was designed for use in the first study of the experimental phase of this programme of research (study four). Two questions were posed for this study, the first sought to determine the extent to which spatial visualisation skill of pre-registration radiotherapy students could be measured. The second asked: does spatial visualisation skill change during the programme of study? A longitudinal controlled study was designed and conducted to measure the baseline spatial visualisation skill of volunteers who were recruited from cohorts of diagnostic imaging and radiotherapy students. The justification for including diagnostic imaging students in this study was based on the fact that the patient positioning and beam path visualisation requirements are similar. While the VERT™ is predominantly a radiotherapy platform, many of its visualisation scenarios would have potential benefit to other professions who work in complex 3-D clinical environments.

The findings demonstrated that analysis of performance scores in the 3-D SVT at baseline could identify whether an individual had low, intermediate or high visualisation skill. In addition, analysis of performance in the MRT and the SBST for spatial perception and visualisation skills provided an insight into the type of focused tutorial input low and intermediate performers might benefit from. This is of particular importance since mental rotation can be likened to positioning a patient correctly in relation to the linear accelerator, while the cross section component of spatial visualisation relates to building a mental model of the beam path and its relationship to the tumour target volume. While spatial visualisation performance scores were shown to improve over time, the gains were small and may have been due to the one size fits all approach to practical workshops not fitting with some participants' developmental requirements.

During the course of this research, Ziemkiewicz et al., (2012, pp. 8-89) identified that, in relation to the perception of visual patterns having a knowledge of individual differences in ability could assist the development of guidelines for the introduction of visualisations of complex tasks. This would facilitate a move beyond a one size fits all approach since there was no single way for a visualisation to support a given task. But they also acknowledged that the field of visualisation theory lacked the tools to analyse the factors which could lead to a difference in performance. Based on the findings of study four in particular, it is proposed that analysis of patterns of incorrect answer choices in the SBST would assist in this analysis.

Study five sought to determine the extent to which baseline visualisation skill might influence the performance of a complex (skin apposition) positioning task using the VERT™ platform. The results of the observational study which involved 12 radiotherapy students carrying out a complex treatment task using VERT™ were inconclusive. The platform was employed as a surrogate for a real clinical task because it has performance metrics built into the virtual patient software that provides information about position accuracy. It was hypothesised that those students with higher spatial visualisation skill at the commencement of their education (based on baseline performance score) would perform the task in a shorter time and with fewer equipment adjustments when compared with those with less well developed skill. This was unproven; regression analysis for the performance metrics for time taken, number of equipment adjustments and baseline spatial visualisation skill, for the 10 participants who completed the task had R^2 values of .13 and .1 respectively. The reasons may be twofold; there could be a lack of confidence with the operation of VERT™ or with the application of the technique itself. An important finding from the study was that the participants who did not complete the task made fewer adjustments over a longer time period. These findings suggest that there is a need to not only reassess how VERT™ is used to perform these positioning tasks, but also how to measure students' performance in them more accurately.

The purpose of study six was to identify factors which may affect the development of spatial visualisation in radiotherapy and diagnostic imaging learners? The study analysed the demographic factors of gender and age, handedness and engagement with sporting and other activities such 3-D perspective drawing, play with construction type toys and computer gaming. These activities are considered to have inherent components of mental rotation, spatial perception and spatial visualisation and have been identified in the literature as being influencing factors in the development of spatial visualisation skill in children and teenagers. Interestingly, the only factor shown to have a significant effect on baseline 3-D spatial visualisation skill was computer gaming. It might be expected that free form model building with Lego® bricks, for example, would involve mental visualisation, perception and the required rotation of individual bricks prior to the physical construction of the proposed model. What is not clear from these findings is whether computer gaming contributes to the development of 3-D spatial visualisation, or whether those with already well developed 3-D skills are drawn to computer gaming. Investigation of these relationships could be incorporated into future radiotherapy, diagnostic imaging or wider health care science visualisation studies.

7.3 Comparison of findings with other spatial visualisation studies

While the study of spatial visualisation skill is not new, research has tended to focus on the determination of different effect sizes between male and female performance, prediction of performance in, for example, engineering drawing and surgical tasks and the contribution that it makes to success in STEM subjects. In radiotherapy specifically, recent spatial visualisation literature has focused on the development of hardware and software platforms for the reconstruction of 3-D spatial data for tumour imaging. The 3-D SVT results for the cohorts studied in this programme of research were triangulated with previous studies reported in radiotherapy and diagnostic imaging. But it has to be recognised that while comparisons have been made, differing scoring conventions and methodologies have been used in these studies, so these comparisons should be treated with caution. In radiotherapy, there have been two previous studies by Appleyard and Coleman (2010) and Green and Appleyard (2011) and reported in chapter 6.5.4, p. 204. Both studies employed one spatial visualisation test instrument, the Vandenberg & Kuse MRT. The mean performance across all participants (n=149) was 42.5% while the 29 radiotherapy participants in this programme of research achieved a mean score, for the mental rotation subcomponent of 40%. In relation to diagnostic imaging, a recent study of 33 novice sonographers by Duce et al., (2016, p. 1164) reported that the mean performance in the revised 24 item Vandenberg and Kuse test was 45%.

One of the main drivers for this programme of research was to determine if, through the measurement of 3-D spatial visualisation, learners who would benefit most from experience with VERT™ could be identified. To achieve this, individuals were grouped into low intermediate and high categories based on their total performance score from both subcomponents of the 3-D SVT. This exercise found that there were proportionally fewer low performers in the radiotherapy group (n = 2, 13.3%) compared to the diagnostic imaging group (n = 11, 28.2%) There are also more high performers in the radiotherapy group (n=3, 20%) compared to the diagnostic imaging group (n= 2, 5.1%). Again it should be noted that any comparison with other studies is limited by the lack of a consensus on the performance boundaries that define, low, average and well developed spatial visualisation skill. However, the study by Duce et al., (2016, p. 1164) identified that of the 33 participants, 10 participants (30.3%) were classified in the low category, 18 (54.5%) were in the intermediate category and five (15.2%) were in the high category, based on their MRT performance. These findings are not dissimilar to those reported within this programme of research, but it is important to note that the important factor is not the comparison of performance with other individuals or groups; it is the identification of individual students who may require and benefit from additional support.

During the course of this programme of research, Veurink and Sorby (2019, p. 156) reported on the use of the 30 item Purdue Spatial Visualisation Test: Rotation for identifying the level of spatial visualisation development. A total of 3,938 first year engineering students in a North American university were screened between 2009 and 2014 and those students correctly identifying 18 items or less (equivalent to a score of 60%) were classified as having lower – weak visualisation. These students were required to participate in a "spatial skills" course consisting of paper based projectional drawings of normal and rotated surfaces, planes and 3-D objects. The students scoring between 19 - 21 and 22 – 27 were classified as marginal and average visualisers respectively and those with a score of 28 – 30 were regarded as having high visualisation skills and were not offered additional support. It should be noted that the rationale for setting the performance boundaries was not reported. Across the total population, 512 students (13%) and 539 (14%) were in the lower – weak and marginal categories respectively and 2,018 (51%) were classed as average, the remaining 869 (22%) were identified as high visualisers. A comparison of these findings with the grouping exercise conducted in study four (chapter 6.5.5, p. 211) demonstrated that, of the 54 diagnostic imaging and radiotherapy students who participated in October 2013 baseline testing, 13 (24%) were classed as having low spatial visualisation skill and five (9%) as having high skill. The remaining 36 students (66%) were in the intermediate group, which is a similar proportion to Veurink and Sorby's marginal and average groups. Also of note is the finding that, in the female engineering student cohort (n = 794, 20%), while small in comparison to the size of the male cohort (n = 3144, 80%), the number of female classified as having lower – weak visualisation skill were 224 (28%) compared to the male cohort where 288 (9%) were in the same category. The corresponding data for the radiography participants in study four showed that 36% of male students (n = 5) and 20% (n = 8) of females were categorised as low. While there was no report of spatial visualisation skill being reassessed following the spatial skills course, Veurink and Sorby (2019, p. 160) did report that the lower – weak visualisers were more successful in introductory engineering modules (gaining a pass at grade C or above). In addition, they identified that these students had higher retention and completion rates than those who were initially in the high visualisation group. While retention and completion data was not collected as part of this programme of research, it is recommended that it is collected in any follow up studies to determine the impact of additional spatial visualisation support.

From an examination of the performance results from study four in this programme of research, it was possible to identify, for each individual, which subcomponents of mental rotation and perception and visualisation (cross sections) might benefit from additional support. The results indicated that, in the control group of diagnostic imaging students, 11 (28%) of the 39 participants would benefit from exercises in both mental rotation and beam's eye view cross sectional activities and 13 (33%) would benefit from either mental rotation or beam's eye view activities. For the experimental group of radiotherapy students, two (13%) from 15 participants could benefit from activities in both components while five (33%) would benefit from one or the other. Of note is that across the population as a whole, the proportion requiring support with either mental rotation (34%) or cross section cutting plane activities (33%) is similar. This might suggest that the current practice of one size fits all approach to practical workshops would benefit from additional activities with emphasis 3-D mental model building and visualisation. The findings of the grouping exercise have shown that it would be possible to identify specific areas of focussed visualisation support for those learners with low and intermediate skills. However in the interests of equity, those identified as having high spatial visualisation skill and potentially requiring little or no additional developmental support over the standard approach, should be encouraged to access whatever additional input they feel they might benefit from.

7.4 Impact of the test-retest time period on performance and change over time

The time interval between measurement time points and the number of sessions that a study participant completes has been identified by Calamia, Markon and Tranel (2012, p. 545) as factors which may have an influence on the magnitude of any performance gain. These factors are known as retest effects and have been defined by Scharfen, Peters and Holling (2018, p. 45) as the change in a test score as a result of retaking the same test under comparable conditions. The authors report that the effect can also be referred to as retest bias and practice effect. They identify three causes of retest effect from which they form a theoretical basis for the impact of retesting. The first is related to the cognitive construct being measured being enhanced by retesting so learning occurs as a result of practice effects. Secondly, participant anxiety is likely to reduce with repeat testing because of increased familiarity and results in better understanding of the requirements of the test leading to improved results. The final component is the development of test taking strategies or test specific solution strategies which can lead to improvements in performance scores.

They also acknowledge that there is debate about how these factors may impact on test validity but highlight that this may only hold true if cognitive ability is viewed as a stable construct. However, Rust and Golombok (2009, p. 12) indicated that 50% of the variation in intelligence test performance is related to inherited characteristics. This would suggest that test performance is also related to learned characteristics or experience. Scharfen et al., (2018, pp. 56-57) also identify that if an individual develops a solution strategy which assists effective test item solution this is likely to occur within the first or second test time point. The impact of which could account for a test improvement of one third of a standard deviation without any further intervention between timepoints. They also indicate that this improvement has a diminishing impact from the point of first improvement with a plateau being reached by the third time of testing. For this programme of research, any improvement beyond time point two would be related to other factors including simulation tutorials and experience with problem solving the positioning of challenging patients in the clinical environment. It should also be recognised that in the transition from novice to expert, a correct solution is equally as important as the strategy chosen.

It is also important to consider the test –retest time interval with Scharfen et al., (2018, p. 58) suggesting that with longer retest intervals the impact of solution strategies reduces. For this programme of research the test - retest time intervals were determined by the timing of campus based practical workshops during the academic taught timetable and experiential learning in the diagnostic imaging and radiotherapy departemnts. Retest performance scores may be highest at short intervals but gains may decrease with time. However, with longer intervals between testing the impact of practice effects may become more difficult to distinguish from actual changes and within participant variability (Calamia et al., 2012, p. 547). The gains in performance for the radiotherapy and diagnostic imaging cohorts observed in study four, although small, are more likely to be due to actual changes in performance rather than practice effects given the relatively long (18 month) duration of the study.

7.5 Limitations

The findings described above must be viewed within the context of the studies reported. This context relates principally to the quantitative design and the relatively small sample size of the radiotherapy cohort in the experimental studies. While these do not necessarily compromise the results, their generalisability to and interpretation for other radiotherapy programmes or any complex environment where well developed 3-D spatial visualisation skill is required, should be done with caution.

7.5.1 Research approach and design

While the individual studies were predominantly quantitative in design, the overall programme of research was positioned within the worldview of pragmatism (research findings informing educational and clinical practice) and employed a mixed QUANT + qual model design. The conceptual framework for this programme of research was to determine if the measurement of the 3-D spatial visualisation of radiotherapy learners could identify them in terms of their relative baseline spatial visualisation skill. This necessitated a predominantly quantitative approach, but also employed qualitative data to add to the understanding of individual spatial visualisation performance scores through exploration of possible influencing factors.

7.5.2 Sample size and impact of missing data

The number of Health Education England commissioned training places limited the recruitment pool and therefore number of potential participants in both phases of this programme of research. Consideration was given to the recruitment of radiotherapy students from other HEI`s, (similar to the sampling strategy employed in the DoH VERT™ evaluation project). However, while this would have increased the potential sample size and therefore power of the studies, each HEI radiotherapy programme tends to have a different clinical placement model in terms of timing and number of hours and therefore its utilisation of VERT™.

During any longitudinal study participation is likely to fluctuate due to participants withdrawing from their programme of study or having other commitments at the time of each of the data collection points. To mitigate this impact, consideration was also given to recruitment of other healthcare science pre-registration students from the host institution. While this may have delivered a wider understanding of the 3-D spatial visualisation skills of learners, it would not have focussed specifically on those skills required in radiotherapy.

7.5.3 Impact of time constraints

Performance differences between males and females, particularly in mental rotation tasks, have been widely reported in the literature, with Peters (2005, p. 176) suggesting that one of the influencing factors may be related to the time allowed to complete the test. In a study to determine the impact of different timing conventions 501 males and 1264 females completed the Vandenberg & Kuse mental 24 item MRT. The test was administered using the recommended timing of three minutes per 12 items and standard scoring of one point for each correct pair identified. The mean number of problems solved correctly was 14.05 (SD = 5.9) for males and 8.96 (SD = 4.4) for females. This difference was reported as significant (F (1), 1764) = 425.6, p <.0001 and an overall effect of gender expressed as Cohen’s d = .97. The study also identified that the number of participants failing to attempt a test item increased as the test progressed. An analysis of the number of participants attempting the final three items (although there is no indication why the focus was just on the last three), showed that 38.7% of males and 19.5% of females attempted solutions for these items.

In a follow up study involving a cohort of 212 students, comprising 88 males (41.5%) and 124 females (58.5%), participants were allocated to one of two timing conditions: T1 (standard timing of three minutes) and T2 (double time of six minutes). The results indicated that the mean number of items in the first half of the test completed by males in the standard timing condition (T1) was 8.04 (67%) while for females it was 6.6 (55%). But under the extended timing condition (T2) the mean number of completed items increased to 11.4 (94.9%) for males and 10.9 (90.9%) for females. The impact of increasing the time available can be demonstrated by analysing the number of participants who attempted the final test item. In timing condition T1 this represented 11% of males and 6.6% of females, while in timing condition T2 this increased to 79.1% of males and 58.6% of females. While females benefitted from having additional time the performance difference was reported as not significant since males also gained from the same benefit of extra time. The analysis of uncompleted items across both phases of this programme of research demonstrated that fewer participants completed test objects towards the end of the test compared to the beginning. While the time allowance could have been increased and was considered, unlimited time is not a luxury that can be afforded in the clinical environment. As Peters (2005, p. 177) also acknowledged, while spatial visualisation skills evolve in response to the environmental demands of a particular task, the environment does not always permit the luxury of unlimited time to complete that task. This is certainly the case in radiotherapy treatment delivery where the increasing demand for radiotherapy services and the increasing use of 3 and 4-D techniques places a constraint on the time available for the mental visualisation of correct tumour target volume alignment in relation to patient position. While automated image matching processes are available for image guided radiotherapy processes, radiotherapy radiographer decision making still needs to be done in a timely manner. There is limited research relating to radiotherapy workflow timings, and where studies are reported they tend to examine the overall time that a patient spends in a treatment room For example, in an analysis of 324 randomly selected radiotherapy treatment sessions, conducted by Van de Werf et al., (2009, p.138) the mean in- room time was reported as 11.6 minutes (SD = 5.9) for conventional 3-D conformal delivery. For more complex treatment delivery, such as a seven field IMRT technique, this time increased to 13.6 minutes (SD = 5.4) and with the addition of IGRT the mean was 17.3 minutes (SD = 6.8). A more recent study by Beech, Burgess and Stratford (2016, p. 207) reported on 1085 treatment sessions over a

400 day period. The mean room occupancy time was found to be 14 minutes (SD = 6), but further analysis of the time between the patient getting on the treatment couch to the initiation of the first radiation beam (equivalent to the positioning time) was seven minutes. These findings would support the decisions made for the overall time limit of eight minutes for the 3-D SVT employed in studies one, two and four in this programme of research. They may also explain the reports of perceived time pressures which contribute to poor clinical experiences for students, reported in chapter one.

7.5.4 Methodological criticisms

The choice of a predominantly quantitative design was informed by the need to place a numerical value on an individual learner's spatial visualisation skill. The longitudinal study (study four) measured performance at four time points over an 18 month period. Between each data collection event, participants had undertaken periods of clinical experience in diagnostic imaging and radiotherapy departments and in the simulation facilities on campus. It would therefore be expected that this experience would have an impact on 3-D SVT performance. While performance gains were reported they were generally small, and different for both the control and experimental groups and the exact nature of the gains, whether from clinical experience or engagement with simulation workshops could not be determined. The decision to use a single control group and a single experimental group may also have had an impact. An alternative approach could have been to employ a four group design as proposed by Solomon (1949, p. 139). It has the advantage of strengthening internal validity by determining the effects of pre-testing an intervention by employing two control groups and two experimental groups. It seeks to determine if a relationship exists between an intervention group who have received a pre-test and those who have not. If a difference is detected between the two groups at post-test, this would indicate that the pre-test had an influence on the intervention. This may also mean that results would not be generalisable in all situations. As study four sought to determine if change in 3-D spatial visualisation could be detected over time, identifying the reasons for that change was not a primary aim. However in future pre and post intervention studies to determine the impact of any additional visualisation training, a four group design should be considered.

To ascertain whether engagement with previous recreational activities with a spatial component would have a bearing on baseline test performance and development, a qualitative questionnaire to derive quantitative data was employed. It is recognised that the use of questionnaires have the potential for response bias, in that, the responses of participants are based on what they think the researcher wants to hear and so any findings need to be treated with some caution. An alternative mixed methods design could have incorporated structured focus group interviews to gain a deeper understanding of individual perceptions of spatial visualisation skill and test performance via the participant voice. However, as the first data collection time point for the experimental phase took place within the first two weeks on campus and prior to involvement in radiotherapy visualisation workshops, participants would have a limited understanding of the role of 3-D spatial visualisation.

7.5.5 Validity and reliability

When Vandenberg and Kuse (1978, pp. 601 – 602) described the development of the MRT, they reported Pearson Product – Moment correlations with the card rotation test of .62 and for Shepard & Metzler identical blocks .68. For other spatial tests, such as hidden figures and form boards these values are lower at .4 and .41 respectively. While not reported by the authors, the lower values observed for some tests may relate to them being 2-D tests. In an early study involving 3264 teenagers (≥ 14 years of age) and adults, a Kuder-Richardson 20 (K- 20) coefficient of .88 was reported, while in a separate study of 336 subjects the K-R 20 was reported as .83 with a test retest reliability at one year of .70, although it should be noted that there was no report of the demographic profile for the participants in this second study.

While the initial validity reported for the MRT was determined by correlations with what might be argued to be tests for general spatial skills, it has been widely adopted in visualisation studies since its introduction over four decades ago. When Cohen and Hegarty (2007, p. 181) reported on the development of the SBST, in a study of 59 psychology students, they also employed the MRT and the Visualization of Views Test. They reported that the performance in both tests was highly correlated ($r = .47$) and using averaged score from both tests, which they referred to as the spatial score, reported a correlation of .5 ($p < .01$) with all types of test figures and cutting planes in the SBST. They also reported a split half Cronbach Alpha for internal consistency for the 29 test items in the SBST of 0.86, which they referred to as a satisfactory.

In the pilot phase of this programme of research the consistency of the results achieved in the paper based and online test platforms, the performance of the 10 participants who completed both versions of the test was analysed by comparing the number of correct answer choices in the MRT and SBST subcomponents. Performance in each subcomponent was compared using Cronbach's alpha coefficient. For the MRT and the SBST paper and online subcomponents, Cronbach's alpha for both was .6. A comparison between the paper based MRT and SBST items demonstrated a Cronbach alpha of .52 and for the online items the value was .76. Coefficient values above 0.7 indicate good consistency as identified by Walker and Almond (2010, p. 86). While Tavakol & Dennick (2011, p.54) identified different reports about the acceptable values of alpha, ranging from 0.70 to 0.95, they refer to earlier work by Bland and Altman (1997, p.572) who suggested that for scales which are used as research tools to compare groups, alpha values between 0.7 to 0.8 can be regarded as satisfactory. This is the level reported by DeVon et al., (2007, p. 160) as being acceptable for a new scale, although as Bailey et al., (2018, p. 345) reported, there is no definitive cut-off point. But it also has to be acknowledged that test-retest performance scores within the same cohort can be expected to change as a result of clinical experience. For any test instrument it is conceivable that every sample may result in a unique set of scores which may result in different reliability coefficients. So across a longitudinal study, such as study four in the experimental phase, it should not be expected that each set of scores will produce similar reliability.

In addition, the small number of participants in the pilot studies is acknowledged. This could have been improved by recruiting student volunteers from other programmes delivered within the School of Health Sciences & Social Work portfolio or the wider University of Portsmouth undergraduate community. While this could have increased the power of the findings, it may have decreased the richness of detail and understanding of the spatial visualisation characteristics of diagnostic imaging and radiotherapy students.

7.6 Contribution to existing knowledge

This programme of research has shown that it is possible to measure students' spatial visualisation skill at the start of a programme of study and band performance into high intermediate or low categories. Using this method, it is possible to apply different additional training support to the three levels of skill and tailor this to individual needs.

This would support development of best practice in the use of VERT™ in the classroom. The research questions formulated for this programme of research had not previously been addressed in the published literature. One of the recommendations of the Department of Health (England) final project report on VERT™ was to measure the inherent spatial ability of students to identify those most likely to benefit from experience in VERT™. If this is being done, it is being done at individual HEI level and has not been reported in the wider radiotherapy educational research community. The research findings reported in this thesis can inform and support the development of specific spatial visualisation syllabus content. It is proposed, therefore, that gaining an understanding of the spatial visualisation performance of individual learners will lead to a more focussed use of VERT™ which goes beyond its current role. This in turn will lead to the development of additional learning outcomes for VERT ™ in the academic and clinical settings which have a clear and direct link to the HCPC Standards of Proficiency relating to spatial ability.

7.6 Contribution to practice

The findings of this programme of research indicate that gains in spatial visualisation skill are achieved over time as a result of experiential learning in the clinic combined with practical workshops in the simulation environment. It is posited that the development and introduction of a more focused, integrated approach to the development of spatial visualisation skill within VERT™ will lead to more effective use of clinical placement time in supporting 3-D soft tissue image interpretation and reduce the ad-hoc, opportunistic learning of spatial visualisation skills in the clinic.

This research is distinctive because it has developed a composite online 3-D spatial visualisation skill measurement tool based on previously validated paper based test instruments. The measurement tool contains 3-D mental rotation and cross sectional objects which require the use of visualisation, perception and mental transformation (rotation) skills similar to those required when mentally visualising the impact of patient position on the position of internal anatomical structures. It has demonstrated that baseline spatial visualisation skill can be determined, learners can be categorised as having high, intermediate and low spatial visualisation skill and that it does change over time, thus confirming that it can be developed through practice. The ability to identify learners who may have difficulty in visualising complex principles because of less well developed spatial visualisation skills at the start of their radiography education provides the opportunity to identify their specific developmental needs and to develop bespoke visualisation activities to support that development.

7.7 Implications for future practice

The current radiotherapy curriculum is designed to support the declarative knowledge base and procedural understanding of complex radiotherapy processes and must still be delivered. However the findings of this programme of research suggest this delivery may be enhanced by making the links between the spatial visualisation components of the process more overt. This can be achieved by focusing attention on the pattern recognition of internal normal and altered anatomical position based on external cues, beam path geometry and direction and 3-D interpretation of soft tissue data from computed tomography. The analysis of an individual learner's test performance in the subcomponents of mental rotation and visual perception will also provide insight into the specific 3-D visualisation development requirements of each student.

Within the Centre for Simulation in Health at the University of Portsmouth there are a number of platforms that can support the development of 3-D spatial visualisation for those learners whose test performance score has identified that they may benefit from additional 3-D spatial visualisation support. Specifically, these may encompass the use of virtual reality anatomy platforms such as Anatomage which can support the visualisation of relational anatomy and CT pattern recognition. The VERT™ platform supports the viewing of a tumour target volume from different angles which would assist the development of mental rotation skills. In combination with the beam's eye view function the visualisation of relationships and the development of visual penetrative ability can be supported.

7.8 Future development of this programme of research

The purpose of this programme of research was to develop a body of knowledge relating to the spatial visualisation characteristics of radiography students, with a specific focus on pre-registration radiotherapy learners in a single institution. The aim was to develop a method for measuring spatial visualisation in order to identify those learners who may have less well developed spatial visualisation skills when joining a programme of study. This aim was in line with one of the recommendations of the DoH (England) VERT™

project report reported above. The desire, therefore, would be a continuation of this work both within radiotherapy and beyond, in collaboration with other academic colleagues involved in health and care education. This would offer further insight for other educators and researchers who may wish to identify those learners who would benefit from additional support with the visualisation and mental modelling of complex spatial concepts. Therefore, by repeating study four, which measured baseline visualisation skill and tracked change over time, with larger populations, both within this and other institutions, should offer more generalisable results.

7.9 Developing the spatial visualisation curriculum

Practical workshops using the VERT™ platform, embedded in the clinical learning modules, focus mainly on the development of technical and motor skills and the overarching conceptual framework for radiotherapy. This is done within semi rigid lesson plans with standard syllabus content and learning outcomes. While the order of delivery may be adjusted to match the speed of learning of each group or to incorporate themes that have been introduced via didactic delivery within academic units, there is no structured method for identifying those learners with less well developed 3-D spatial visualisation skill. Practical workshop facilitators are reliant on an individual student informing them that they are not understanding or able to visualise a particular concept or process.

If the educator does not have an insight and understanding of an individual students` visualisation skill then there is a strong possibility that those students with less well developed SVS may experience a theory practice gap in understanding and applying these principles. Identifying learners at risk of not understanding during the early stages of their radiotherapy education would lead to the development of a more strategic approach to the use of VERT™ as a technology enhanced active learning platform. Embedding focussed visualisation exercises incorporating, for example, 3-D pattern recognition into the early stages of delivery of the standard curriculum would also facilitate the use of additional interventions and support for those learners with low spatial visualisation skill. These are the students identified by Cohen and Hegarty (2007, p. 182) who are likely to have less well developed spatial visualisation skills.

7.10 Conclusions and recommendations

The measurement of individual learner performance in spatial visualisation tasks, carried out during this programme of research, has demonstrated that it is possible to differentiate between those who have, high, intermediate and low performance at the start of their pre-registration radiographic studies. By analysing subcomponent mental rotation and cutting planes scores has provided an insight into specific areas of potential visualisation difficulty when translated to real case positioning accuracy during the radiotherapy process. This additional insight can support the development of an updated spatial visualisation curriculum and syllabus content based on individual learner needs. This approach will enhance the development of the spatial visualisation skills required to support current and future radiotherapy practice and goes beyond the standard one size fits all lesson plans currently in place. The findings will also add to the radiotherapy communities understanding of VERT™ best practice.

It should be recognised that the finding of less well developed 3-D spatial visualisation skill at the start of a programme of study will not automatically result in a graduate with low spatial visualisation skill. It may however mean that an individual with a low performance score at the outset may take longer to build those mental model relationships between the position of organs in relation to a tumour and the intended radiation beam path. As Veurink and Sorby (2019, p. 160) found in their study of first year engineering students, identifying learners with less well developed spatial visualisation skill and providing additional visualisation activities contributes to success. The findings of the longitudinal study (study four) demonstrated an improvement in performance over time. Given the relatively long time period between baseline and end of study testing, the changes seen may be attributed to actual changes in spatial visualisation rather than test-retest practice effects.

Given the conclusions identified above, there exists a need to conduct further research to determine the spatial visualisation skills of future radiotherapy students. If the results are replicable, then the ability of educators to determine which learners may be at risk of less well developed skills at the commencement of studies will be proven. Evidence from the previous research would suggest that spatial visualisation skill is part innate and part developmental. The measurement of baseline entry level skill can, therefore, assist in the identification of those individuals who may benefit from additional focussed tutorials. Future research themes which are focussed on the impact of spatial visualisation exercises, the employment of platforms such as the Anatomage virtual dissection table for anatomical relationships and VERT™ for beam path visualisation should be developed. If these were to incorporate eye tracking studies, patterns of visual information acquisition could be determined and may further refine the focus of this additional spatial visualisation learning support. In doing so, the findings of this programme of research would transferable to other health and care professions where a high level of spatial visualisation is required in complex clinical situations. The recommendations developed from these conclusions are summarised below in table 7.2.

Table 7.2: Summary of recommendations

1	Conduct a large scale collaborative study to benchmark appropriate values for the banding of high, intermediate and low spatial visualisation skill learners in radiotherapy
2	Develop performance metrics and design a new practical task using VERT™ to determine the impact of additional spatial visualisation training and transfer to practice
3	Conduct a review of the 3-D SVT online test platform and implement in other health science subject areas where spatial visualisation skill is a fundamental component of complex 3-D practice
4	Develop additional mental rotation and cross sectional cutting plane activities to supplement tutorials and practical simulation workshops
5	Develop a screening tool to evaluate the effectiveness of these activities
6	Identify opportunities for integration of the VERT™ platform with other virtual reality visualisation environments to support the development of 3-D spatial visualisation skill

7.11 The contributions of this programme of research to self-development: a reflection

Completing this programme of research and study has enabled considerable personal development and insight. The researcher entered the radiotherapy profession with a professional diploma, progressing after consolidation of clinical skills to a higher diploma. The intention at that point was to move into an academic career by studying for a professional body teaching diploma. However, two events occurring almost simultaneously resulted in a different career direction. The first was the transfer of practical training and academic education from a predominantly local, hospital based provision, to higher education intuitions and graduate entry to the profession during the early 1990`s. This arose as a result of the development of more complex technology and the expansion of the boundaries of the therapeutic radiography profession. The second was an appointment as an operational manager with key responsibilities for the supervision, education and assessment of learners in the clinic. Following a break from studying for several years and not wishing to get left behind as the new technology to support IMRT became widely available in the researchers department, a Master's Degree programme in Radiotherapy and Oncology was embarked upon. By this time extensive clinical experience was being used to support both undergraduate and postgraduate teaching. The idea of studying at a level higher combined with the development of practice to support advanced radiotherapy techniques and the ability to pass this experience on to others was a challenge to be relished and it was at this point that a full time academic career was considered. Shortly after this academic career commenced, the researcher attended a Department of Health (England) briefing sessions ahead of the introduction of VERT™. Looking back through the notes from that meeting, it was evident that a map of the platform's potential contribution to the existing curriculum was already emerging. Prior to the installation of the platform, the researcher was appointed the VERT™ lead and has represented the University as a member of the international user group since its inception. This has enabled the sharing of research interests within the wider VERT™ community.

The idea of moving from Masters level study to Doctoral was research mooted at an annual professional development review and was immediately taken up as a challenge, something that even five years previously would never have been considered. However the start of the research journey commenced many years before embarking on this programme of research. The technological inventions and developments through the 1980`s drove the development and application of advanced 3-D radiotherapy techniques through the 1990`s and continue to this day. Driven partly by curiosity and partly by a need to provide an explanation of these increasingly complex techniques from first principles to learners in the clinic, an interest in approaches to and styles of learning was developed. The arrival of VERT™ delivered new opportunities to visualise and build mental models of complex treatment delivery scenarios. Initial thoughts relating to the use of VERT™ as a research platform centred on its potential as an alternative to the assessment of clinical skills in the often pressured environment of the radiotherapy clinical environment. However during the development of the conceptual framework for the research, the emphasis changed to a focus on how individuals engage with and learn in virtual reality environments. If a learner cannot engage with virtual reality platforms then the assessment framework may disadvantage that student and could therefore be inequitable. This led to thinking about visualisation skills in general and then homing on how spatial visualisation skill may be measured and developed.

The researchers' career has evolved from being a clinician interested in education into an educator interested in the development of learners' clinical practice through educational research and finally to a researcher using the evidence base to inspire the next generation of learners to think about the potential role of advanced technology in both clinical practice and education. To support this role I intend to continue and develop my role as a researcher and look forward to encouraging and supporting others to do likewise.

References

Abel, E., Silander, E., Nyman, J., Bove, M., Johansson, L., Björk-Eriksson, T., & Hammerlid, E. (2017). Impact on quality of life of IMRT versus 3-D conformal radiation therapy in head and neck cancer patients: A case control study. *Advances in Radiation Oncology*, 2(3), 346-353.

Abrams, J., Conley, B., Mooney, M., Zwiebel, J., Chen, A., Welch, J. J., Takebe, N., (....), & Doroshow, J. (2014). National Cancer Institute's precision medicine initiatives for the new National Clinical Trials Network. *American Society of Clinical Oncology Educational Book*, 34(1), 71-76.

Agee, J. (2009). Developing qualitative research questions: a reflective process. *International Journal of Qualitative Studies in Education*, *22*(4), 431-447.

Akasah, Z. A., & Alias, M. (2010). Bridging the spatial visualisation skills gap through engineering drawing using the whole-to-parts approach. *Australasian Journal of Engineering Education*, 16(1), 81-86.

Alias, M., Black, T.R., & Gray, D.E. (2002). Effect of instruction on spatial visualisation ability in civil engineering students. *International Education Journal*, 3(1), 1-12.

Alinier, G. (2007). A typology of educationally focused medical simulation tools. *Medical Teacher*, *29*(8), e243-e250.

Almhagen, E., Boersma, D. J., Nyström, H., & Ahnesjö, A. (2018). A beam model for focused proton pencil beams. *Physica Medica*, 52, 27-32.

Anders Ericsson, K. (2008). Deliberate practice and acquisition of expert performance: a general overview. *Academic Emergency Medicine*, 15(11), 988-994.

Anders Ericsson, K., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, 47(1), 273-305.

Andersen, K. W., & Siebner, H. R. (2018). Mapping dexterity and handedness: recent insights and future challenges. *Current Opinion in Behavioral Sciences*, 20, 123-129.

Appleyard, R., & Coleman, L. (2010). Virtual environment for radiotherapy training (VERT) Final project report–Executive summary; Department of Health for England, Cancer Action Team. London: Society and College of Radiographers.

Auer, T., Schwarcz, A., Aradi, M., Kalmár, Z., Pendleton, C., Janszky, I., (....) & Janszky, J. (2008). Right–left discrimination is related to the right hemisphere. *Laterality*, 13(5), 427-438.

Baartman, L. K., & De Bruijn, E. (2011). Integrating knowledge, skills and attitudes: Conceptualising learning processes towards vocational competence. *Educational Research Review*, 6(2), 125-134.

Baddeley, A. (1998). Recent developments in working memory. *Current Opinion in Neurobiology*, 8(2), 234-238.

Baddeley, A. D., & Hitch, G. J. (2000). Development of working memory: Should the Pascual-Leone and the Baddeley and Hitch models be merged? *Journal of Experimental Child Psychology*, 77(2), 128-137.

Bailey, S. K., Neigel, A. R., Dhanani, L. Y., & Sims, V. K. (2018). Establishing measurement equivalence across computer and paper based tests of spatial cognition. *Human Factors*, 60(3), 340-350.

Barker, C., Lowe, M., & Radhakrishna, G. (2019). An introduction to proton beam therapy. *British Journal of Hospital Medicine*, 80(10), 574-578.

Beavis, A. W., & Ward, J. W. (2014). The use of a virtual reality simulator to explore and understand the impact of Linac mis-calibrations. *Journal of Physics Conference Series* 489 p.012086 https//: dx.doi:10.1088/1742-6596/489/1/012086

Beech, R., Burgess, K., & Stratford, J. (2016). Process evaluation of treatment times in a large radiotherapy department. *Radiography*, 22(3), 206-216.

Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., Pitkin, R., Rennie, D., Schulz, K.F., Simel, D. (....), & Stroup, D. F. (1996). Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *Journal of the American Medical Association*, 276(8), 637-639.

Benner, P. (1984). From novice to expert. Menlo Park: Addison-Wesley

Berney, S., Bétrancourt, M., Molinari, G., & Hoyek, N. (2015).How spatial abilities and dynamic visualizations interplay when learning functional anatomy with 3D anatomical models. *Anatomical Sciences Education*, 8(5), 452-462.

Berthelsen, A. K., Dobbs, J., Kjellén, E., Landberg, T., Möller, T. R., Nilsson, P., Specht, L, & Wambersie, A. (2007). What's new in target volume definition for radiologists in ICRU Report 71? How can the ICRU volume definitions be integrated in clinical practice? *Cancer Imaging*, 7(1), 104-116.

Bhide, S. A., & Nutting, C. M. (2010). Recent advances in radiotherapy. *BMC Medicine*, *8* (1), Retrieved from https://bmcmedicine.biomedcentral.com/articles/10.1186/1741-7015-8-25

Biggs, J. B. (2011). Teaching for quality learning at university: What the student does. Maidenhead: McGraw-Hill Education.

Bland, J.M. & Altman, D.G. (1997). Statistics Notes: Cronbach's Alpha*. British Medical Journal*, 314 (7080), p. 572. https://doi:10.1136/bmj.314.7080.572

Böckers, A., Mayer, C., & Böckers, T. M. (2014). Does learning in clinical context in anatomical sciences improve examination results, learning motivation, or learning orientation? *Anatomical Sciences Education*, *7*(1), 3-11.

Bodner, G.M., & Guay, R.B. (1997). The Purdue visualization of rotations test. *The Chemical Educator*, 2(4), 1-17.

Borras, J. M., Lievens, Y., Dunscombe, P., Coffey, M., Malicki, J., Corral, J., Gasparottoh, C., (....), & Grau, C. (2015). The optimal utilization proportion of external beam radiotherapy in European countries: an ESTRO-HERO analysis. *Radiotherapy and Oncology*, 116(1), 38-44.

Bossuyt, P.M., Reitsma, J.B., Bruns, D.E., Gatsonis, C.A., Glasziou, P.P., Irwig, L.M., Moher, D., (….), & Lijmer, J.G. (2003). The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clinical Chemistry*, 49(1), 7-18.

Bourke, B. (2014). Positionality: Reflecting on the research process. *The Qualitative Report*, 19(33), 1-9.

Bradbury-Jones, C., & Alcock, J. (2010). Nursing students as research participants: a framework for ethical practice. *Nurse Education Today*, 30(2), 192-196.

Bragge, P. (2010). Asking good clinical research questions and choosing the right study design. *Injury*, 41, S3-S6.

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper (Ed.), *APA handbook of research methods in psychology, Vol. 2. Research designs* (pp. 57–71). Washington DC: American Psychological Association.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. https://doi: 10.3322

Bridge, P., Appleyard, R. M., Ward, J. W., Philips, R., & Beavis, A. W. (2007). The development and evaluation of a virtual radiotherapy treatment machine using an immersive visualisation environment. *Computers & Education*, 49(2), 481-494.

Bridge, P., & Carmichael, M. A. (2014). Factors influencing radiation therapy student clinical placement satisfaction. *Journal of Medical Radiation Sciences*, 61(1), 45-50.

Bridge, P., Fielding, A., Pullar, A., & Rowntree, P. (2016). Development and initial evaluation of a novel 3D volumetric outlining system. *Journal of Radiotherapy in Practice*, 15(1), 38-44.

Bridge, P., Giles, E., Williams, A., Bøjen, A., Appleyard, R., & Kirby, M. (2017). International audit of virtual environment for radiotherapy training usage. *Journal of Radiotherapy in Practice*, 16(4), 375-382.

Bridge, P., Kirby, M. C., & Callender, J. A. (2019). Evaluating VERT as a radiotherapy plan evaluation tool: comparison with treatment planning software. *Journal of Radiotherapy in Practice, 1-5.* https://doi: 10.1017/S1460396919000797

Bridge, P., & Tipper, D. J. (2011). *CT anatomy for radiotherapy*. Keswick: M&K Update Ltd.

Bunniss, S., & Kelly, D. R. (2010). Research paradigms in medical education research. *Medical Education*, 44(4), 358-366.

Bristow, R. G., Alexander, B., Baumann, M., Bratman, S. V., Brown, J. M., Camphausen, K., Choyke, P., (….), & Harari, P.M. (2018). Combining precision radiotherapy with molecular targeting and immunomodulatory agents: a guideline by the American Society for Radiation Oncology. *The Lancet Oncology*, 19(5), e240-e251.

Burgess, R. G. (Ed.). (2005). *The ethics of educational research*. Lewes: Taylor & Francis.

Calamia, M., Markon, K., & Tranel, D. (2012). Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. *The Clinical Neuropsychologist*, 26(4), 543-570.

Carroll, J. B. (1997). Psychometrics, intelligence and public perception. *Intelligence*, 24(1), 25-52.

Carthey, J. (2003). The role of structured observational research in health care. *BMJ Quality & Safety*, 12(2), 13-16.

Casey, M. B. (1996). Understanding individual differences in spatial ability within females: A nature/nurture interactionist framework. *Developmental Review*, 16(3), 241-260.

Chamunyonga, C., Burbery, J., Caldwell, P., Rutledge, P., Fielding, A., & Crowe, S. (2018). Utilising the virtual environment for radiotherapy training system to support undergraduate teaching of IMRT, VMAT, DCAT treatment planning, and QA concepts. *Journal of Medical Imaging and Radiation Sciences*, 49(1), 31-38.

Chamunyonga, C., Edwards, C., Caldwell, P., Rutledge, P., & Burbery, J. (2018). Exploring the role and application of the deliberate practice concept in radiation therapy. *Journal of Medical Imaging and Radiation Sciences*, 49(3), 237-242.

Cherry, P., & Duxbury, A. (Eds.). (2009). Practical radiotherapy: physics and equipment. Chichester: Wiley-Blackwell.

Cilesiz, S. (2011). A phenomenological approach to experiences with technology: Current state, promise, and future directions for research. *Educational Technology Research and Development*, 59(4), 487-510.

Clem, D., Anderson, S., Donaldson, J., & Hdeib, M. (2010). An exploratory study of spatial ability and student achievement in sonography. *Journal of Diagnostic Medical Sonography*, 26(4), 163-170.

Cohen, C.A., & Hegarty, M. (2007). Sources of difficulty in imagining cross sections of 3D objects. In Proceedings of the Twenty-Ninth Annual Conference of the Cognitive Science Society (pp. 179-184). Austin: Cognitive Science Society.

Cohen, C.A., & Hegarty, M. (2012). Inferring cross sections of 3-D objects: a new spatial thinking test. *Learning and Individual Differences*, 22(6), 868-874.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education*. New York: Routledge.

Constant, M., & Mellet, E. (2018). The Impact of Handedness, Sex, and Cognitive Abilities on Left–Right Discrimination: A Behavioral Study. *Frontiers in Psychology*, https://doi.org/10.3389/fpsyg.2018.00405

Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: theory and application. *The American Journal of Medicine*, 119(2), 7-16.

Cook, C., Cleland, J., & Huijbregts, P. (2007). Creation and critique of studies of diagnostic accuracy: use of the STARD and QUADAS methodological quality assessment tools. *Journal of Manual & Manipulative Therapy*, 15(2), 93-102.

Creswell, J. W. (2009). *Research design. Qualitative, quantitative and mixed methods approaches* (3rd ed.). Thousand Oaks: Sage Publications.

Creswell, J.W & Plano-Clark, V.L. (2011) *Choosing a mixed methods design. Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks: Sage Publications.

Cunha, J. A. M., Flynn, R., Bélanger, C., Callaghan, C., Kim, Y., Jia, X., Chen, Z., & Beaulieu, L. (2019). Brachytherapy Future Directions. *Seminars in Radiation Oncology*. 30 (1) 94-106.

Dalgarno, B., Lee, M.J.W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, 41(1), 10-32.

Damianakis, T., & Woodford, M.R. (2012). Qualitative research with small connected communities: generating new knowledge while upholding research ethics. *Qualitative Health Research*, 22(5), 708-718.

de Jong, T., & Fergusson-Hessler, M.G.M. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31(2), 105-113.

De Lisi, R., & Cammarano, D.M. (1996). Computer experience and gender differences in undergraduate mental rotation performance. *Computers in Human Behavior*, 12(3), 351-361.

Delwiche, F. A. (2013). Mapping the literature of radiation therapy. *Journal of the Medical Library Association*, 101(2), 120-127.

Department of Health. (2007).The National Health Service Cancer Reform Strategy. London: Department of Health.

DeVita, V. T., & Chu, E. (2008). A history of cancer chemotherapy. *Cancer Research*, *68*(21), 8643-8653.

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., Savoy, S.M. & Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship*, 39(2), 155-164.

D`Oliveira, T.C. (2004). Dynamic spatial ability: an exploratory analysis and a confirmatory study. *The International Journal of Aviation Psychology*, 14(1), 19-38.

Donaldson, L.J., Panesar, S.S., & Darzi, A. (2014) Patient-Safety-Related Hospital Deaths in England: Thematic Analysis of Incidents Reported to a National Database, 2010–2012. PLoS Med 11(6): e1001667. https://doi.org/10.1371/journal.pmed.1001667

Doyle, L., Brady, A. M., & Byrne, G. (2009). An overview of mixed methods research. *Journal of Research in Nursing*, 14(2), 175-185.

Duce, N. A., Gillett, L., Descallar, J., Tran, M. T., Siu, S. C. M., & Chuan, A. (2016). Visuospatial ability and novice brachial plexus sonography performance. *Acta Anaesthesiologica Scandinavica*, *60*(8), 1161-1169.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). Manual for kit of factor-referenced cognitive tests. Princeton, NJ: Educational Testing Service.

Elias, S. M., Smith, W. L., & Barney, C. E. (2012). Age as a moderator of attitude towards technology in the workplace: work motivation and overall job satisfaction. *Behaviour & Information Technology*, 31(5), 453-467.

Eliot, J., & Smith, I. M. (1983). An international directory of spatial tests. Windsor: NFER-Nelson Publishing.

Erridge, S., Toy, E., & Campbell, S. (2012). *Radiotherapy for thoracic tumours* (2^{nd} ed., pp 101-124) In P. Hoskin (Ed.), Radiotherapy in practice: external beam radiotherapy. (2^{nd} ed.). Oxford: Oxford University Press.

Evans, J. G., & Dirks, S. J. (2001). Relationships of admissions data and measurements of psychological constructs with psychomotor performance of dental technology students. *Journal of Dental Education*, *65*(9), 874-882.

Flinton, D. (2013). Competency based assessment using a virtual environment for radiotherapy. *Procedia Computer Science*, 25, 399-401.

Flinton, D. (2015). Competency based assessment using virtual reality (VERT): is it a realistic possibility? (Unpublished doctoral thesis). CASS School of Education. University of East London. Retrieved from https://repository.uel.ac.uk/item/8547q

Flinton, D., & White, N. (2009). Preliminary findings on the Virtual Environment for Radiotherapy Training (VERT) system: simulator sickness and presence. *Journal of Radiotherapy in Practice*, 8(4), 169-176.

Fontela, P.S., Pant Pai, N., Schiller, I., Dendukuri, N., Ramsay, A., & Pai, M. (2009) Quality and Reporting of Diagnostic Accuracy Studies in TB, HIV and Malaria: Evaluation Using QUADAS and STARD Standards. PLoS ONE 4(11): e7753. https://doi.org/10.1371/journal.pone.0007753

Ford, E. C., Gaudette, R., Myers, L., Vanderver, B., Engineer, L., Zellars, R., Song, D.Y., (….) & DeWeese, T. L. (2009). Evaluation of safety in a radiation oncology setting using failure mode and effects analysis. *International Journal of Radiation Oncology, Biology & Physics*, 74(3), 852-858.

Franks, K. N., Jain, P., & Snee, M. P. (2015). Stereotactic ablative body radiotherapy for lung cancer. *Clinical Oncology*, 27(5), 280-289.

Gaba, D. M. (2004). The future vision of simulation in health care. *Quality and safety in Health Care*, *13* (Supplement 1), i2-i10.

Geiser, C., Lehmann, W., & Eid, M. (2006). Separating "rotators" from "non-rotators" in the Mental Rotations Test: a multigroup latent class analysis. *Multivariate Behavioral Research*, 41(3), 261-293.

Gegenfurtner, A., Kok, E., Geel, K., Bruin, A., Jarodzka, H., Szulewski, A., & Merriënboer, J. J. (2017). The challenges of studying visual expertise in medical image diagnosis. *Medical Education*, 51(1), 97-104.

Gillan, C., Li, W., & Harnett, N. (2013). Radiation therapist perspectives on cone-beam computed tomography practices and response to information. *Journal of Radiotherapy in Practice*, 12(3), 237-244.

Gillan, C., & Liszewski, B. (2016). Is the Practice of Medical Radiation Technologists Being 'Dumbed Down' by Advancing Technology, Risking Our Obsolescence as a Profession? *Journal of Medical Imaging and Radiation Sciences*, 47(1), 5-8.

Giuliano, K. K., Tyer-Viola, L., & Lopez, R. P. (2005). Unity of knowledge in the advancement of nursing knowledge. *Nursing Science Quarterly*, 18(3), 243-248.

Goldstein, D., Haldane, D., & Mitchell, C. (1990). Sex differences in visual-spatial ability: the role of performance factors. *Memory & Cognition*, 18(5), 546-550.

Goleń, M., Składowski, K., Wygoda, A., Przeorek, W., Pilecki, B., Syguła, M., Maciejewski, B., & Kołosza, Z. (2005). A comparison of two scoring systems for late radiation toxicity in patients after radiotherapy for head and neck cancer*. Reports of Practical Oncology & Radiotherapy*, 10(4), 179-192.

Gorska, R., Sorby, S.A., & Leopold, C. (1998). Gender differences in visualization skills – an international perspective. *Engineering Design Graphics Journal*, 62(3).

Guilford, J. P., & Zimmerman, W. S. (1948). The Guilford-Zimmerman Aptitude Survey. *Journal of Applied Psychology*, 32(1), 24-34.

Green, D., & Appleyard, R. (2011). The influence of VERT™ characteristics on the development of skills in skin apposition techniques. *Radiography*, 17(3), 178-182.

Greenhalgh, T., Thorne, S., & Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews? *European Journal of Clinical Investigation*, 48(6), 12931-12936.

Grix, J. (2002). Introducing students to the generic terminology of social research. *Politics*, 22(3), 175-186.

Hamilton, C. S., & Ebert, M. A. (2005). Volumetric uncertainty in radiotherapy. *Clinical Oncology*, 17(6), 456-464.

Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, 100(1), 57-70.

Hand, C. M., Kim, S., & Waldow, S. M. (2004). Overview of radiobiology. In: Washington, C.M., & Leaver, D. (Eds.), *Principles and practice of radiation therapy* (2nd edn). St Louis: Mosby.

Harrington, J., & White, J. (2017). The late medical effects of cancer treatments: a growing challenge for all medical professionals. *Clinical Medicine*, 17(2), 137-139.

Harris, J., Hirsh-Pasek, K., & Newcombe, N. S. (2013). Understanding spatial transformations: similarities and differences between mental rotation and mental folding. *Cognitive Processing*, 14(2), 105-115.

Harris, P., Snell, L., Talbot, M., Harden, R. M., & International CBME Collaborators. (2010). Competency-based medical education: implications for undergraduate programs. *Medical Teacher*, 32(8), 646-650.

Health & Care Professions Council. (2009). Standards of Education and Training, London: Health & Care Professions Council.

Health & Care Professions Council. (2013). Standards of Proficiency for Radiographers, London: Health & Care Professions Council.

Health & Care Professions Council. (2017). Standards of Education and Training, London: Health & Care Professions Council.

Health & Care Professions Council. (2018). *Statistics gender*. Retrieved from the HCPC website: https://www.hcpc-uk.org/resources/data/2018/registrant-snapshot-november-2018/

Health Education England. (2017). *Cancer workforce plan—phase 1: delivering the cancer strategy to 2021*. Retrieved from the Health Education England website: https://hee.nhs.uk/sites/default/files/documents/Cancer%20Workforce%20Plan.pdf

Health Education England. (2018). *RePAIR (Reducing Pre-registration Attrition and Improving Retention)*. Retrieved from the Health Education England website: https://www.hee.nhs.uk/our-work/reducing-pre-registration-attrition-improving-retention

Hedman, L., Klingberg, T., Enochsson, L., Kjellin, A., & Felländer-Tsai, L. (2007). Visual working memory influences the performance in virtual image-guided surgical intervention. *Surgical Endoscopy*, 21(11), 2044-2050.

Hegarty, M., Keehner, M., Khooshabeh, P., & Montello, D.R. (2009). How spatial abilities enhance and are enhanced by dental education. *Learning and Individual Differences*, 19(1), 61-70.

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*(2), 175-191.

Henry, S. G. (2010). Polanyi's tacit knowing and the relevance of epistemology to clinical medicine. *Journal of Evaluation in Clinical Practice*, 16(2), 292-297.

Hertzog, M.A. (2008). Considerations in determining sample size for pilot studies. *Research in Nursing & Health*, 31,180-191.

Higher Education Statistics Agency (2020). *Who's studying in HE?* Retrieved from https://www.hesa.ac.uk/data-and-analysis/students/whos-in-he

Holch, P., Henry, A. M., Davidson, S., Gilbert, A., Routledge, J., Shearsmith, L., Franks, K., (….) & Velikova, G (2017). Acute and late adverse events associated with radical radiation therapy prostate cancer treatment: A systematic review of clinician and patient toxicity reporting in randomized controlled trials. *International Journal of Radiation Oncology, Biology & Physics*, 97(3), 495-510.

Hopkins, M. G., Callan, S. A., Cant, J. G., & Dempsey, S. (2007). Technical knowledge and skills in undergraduate radiation therapy: The University of Newcastle experience. *Radiographer*, *54*(1), 6-10.

Horii, C. V. (2007). Teaching insights from adult learning theory. *Journal of Veterinary Medical Education*, 34(4), 369-376.

Hoskin, P.J., & Coyle, C. (2011). Introduction. In Hoskin, P., & Coyle, C. (Eds.), *Radiotherapy in practice-brachytherapy* (2nd ed., pp1-3). Oxford: Oxford University Press.

House of Commons Public Accounts Committee. (2006). The NHS Cancer Plan: a progress report. London. The Stationery Office.

Hoyek, N., Collet, C., Rastello, O., Fargier, P., Thiriet, P., & Guillot, A. (2009). Enhancement of mental rotation abilities and its effect on anatomy learning. *Teaching and Learning in Medicine*, *21*(3), 201-206.

Høyer, M., Thor, M., Thörnqvist, S., Søndergaard, J., Lassen-Ramshad, Y., & Muren, L. P. (2011). Advances in radiotherapy: from 2D to 4D. *Cancer Imaging*, 11, s147-s152.

Huk, T. (2006). Who benefits from learning with 3D models? The case of spatial ability. *Journal of Computer Assisted Learning*, 22(6), 392-404.

Independent Cancer Taskforce (2015). *Achieving world-class cancer outcomes: a strategy for England 2015-2020*. Retrieved from the Cancer Research UK website: http://www.cancerresearchuk.org/sites/default/files/achieving_world-class_cancer_outcomes_-_a_strategy_for_england_2015-2020.pdf

International Commission on Radiation Units and Measurements. (1993). ICRU report 50: prescribing, recording and reporting photon beam. Bethesda: ICRU.

International Commission on Radiation Units and Measurements. (2010). ICRU report 83: prescribing, recording and reporting photon beam intensity modulated radiation therapy (IMRT). Oxford University Press.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods*, 18(1), 3-20.

James, S., & Dumbleton, C. (2013). An evaluation of the utilisation of the virtual environment for radiotherapy training (VERT) in clinical radiotherapy centres across the UK. *Radiography*, 19 (2), 142-150.

Jansen, P., & Pietsch, S. (2010). Physical activity improves mental rotation performance. *Creative Education*, 1(1), 58-61.

Jeffreys, M. R. (2015). Jeffreys's Nursing Universal Retention and Success Model: Overview and action ideas for optimizing outcomes A–Z. *Nurse Education Today*, 35 (3) 426-431.

Jeggo, P. A., Pearl, L. H., & Carr, A. M. (2016). DNA repair, genome stability and cancer: a historical perspective. *Nature Reviews Cancer*, 16(1), 35-42.

Jimenez, Y. A., Hansen, C. R., Juneja, P., & Thwaites, D. I. (2017). Successful implementation of Virtual Environment for Radiotherapy Training (VERT) in Medical Physics education: The University of Sydney's initial experience and recommendations. *Australasian Physical & Engineering Sciences in Medicine*, 40(4), 909-916.

Johanson, G. A., & Brooks, G. P. (2010). Initial scale development: sample size for pilot studies. *Educational and Psychological Measurement*, 70(3), 394-400.

Johnson, W., & Bouchard, T.J. (2005). The structure of human intelligence: it is verbal, perceptual and image rotation (VPR), not fluid and crystallized. *Intelligence*, 33(4), 393-416.

Johnson, W., Bouchard, T.J., Krueger, R.F., McGue, M., & Gottesman, I.I. (2004). Just one g: consistent results from three test batteries. *Intelligence*, 32(1), 95-107.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14-26.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2), 112-133.

Julious, S. A. (2005). Sample size of 12 per group rule of thumb for a pilot study. Pharmaceutical Statistics, 4, 287-291.

Joustra, P. E., Kolfin, R., van Dijk, N. M., Koning, C. C. E., & Bakker, P. J. M. (2012). Reduce fluctuations in capacity to improve the accessibility of radiotherapy treatment cost-effectively. *Flexible Services and Manufacturing Journal*, 24(4), 448-464.

Kane, P. (2018). Simulation-based education: A narrative review of the use of VERT in radiation therapy education. *Journal of Medical Radiation Sciences*, 65(2), 131-136.

Kanter, S. L. (2009). Ethical approval for studies involving human participants: academic medicine's new policy. *Academic Medicine*, 84(2), 149-150.

Karnieli-Miller, O., Strier, R., & Pessach, L. (2009). Power relations in qualitative research. *Qualitative Health Research*, 19(2), 279-289.

Katrak, P., Bialocerkowski, A.E., Massy-Westropp, N., Saravana Kumar, V.S., & Grimmer, K.A. (2004). A systematic review of the content of critical appraisal tools. *BMC Medical Research Methodology*, 4:22.

Kaufman, S.B. (2007). Sex differences in mental rotation and spatial visualization ability: can they be accounted for by differences in working memory capacity? *Intelligence*, 35(3), 211-223.

Keehner, M., Lippa, Y., Montello, D. R., Tendick, F., & Hegarty, M. (2006). Learning a spatial skill for surgery: How the contributions of abilities change with practice. *Applied Cognitive Psychology*, 20(4), 487-503.

Keehner, M. M., Tendick, F., Meng, M. V., Anwar, H. P., Hegarty, M., Stoller, M. L., & Duh, Q. Y. (2004). Spatial ability, experience, and skill in laparoscopic surgery. *The American Journal of Surgery*, 188(1), 71-75.

Kennedy, I. (2001). The report of the public inquiry into children's heart surgery at the Bristol Royal Infirmary 1984–1995: learning from Bristol. London: Stationery Office.

Ker, J., & Bradley, P. (2010). Simulation in medical education. *Understanding Medical Education: Evidence, Theory and Practice*, 164-180.

Keszei, A. P., Novak, M., & Streiner, D. L. (2010). Introduction to health measurement scales. *Journal of Psychosomatic Research*, 68(4), 319-323.

Kilminster, S., Cottrell, D., Grant, J., & Jolly, B. (2007). AMEE Guide No. 27: Effective educational and clinical supervision. *Medical Teacher*, 29(1), 2-19.

Kirby, M.C. (2015). Teaching radiotherapy physics concepts using simulation: experience with student radiographers in Liverpool, U.K. *Medical Physics International,* 3(2), 87-93.

Kolb, A. Y., & Kolb, D. A. (2009). Experiential learning theory: A dynamic, holistic approach to management learning, education and development. In S.J. Armstrong & C. V. Fukami (Eds.), *The SAGE handbook of management learning, education and development* (pp. 42 – 68). London: Sage

Kong, A., Hodgson, Y., & Druva, R. (2015). The role of simulation in developing clinical knowledge and increasing clinical confidence in first-year radiography students. *Focus on Health Professional Education: A Multi-disciplinary Journal*, 16(3), 29-44.

Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002). Revising the visualizer-verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction*, 20(1), 47-77.

Kron, T. (2011). Image guidance in the radiotherapy treatment room: Can ten years of rapid development prepare us for the future? *Journal of Radiotherapy in Practice*, 10(2), 71-75.

Lammers, R. L., Davenport, M., Korley, F., Griswold-Theodorson, S., Fitch, M. T., Narang, A. T. Evans, L.V., (….), & Robey, W.C. (2008). Teaching and assessing procedural skills using simulation: metrics and methodology. *Academic Emergency Medicine*, 15(11), 1079-1087.

Llaurens, V., Raymond, M., & Faurie, C. (2009). Why are some people left-handed? An evolutionary perspective. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, 364(1519), 881-894.

Lee, C., & Lowe, G. (2013). Isotopes and delivery systems for brachytherapy. In Hoskin, P., & Coyle, C. (Eds.), *Radiotherapy in practice-brachytherapy* (2nd ed., pp6-24). Oxford: Oxford University Press.

Leeflang, M. M., Deeks, J. J., Gatsonis, C., & Bossuyt, P. M. (2008). Systematic reviews of diagnostic test accuracy. *Annals of Internal Medicine*, 149(12), 889-897.

Leentjens, A. F., & Levenson, J. L. (2013). Ethical issues concerning the recruitment of university students as research subjects. *Journal of Psychosomatic Research*, 75(4), 394-398.

Leong, A., Herst, P., & Kane, P. (2018). VERT, a virtual clinical environment, enhances understanding of radiation therapy planning concepts. *Journal of Medical Radiation Sciences*, 65(2), 97-105.

Linn, M.C. & Petersen, A.C. (1985). Emergence and characterization of sex differences in spatial ability: a meta-analysis. *Child Development*, 56(6), 1479-1498.

Liu, A., Tendick, F. Cleary, K., & Kaufman, C. (2003). A survey of surgical simulation: applications, technology and education. *Presence: Teleoperators and Virtual Environments*, 12(6), 599-614.

Lodico, M. G., Spaulding, D. T., & Voegtle, K. H. (2010). Methods in educational research: From theory to practice (2 ed). San Francisco: Jossey-Bass.

Lohman, D. F. (1979). Spatial Ability: A Review and Reanalysis of the Correlational Literature (No. TR-8). Stanford: Stanford University.

Logan, T. (2015). The influence of test mode and visuospatial ability on mathematics assessment performance. *Mathematics Education Research Journal*, 27(4), 423-441.

Lozano, R. G. (2012). A review of literature: learning conditions of radiation therapists. *Radiation Therapist*, 21(1), 1-11

Luursema, J. M., Buzink, S. N., Verwey, W. B., & Jakimowicz, J. J. (2010). Visuo-spatial ability in colonoscopy simulator training. *Advances in Health Sciences Education*, 15(5), 685-694.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence,* 37, 1-10

McKenzie, A. & Thwaites, D. (2007). Electron beam treatment planning. In P. Mayles, A. E. Nahum, & J. C. Rosenwald (eds). *Handbook of radiotherapy physics: theory and practice* (pp 701-718). London: Taylor & Francis.

McPake, M. (2019). Radiographers' and students' experiences of undergraduate radiotherapy practice placement in the United Kingdom. *Radiography,* 25, 220-226.

Mackay, R. I., Staffurth, J., Poynter, A., & Routsis, D. (2010). UK guidelines for the safe delivery of intensity-modulated radiotherapy. *Clinical Oncology*, 22(8), 629-635.

Mantovani, F., Castelnuovo, G., Gaggioli, A., & Riva, G. (2003). Virtual reality training for health-care professionals. *CyberPsychology & Behavior*, 6(4), 389-395.

Maran, N. J., & Glavin, R. J. (2003). Low-to high-fidelity simulation–a continuum of medical education? *Medical Education*, 37 (Supplement 1), 22-28.

Massagué, J., & Obenauf, A. C. (2016). Metastatic colonization by circulating tumour cells. *Nature*, 529(7586), 298-306.

Mazur, L. M., Marks, L. B., McLeod, R., Karwowski, W., Mosaly, P., Tracton, G., Adams, R.B., (...), & Chera, B. (2018). Promoting safety mindfulness: Recommendations for the design and use of simulation-based training in radiation therapy. *Advances in Radiation Oncology*, 3(2), 197-204.

Meads, C. A., & Davenport, C. F. (2009). Quality assessment of diagnostic before-after studies: development of methodology in the context of a systematic review. *BMC Medical Research Methodology*, 9(1), 3-13.

Merriam, S. B., Johnson-Bailey, J., Lee, M., Kee, Y., Ntseane, G., & Muhamad, M. (2001). Power and positionality: Negotiating insider/outsider status within and across cultures. *International Journal of Lifelong Education*, 20, 405–416.

Middleton, M., Bradford, C., Frantzis, J., Ambler, A., Sisson, T., Montgomerie, D., & Martin, J. (2009). Paperless and paper-based processes in the modern radiotherapy department. *Radiography*, 15(4), 300-305.

Miller, G.E. (1990). The assessment of clinical skills, competence and performance. *Academic Medicine*, 65(9), S63-7.

Mistry, M., Parkin, D. M., Ahmad, A. S., & Sasieni, P. (2011). Cancer incidence in the United Kingdom: projections to the year 2030. *British Journal of Cancer*, 105 (11), 1795-1803.

Mohler, J. L. (2008). The Impact of visualization methodology on spatial problem solutions among high and low visual achievers. *Journal of Industrial Technology*, 24(1), 1-9.

Monahan, J.S., Harke, M.A., & Shelley, J.R. (2008). Computerizing the Mental Rotations Test: are gender differences maintained? *Behavior Research Methods*, 40(2), 422-427.

Monaghan, T. (2015). A critical analysis of the literature and theoretical perspectives on theory–practice gap amongst newly qualified nurses within the United Kingdom. *Nurse Education Today*, 35(8), e1-e7.

Montgomerie, D., Kane, J. P., Leong, A., & Mudie, B. (2016). Enhancing conceptual knowledge: an approach to using Virtual Environment for Radiotherapy Training in the classroom. *Journal of Radiotherapy in Practice*, 15(2), 203-206.

Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative Health Research*, 8(3), 362-376.

Morris, M. and Venkatesh, V. 2000. Age differences in technology adoption decisions: implications for a changing work force. *Personnel Psychology*, 53(2), 375–403

Müller, K. W., Janikian, M., Dreier, M., Wölfling, K., Beutel, M. E., Tzavara, C., Richardson, C. & Tsitsika, A. (2015). Regular gaming behavior and internet gaming disorder in European adolescents: results from a cross-national representative survey of prevalence, predictors, and psychopathological correlates. *European Child & Adolescent Psychiatry*, *24*(5), 565-574.

Mumford, A. (1996). Four approaches to learning from experience. *Employee Counselling Today: The Journal of Workplace Learning*, 8 (5), 22-29.

National Health Service Executive. (2000). The NHS cancer plan: a plan for investment: a plan for reform. Leeds: Department of Health.

National Radiotherapy Advisory Group (2007). Radiotherapy: developing a world class service for England: report to ministers. London: National Health Service.

Neal, A. J., & Hoskin, P. J. (2012). Clinical Oncology: Basic Principles and Practice (4^{th} Edn.), London: CRC Press.

Needham, A., Hutton, D., & Baker, A. (2011). The introduction of lung stereotactic body radiotherapy in the UK...it's now a reality! *Journal of Radiotherapy in Practice*, 11(1), 7-15.

Németh, B. (2007). Measurement of the development of spatial ability by Mental Cutting Test. *Annales Mathematicae et Informaticae*, 34, 123-128.

Newcombe, N., Bandura, M.M. & Taylor, D.G. (1983). Sex differences in spatial ability and spatial activities. *Sex Roles*, 9, 377-386.

Newcombe, N. S. & Terlecki, M. S. (2005). How important is the digital divide? The relation of computer and videogame usage to gender differences in mental rotation ability. *Sex Roles*, *53*(5-6), 433-441.

Newcombe, N.S., & Stieff, M. (2012). Six myths about spatial thinking. *International Journal of Science Education*, 34(6), 955-971.

Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational Research Methods*, 17(4), 372-411.

Ng, C. K., & White, P. (2005). Qualitative research design and approaches in radiography. *Radiography*, 11(3), 217-225.

Nisbet, H., & Matthews, S. (2011). The educational theory underpinning a clinical workbook for VERT. *Radiography*, *17*(1), 72-75.

Ofcom (2017) *Adults media use and attitudes report*. Retrieved from the Ofcom website: https://www.ofcom.org.uk/__data/.../Adults-Media-Use-and-Attitudes-Report-2018.pdf

Office of National Statistics (2019) *Cancer registration statistics, England: 2017*: Retrieved from the ONS website: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancerregistrationstatisticsengland/2017

Oikonomidoy, E., & Wiest, L. R. (2017). Navigating cross-boundary connections in educational research. *International Journal of Research & Method in Education*, 40(1), 53-65.

Okuda, Y., Bryson, E.O., DeMaria Jr, S., Jacobson, L., Quinines, J., Shen, B., & Levine, A.I. (2009). The utility of simulation in medical education: what is the evidence? *Mount Sinai Journal of Medicine*, 76(4), 330-343.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113.

Oliver, S. E., Roman, E., Crouch, S., Bolton, E., & Ferguson, B. (2013). Comment on Cancer incidence in the United Kingdom: projections to the year 2030. *British Journal of Cancer*, 108(5), 1213-1214.

Oluwatayo, J. A. (2012). Validity and reliability issues in educational research. *Journal of Educational and Social Research*, 2(2), 391-400.

Onwuegbuzie, A. J., & Leech, N. L. (2005). On becoming a pragmatic researcher: The importance of combining quantitative and qualitative research methodologies. *International Journal of Social Research Methodology,* 8(5), 375-387.

Onwuegbuzie, A. J., & Leech, N. L. (2007). Sampling designs in qualitative research: Making the sampling process more public. *The Qualitative Report*, *12*(2), 238-254.

Orth, M., Lauber, K., Niyazi, M., Friedl, A. A., Li, M., Maihöfer, C., Schüttrumpf, L., (….), & Belka, C. (2014). Current concepts in clinical radiation oncology. *Radiation and Environmental Biophysics*, 53(1), 1-29.

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review*, 24(1), 27-45.

Pak, R., Czaja, S. J., Sharit, J., Rogers, W. A., & Fisk, A. D. (2008). The role of spatial abilities and age in performance in an auditory computer navigation task. *Computers in Human Behavior*, 24(6), 3045-3051.

Palma, D. (2017). Is it time for new target volumes in radiation oncology? *International Journal of Radiation Oncology, Biology & Physics*, 98(3), 508-510.

Parodi, K. (2017). Imaging in Radiotherapy. *CERN Yellow Reports: School Proceedings*, *1*, 71. edited by R. Bailey, CERN Yellow Reports: School Proceedings, Vol. 1/2017, CERN-2017-004-SP (CERN, Geneva, 2017)

Parsons, T., Larson, P., Kratz, K., Thiebaux, M., Bluestein, B., Buckwalter, J.G., & Rizzo, A.A. (2004). Sex differences in mental and spatial rotation in a virtual environment. *Neuropsychologia*, 42(4), 555-562.

Patel, V.L., Yoskowitz, N.A., Arocha, J.F., & Shortliffe, E.H. (2009). Cognitive and learning sciences in biomedical and health instructional design: a review with lessons for biomedical informatics education. *Journal of Biomedical Informatics*, 42(1), 176-197.

Paterson, A. (2012). Cancer: Implications for pre-registration radiography curricula. *Radiography*, 18(1), 47-50.

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research*, 74(4), 525-556.

Perez, C. A., & Mutic, S. (2013). Advances and future of Radiation Oncology. *Reports of Practical Oncology & Radiotherapy*, 18(6), 329-332.

Peters, M. (2005). Sex differences and the factor of time in solving Vandenberg and Kuse mental rotation problems. *Brain and Cognition*, 57(2), 176-184.

Peters, M., & Battista, C. (2008). Applications of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library. *Brain and Cognition*, 66(3), 260-264.

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test: different versions and factors that affect performance. *Brain and Cognition*, 28(1), 39-58.

Phillips, R., Ward, J.W., & Beavis, A.W. (2005). Immersive visualisation training of radiotherapy training. Proceedings of Medicine Meets Virtual Reality 15. *Studies in Health Technology and Informatics*, 125, pp 491-497.

Pittalis, M., & Christou, C. (2010). Types of reasoning in 3D geometry thinking and their relation with spatial ability. *Educational Studies in Mathematics*, 75(2), 191-212.

Plomin, R., & Petrill, S.A. (1997). Genetics and intelligence: what's new? *Intelligence*, 24(1), 53-77.

Polanyi, M., & Grene, M. (1969). *Knowing and being: essays by Michael Polanyi*. Chicago: University of Chicago Press.

Polgar, S. & Thomas, S. (2013). *Introduction to research in health sciences.* 6th ed. London: Elsevier Churchill Livingston.

Pratt, S., & Adams, C. (2003). How to create a degree course in radiography: a recipe. *Radiography*, 9(4), 317-322.

Pradu, K.L., Starkshall, G., & Mohan, R. (2007). Three-dimensional conformal radiotherapy. In F.M. Khan (Ed.), *Treatment planning in radiation oncology*. Philadelphia: Lippincott Williams & Wilkins. (2nd ed., pp116-141).

Probst, H., Hutton, D., Collins, M. L., & Adams, R. (2014). Response to Editorial in the JRP by Robson, Clark and White 2014 patient safety: the journey towards safer radiotherapy. *Journal of Radiotherapy in Practice*, 13(2), 247-250.

Public Health England. (2018). Guidance: National Diagnostic Reference Levels (NDRLs): 15 November 2018 onwards. Retrieved from the Public Health England website: https://www.gov.uk/government/publications/diagnostic-radiology-national-diagnostic-reference-levels-ndrls/ndrl

Puusa, A., & Eerikäinen, M. (2010). Is tacit knowledge really tacit? *Electronic Journal of Knowledge Management*, 8(3), 307 – 318.

Rafi, A., Anuar, K., Samad, A., Hayati, M., & Mahadzir, M. (2005). Improving spatial ability using a Web-based Virtual Environment (WbVE). *Automation in Construction*, 14 (6), 707-715.

Ramani, S., & Leinster, S. (2008). AMEE Guide no. 34: teaching in the clinical environment. *Medical Teacher*, 30(4), 347-364.

Reedy, G. B., (2015). Using cognitive load theory to inform simulation design and practice.

Clinical Simulation in Nursing, 11(8), 355-360. http://dx.doi.org/10.1016/j.ecns.2015.05.004

Reinders, H. (2010). The importance of tacit knowledge in practices of care. *Journal of Intellectual Disability Research*, 54, 28-37.

Reitsma, J.B., Rutjes, A.W.S., Whiting, P., Vlassov, V.V., Leeflang, M.M.G., & Deeks, J.J. (2009). Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0.* Retrieved from the Cochrane Collaboration website: http://srdta.cochrane.org/

Robson, W., Clark, D., & White, N. (2014). Patient safety: the journey towards safer radiotherapy. *Journal of Radiotherapy in Practice*, 13(2), 129-130.

Round, C.E., Williams, M.V., Kirkby, N.F., Cooper, T., Hoskin, P., & Jena, R. (2013). Radiotherapy demand and activity in England 2006-2020. *Clinical Oncology*, 5(9), 522-530.

Rosenzweig, K. E., Fox, J. L., Yorke, E., Amols, H., Jackson, A., Rusch, V., Kris, M.G., (….) & Leibel, S. A. (2005). Results of a phase I dose-escalation study using three-dimensional conformal radiotherapy in the treatment of inoperable non-small cell lung carcinoma. *Cancer*, 103(10), 2118-2127.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3), 581-592.

Russell, C., & Churches, A. (2010). What do we really want to know about spatial visualization skills among engineering students? In Proceedings of the 21st Annual Conference for the Australasian Association for Engineering Education (p. 567). Engineers Australia.

Rust, J., & Golombok, S. (2009). Modern Psychometrics. Hove: Routledge

Rutherford-Hemming, T. (2012). Simulation methodology in nursing education and adult learning theory. *Adult Learning*, 23(3), 129-137.

Salami, S. S., Obedian, E., Zimberg, S., & Olsson, C. A. (2016). Urinary quality of life outcomes in men who were treated with image-guided intensity-modulated radiation therapy for prostate cancer. *Advances in Radiation Oncology*, 1(4), 310-316.

Sandler, H. M., McLaughlin, P. W., Ten Haken, R. K., Addison, H., Forman, J., & Lichter, A. (1995). Three dimensional conformal radiotherapy for the treatment of prostate cancer: low risk of chronic rectal morbidity observed in a large series of patients. *International Journal of Radiation Oncology, Biology & Physics*, 33(4), 797-801.

Sandoval, W. (2014). Science education's need for a theory of epistemological development. *Science Education*, 98(3), 383-387.

Salthouse, T. A., Babcock, R. L., Skovronek, E., Mitchell, D. R., & Palmon, R. (1990). Age and experience effects in spatial visualization. *Developmental Psychology*, 26(1), 128-136.

Scotland, J. (2012). Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms. *English Language Teaching*, 5(9), 9-16.

Scharfen, J., Peters, J. M., & Holling, H. (2018). Retest effects in cognitive ability tests: A meta-analysis. *Intelligence*, 67, 44-66.

Schuwirth, L. W., & Van Der Vleuten, C. P. (2006). A plea for new psychometric models in educational assessment. *Medical Education*, 40(4), 296-300.

Senan, S., De Koste, J. V. S., Samson, M., Tankink, H., Jansen, P., Nowak, P. J., Kroll, A.D.G., (….), & Lagerwaard, F. J. (1999). Evaluation of a target contouring protocol for 3D conformal radiotherapy in non-small cell lung cancer. *Radiotherapy and Oncology*, 53(3), 247-255.

Shah, U., & Williams, A. (2010). How to use VERT for interactive CT anatomy for post-registration training. *Imaging & Therapy Practice*. Retrieved from: https://www.sor.org/learning/library-publications/imaging-therapy-practice/july-2010/how-use-vert-interactive-ct-anatomy-post-registration

Shen, Z. (2011). Genomic instability and cancer: an introduction. *Journal of Molecular Cell Biology*, 3(1), 1-3.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171(3972), 701-703.

Sibtain, A., Morgan, A., & MacDougall, N. (Eds.). (2012). Physics for Clinical Oncology. Oxford: Oxford University Press.

Skills for Health. (2011). Statements of competence for radiotherapy. Retrieved from the National Occupational Standards website: https://tools.skillsforhealth.org.uk/competence_search/?search=Radiotherapy

Smittenaar, C. R., Petersen, K. A., Stewart, K., & Moitt, N. (2016). Cancer incidence and mortality projections in the UK until 2035. *British Journal of Cancer*, 115(9), 1147-1155.

Smoker, W.R.K., Berbaum, K.S., Luebke, N.H., & Jacoby, C.G. (1984). Spatial perception testing in diagnostic radiology. *American Journal of Roentgenology*, 143(5), 1105-1109.

Society & College of Radiographers (2013) Code of Professional Conduct. London: Society & College of Radiographers.

Society & College of Radiographers. (2013). Education and career framework for the radiography workforce. London: Society & College of Radiographers.

Solomon, R. L. (1949). An extension of control group design. *Psychological Bulletin*, 46(2), 137 - 150.

Sorby, S.A. (1999). Developing 3-D spatial visualization skills. *Engineering & Graphics Journal*. 63(2), 21-31.

Sorby, S.A. (2007). Developing 3-D spatial skills for engineering students. *Australasian Journal of Engineering Education*, 13(1), 1-11.

Sorby, S. A., & Baartmans, B. J. (2000). The development and assessment of a course for enhancing the 3-D spatial visualization skills of first year engineering students. *Journal of Engineering Education*, 89(3), 301-307.

Srinivasan, K., Mohammadi, M., & Shepherd, J. (2014). Applications of linac-mounted kilovoltage Cone-beam Computed Tomography in modern radiation therapy: A review. *Polish Journal of Radiology*, 79, 181-193.

Srinivasan, M., Hwang, J. C., West, D., & Yellowlees, P. M. (2006). Assessment of clinical skills using simulator technologies. *Academic Psychiatry*, 30(6), 505-515.

Stefanidis, D., Arora, S., Parrack, D. M., Hamad, G. G., Capella, J., Grantcharov, T., Urbach, D.R, (….), & Association for Surgical Education Simulation Committee. (2012). Research priorities in surgical simulation for the 21st century. *The American Journal of Surgery*, 203(1), 49-53.

Stewart-Lord, A. (2016). From education to research: a journey of utilising virtual training. *Journal of Radiotherapy in Practice*, 15(1), 85-90.

Strydom, W. Parker, W & Olivares, M. (2005). Electron beams: physical and clinical aspects. In E. B. Podgorsak (Ed.), *Radiation Oncology Physics: A Handbook for Teachers and Students* (pp 273-299). Vienna: IAEA.

Sutton, K. J., & Williams, A. P. (2007). Spatial cognition and its implications for design. International Association of Societies of Design Research, Hong Kong, China.

Symonds, J. E., & Gorard, S. (2010). Death of mixed methods? Or the rebirth of research as a craft. *Evaluation & Research in Education*, 23(2), 121-136.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55.

Tavakol, M., Mohagheghi, M. A., & Dennick, R. (2008). Assessing the skills of surgical residents using simulation. *Journal of Surgical Education*, 65(2), 77-83.

Techentin, C., Voyer, D., & Voyer, S. D. (2014). Spatial abilities and aging: A meta-analysis. *Experimental Aging Research*, 40(4), 395-425.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research*, 1(1), 77-100.

Ten Cate, O. (2009). Why the ethics of medical education research differs from that of medical research. *Medical Education*, 43(7), 608-610.

Teoh, M., Clark, C. H., Wood, K., Whitaker, S., & Nisbet, A. (2011). Volumetric modulated arc therapy: a review of current literature and clinical use in practice. *The British Journal of Radiology*, 84(1007), 967-996.

Terlecki, M. S., & Newcombe, N. S. (2005). How important is the digital divide? The relation of computer and videogame usage to gender differences in mental rotation ability. *Sex Roles*, 53(5), 433-441.

Terlecki, M. S., Newcombe, N. S., & Little, M. (2008). Durable and generalized effects of spatial experience on mental rotation: Gender differences in growth patterns. *Applied Cognitive Psychology*, 22(7), 996-1013.

Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., Robson, R., (….), & Goldsmith, C. H. (2010). A tutorial on pilot studies: the what, why and how. *BMC Medical Research Methodology*, 10(1), 1-10.

Thariat, J., Hannoun-Levi, J. M., Myint, A. S., Vuong, T., & Gérard, J. P. (2013). Past, present, and future of radiotherapy for the benefit of patients. *Nature Reviews Clinical Oncology*, 10(1), 52-60.

The Royal College of Radiologists, Society and College of Radiographers, & Institute of Physics and Engineering in Medicine. (2008). *On Target: ensuring geometric accuracy in radiotherapy.* London: The Royal College of Radiologists.

The Royal College of Radiologists, Society and College of Radiographers, Institute of Physics and Engineering in Medicine, National Patient Safety Agency, & British Institute of Radiology. (2008). *Towards safer radiotherapy*. London: The Royal College of Radiologists.

Thoirs, K., Giles, E., & Barber, W. (2011). The use and perceptions of simulation in medical radiation science education. *Radiographer: The Official Journal of the Australian Institute of Radiography*, 58(3), 5-11.

Thomas, S. J., Vinall, A., Poynter, A., & Routsis, D. (2010). A multicentre timing study of intensity-modulated radiotherapy planning and delivery. *Clinical Oncology*, 22(8), 658-665.

Thwaites, D. I., & Tuohy, J. B. (2006). Back to the future: the history and development of the clinical linear accelerator. *Physics in Medicine and Biology*, 51(13), 343 –362.

Timmerman, R., Paulus, R., Galvin, J., Michalski, J., Straube, W., Bradley, J., Fakiris, A., (….), & Fowler, J. (2010). Stereotactic body radiation therapy for inoperable early stage lung cancer. *Journal of the American Medical Association*, 303(11), 1070-1076.

Titus, S. & Horsman, E. (2009).Characterizing and improving spatial visualisation skills. *Journal of Geoscience Education*, 57(4), 242-254.

United Kingdom Interactive Entertainment. (2018). *UK video games factsheet*: Retrieved from the UKIE website: http://ukie.org.uk/sites/default/files/UK%20Games%20Industry%20Fact%20Sheet%20June%202018_0.pdf

Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why and how. *Psychology of Learning and Motivation*, 57(2), 147-181.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, 47(2), 599-604.

Van de Werf, E., Lievens, Y., Verstraete, J., Pauwels, K., & Van den Bogaert, W. (2009). Time and motion study of radiotherapy delivery: economic burden of increased quality assurance and IMRT. *Radiotherapy and Oncology*, 93(1), 137-140.

Van Der Land, S., Schouten, A. P., Feldberg, F., Van Den Hooff, B., & Huysman, M. (2013). Lost in space? Cognitive fit and cognitive load in 3D virtual environments. *Computers in Human Behavior*, 29(3), 1054-1064.

Van Merriënboer, J. J., & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical Education*, 44(1), 85-93.

Velez, M. C., Silver, D., & Tremaine, M. (2005). Understanding visualization through spatial ability differences. In Visualization. VIS 05. IEEE (pp. 511-518). IEEE.

Verellen, D., De Ridder, M., Linthout, N., Tournel, K., Soete, G., & Storme, G. (2007). Innovations in image-guided radiotherapy. *Nature Reviews Cancer*, 7(12), 949-960.

Verellen, D., De Ridder, M., & Storme, G. (2008). A (short) history of image-guided radiotherapy. *Radiotherapy and Oncology*, 86(1), 4-13.

Veurink, N. L., & Sorby, S. A. (2019). Longitudinal study of the impact of requiring training for students with initially weak spatial skills. European Journal of Engineering Education, 44(1-2), 153-163. https://doi.org/10.1080/03043797.2017.1390547

Vieira, B., Hans, E., Van Vliet-Vroegindeweij, C., Van de Kamer, J., & Van Harten, W. (2017). Operations research for resource planning and-use in radiotherapy: a literature review. *European Journal of Cancer*, 72, 129-139.

Virtanen, V., & Lindblom-Ylänne, S. (2010). University students' and teachers' conceptions of teaching and learning in the biosciences. *Instructional Science*, 38(4), 355-370.

Vogelstein, B., & Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nature Medicine*, 10(8), 789-799.

Voyer, D., Voyer, S., & Bryden, M.M. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological Bulletin*, 117(2), 250-270.

Voyer, D., & Jansen, P. (2017). Motor expertise and performance in spatial tasks: A meta-analysis. *Human Movement Science*, 54, 110-124.

Walker, J., & Almond, P. (2010). Interpreting statistical findings: a guide for health professionals and students. Maidenhead: McGraw-Hill Education.

Wang, H-C., Chang, C-Y., & Li, T-Y. (2007). The comparative efficacy of 2- versus 3D-based media design for influencing spatial visualization skills. *Computers in Human Behavior*, 23(4), 1943-1957.

Watson, M., Barrett, A., Spence, R.A.J. & Twelves, C.*Oncology* (2nd Edn.), Oxford: Oxford University Press.

Waywell, L. & Bogg, J. (1999).Spatial ability assessment: an aid to student selection for therapy radiography training. *Journal of Radiotherapy in Practice*, 1(02), 89-95.

White, E. & Kane, G. (2007). Radiation medicine practice in the image-guided radiation therapy era: new roles and new opportunities. *Seminars in Radiation Oncology*, 17 (4), 298-305.

White, K. L., Varrassi, E., Routledge, J. A., Barraclough, L. H., Livsey, J. E., McLaughlin, J. & Davidson, S. E. (2018). Does the Use of Volumetric Modulated Arc Therapy Reduce Gastrointestinal Symptoms after Pelvic Radiotherapy? *Clinical Oncology*, 30(1), e22-e28.

Whiting, P., Harbord, R., & Kleijnen, J. (2005). No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Medical Research Methodology*, 5-19 https://doi.org/10.1186/1471-2288-5-19

Whiting, P., Rutjes, A.W.S., Reitsma, J.B., Bossuyt, P.M.M. & Kliejnen, J. (2003). The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology*, 3(1)25 https://doi.org/10.1186/1471-2288-3-25 .

Whiting, P., Weswood, M. E., Rutjes, A. W., Reitsma, J. B., Bossuyt, P. N., & Kleijnen, J. (2006). Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Medical Research Methodology*, 6(1)9 https://doi.org/10.1186/1471-2288-6-9

Woolley, N.N., & Jarvis, Y. (2006). Situated cognition and cognitive apprenticeship: a model for teaching and learning clinical skills in a technologically rich and authentic learning environment. *Nurse Education Today*, 27(1), 73-79.

World Health Organisation. (2018). *WHOQOL: Measuring Quality of Life*. Retrieved from the World Health Organisation website http://www.who.int/healthinfo/survey/whoqol-qualityoflife/en/

Yilmaz, H. B. (2009). On the development and measurement of spatial ability. *International Electronic Journal of Elementary Education*, 1(2), 83-96.

Yilmaz, K. (2013). Comparison of quantitative and qualitative research traditions: Epistemological, theoretical, and methodological differences. *European Journal of Education*, 48(2), 311-325.

Young, J. Q., Van Merrienboer, J., Durning, S., & Ten Cate, O. (2014). Cognitive load theory: Implications for medical education: AMEE guide no. 86. *Medical Teacher*, *36*(5), 371-384.

Zacks, J. M. (2008). Neuroimaging studies of mental rotation: A meta-analysis and review. *Journal of Cognitive Neuroscience*, 20(1), 1-19.

Zacks, J.M., Mires, J., Tversky, B., & Hazeltine, E. (2000). Mental spatial transformations of objects and perspective. *Spatial Cognition and Computation*, 2(4), 315-332.

Zelefsky, M. J., Kollmeier, M., Cox, B., Fidaleo, A., Sperling, D., Pei, X., Carver, B., (….), & Hunt, M. (2012). Improved clinical outcomes with high-dose image guided radiotherapy compared with non-IGRT for the treatment of clinically localized prostate cancer. *International Journal of Radiation Oncology, Biology & Physics*, 84(1), 125-129.

Zelefsky, M.J., Yamada, Y., Fuks, Z., Zhang, Z., Hunt, M., Cahlon, O., Park, J., & Shippy, A. (2008). Long-term results of conformal radiotherapy for prostate cancer: impact of dose escalation on biochemical tumor control and distant metastases-free survival outcomes. *International Journal of Radiation Oncology, Biology & Physics*, 71(4), 1028-1033.

Ziemkiewicz, C., Ottley, A., Crouser, R. J., Chauncey, K., Su, S. L., & Chang, R. (2012). Understanding visualization by understanding individual users. *Computer Graphics and Applications*, 32(6), 88-94.

Appendices

Appendix 1

VERT™ Practical Workshop Session Plans & Learning Outcomes

BSc (Hons) Therapeutic Radiography, Year 1 Introduction to clinical practice 2017-18

Introduction to VERT™ lesson plan for session 1: Academic week 2

Activity	Content	Learning Objectives / Learning Outcomes
Outline of session (15 mins)	What VERT™ does & how we use it. Include safety considerations for screen and controls	
Common terminology & demonstration of functions (15mins)	Tutor led demonstration showing use of OEM hand pendant. Explanation of terminology / definitions: X,Y,Z axes Gantry Collimator Field size and beam definition devices Couch movements (vertical, longitudinal & lateral) Link the above to principles of isocentre Collision detection	Outline the main principles and aims of radiotherapy treatment delivery
Demonstrate alignment of equipment to patient positioning points (skin marks) with gantry at 0^o (20 mins)	Use Pancreas or Lung Virtual Patient to demonstrate use of: 1. Room monitor to display parameters 2. Location of tumour target volume 3. Skin surface 4. Skin marks 5. Lasers 6. Field light & FSD indicator	
Students to practice positioning of virtual patient (50 mins)	Gantry 0^o Position couch so that skin marks align to laser position using vertical, longitudinal & lateral couch movements. Practice rotation of gantry to achieve different beam directions	Demonstrate the psychomotor skills required for the safe operation of a linear accelerator
Questions, review & introduce next session (10 mins		

Work Based Learning 1, VERT™ Practical session 2: Academic week 3

Activity	Content	Learning Objectives / Learning Outcomes
Recap session 1 content (15 mins)	Formative quiz to cover terminology	
Practice positioning pelvis virtual patient as in practical session 1 (2 x 20 mins + 10 mins concluding discussion	Divide group into 2 teams, team 1 to position patient, team 2 to observe and consider process order and steps involved. Swap and repeat. Conclude with comparison & discussion	Understand the principles and aims of accurate patient positioning and radiotherapy treatment delivery
Concept of patient positioning (set-up) errors & correction methods (60 mins)	Tutor led discussion & demonstration of effect of pitch, tilt & roll in relation to patient position. Students to practice positioning virtual patient & resolution of positioning anomalies	Understand impact of incorrect patient position in relation to position of target volume
Review session, questions and introduce content for next session		

Work Based Learning 1, VERT™ Practical session 3: Academic week 4

Activity	**Content**	**Learning Objectives / Learning Outcomes**
Recap principles covered in session 2 (15 mins)	Formative quiz to cover process and positioning terminology	
Introduce session (10 mins)	2 part communication with patient. Randomly divide group into "patients" & "radiographers"	Develop an appropriate level of professional conduct, patient care & communication
Introduce waiting room scenarios "Dealing with the difficult questions" (40 – 60mins)	Ask patients to pick an envelope & read question, students will call patient & check ID, patient will ask question. Tutor led discussion about appropriate interventions	
Revisit concept of set up error & methods for correction (30 mins)	Tutor led review of the effect of pitch, tilt & roll covered in session 2. Contrasts couch shift / adjustment of patient position. Demonstrate using target volume & skin surface	Develop an understanding of patient position correction strategies

Introduction to Radiotherapy unit: Academic week 5

Activity	**Content**	**Learning Objectives / Learning Outcomes**
Outline session (5 mins)	Mix of PPT slides and Virtual Presenter	
Basic principles of field placement (15 mins)	Outline of decision making considerations required for determining optimum beam direction	Describe and implement simple treatment prescriptions & treatment plans
Demonstrate variety of delivery techniques & field arrangements (45 mins)	Use pelvis phantom with field overlays from VERT™ virtual presenter	
Terminology test (20 mins) & close with Q`s		

Introduction to Radiotherapy unit: Localisation & Patient positioning: Academic week 6

Activity	Content	Learning Objectives / Learning Outcomes
Outline session (5 mins)		
Basic principles of field placement / terminology (10 mins)	Formative test	
Localisation methods & terminology	Combination of MS PowerPoint & VERT™ visualisation	Understand & explain the principles of tumour localisation procedures for radiotherapy planning
Importance of and application of surface markings		
Introduce concept of 2-D & 3-D planning		
Patient Immobilisation		
Differentiate equipment and patient set up tolerances (15 mins)	Use pelvis virtual patient with set up errors	Understand the principles underpinning the safe administration of radiation

Work Based Learning 1, VERT™ Practical: session 4: Academic week 7

Activity	Content	Learning Objectives / Learning Outcomes
Formative test (20mins)	To cover terminology / principles of radiotherapy treatment delivery	
Recap basic principles of field placement , patient position, immobilisation(30 mins)	Demonstrate a range of IMRT & VMAT plans (Pelvis & Head & Neck)	Transfer and relate theoretical knowledge to a range of radiotherapy clinical procedures
Wash-up (50 mins)	Student led Q & A`s plus final guided practice	

BSc (Hons) Therapeutic Radiography, Year 2 Introduction to Skin Apposition Techniques 2016-17

Work Based Learning 2, VERT™ Practical: session 1: Academic week 1

<table>
<tr><th>Activity</th><th>Content</th><th>Learning Objectives / Learning Outcomes</th></tr>
<tr><td>Introduction to session & reflections from year 1 clinical experience (20 mins)</td><td>Students to identify areas of practice where development is needed</td><td rowspan="2">Employ the appropriate knowledge, skill and understanding to identify and undertake therapeutic radiography procedures at level 5.

Relate and transfer theoretical knowledge to therapeutic radiography practice.</td></tr>
<tr><td>Student led (guided) practice (90 mins)</td><td>To cover the areas identified from the exercise above</td></tr>
</table>

Work Based Learning 2, VERT™ Practical: sessions 2 - 4: Academic weeks 2-4

Activity	Content	Learning Objectives / Learning Outcomes
Skin Apposition Techniques (10 mins)	Tutor led introduction to the skin apposition technique	To develop an understanding of the principles of the skin apposition treatment delivery method
Small group (3-4 students) practice (45 mins per group)	Students practice a range of positioning tasks using virtual patient	To practice the motor and visualisation skills required to achieve accurate patient and applicator positioning for the skin apposition technique

Appendix 2

Search strategy

Search strategy

Questions

How can spatial visualisation skill be measured?

Does spatial visualisation change over time?

PICO framework

Population	Intervention	Comparison	Outcomes
Diagnostic Imaging & Radiotherapy students	Measurement	Other students	Performance score Development (change over time)

Search terms (Key words OR Thesaurus terms)

Spatial Visualisation Spatial Perception Visualisation	Assessment Testing Measurement	College and higher education	Development Change Training

Search conducted: 22nd – 23rd January 2011

Initial filters:

Dates: January 1st 1970 – December 31st 2010

Sources: Journal articles

Study designs: Primary research studies (randomised controlled, case controls and cohort studies)

Search mode and search expanders

English language

Full text articles (online and print)

Search within full text articles

Peer reviewed

Automatically remove duplicates from each search combination

Combination of search terms with BOOLEAN operators

Spatial visualisation AND measurement OR assessment AND College / Higher education students= 209

Spatial perception AND mental rotation of 3-D Objects AND visualisation = 65

Spatial visualisation AND measurement AND assessment = 84

College / Higher education Students AND visualisation AND training OR development = 88

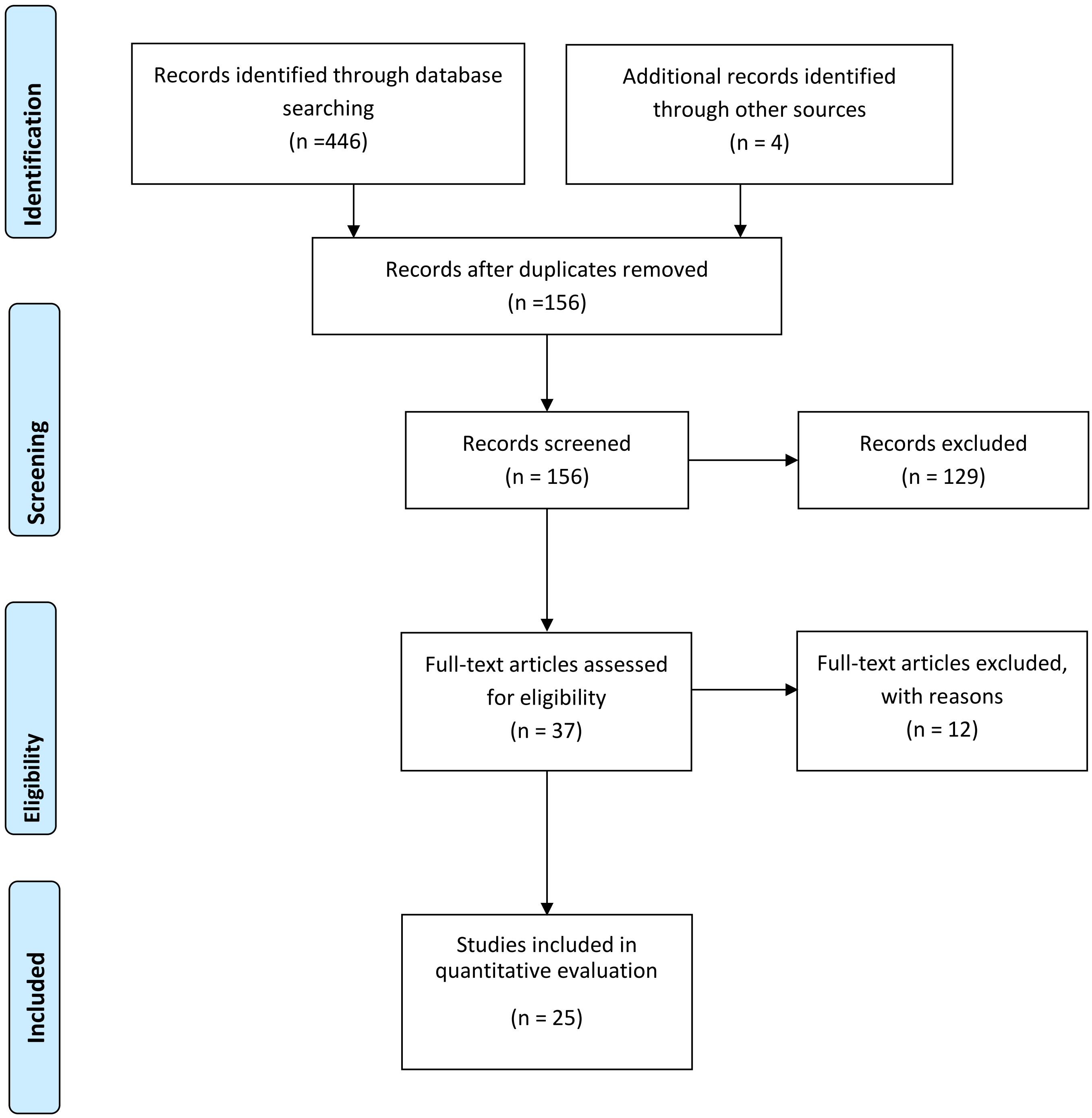
Reasons for excluding records at screening (n = 290):

Duplicates from other searches, wrong age group (children), non-student, health assessment and patients with pathology / cognitive decline, measurement of multiple intelligences / learning styles, conference content lists / abstracts, geosciences / geographic modelling and topography, wayfinding and map reading, maths and reading skills assessment

Reasons for excluding records at eligibility (n = 12)

Spatial memory and WMC processing, children, maths learning, neural function assessment

PRISMA 2009 Flow Diagram



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

Appendix 3

(a) STARD & QUADAS quality checklist item comparison

(b) Justification for selection of checklist items for the evaluation of spatial visualisation testing literature

(c) Modified QUADAS / STARD checklist for SVS performance studies

(d) Completed QUADAS / STARD checklists for SVS performance studies

Appendix 3(a): STARD & QUADAS quality checklist item comparison

STARD Checklist Items	Matching QUADAS Checklist Item
TITLE /ABSTRACT/KEYWORDS	
1 Identify the article as a study of diagnostic accuracy (recommend MeSH heading 'sensitivity and specificity').	
INTRODUCTION	
2 State the research questions or study aims, such as estimating diagnostic accuracy or comparing accuracy between tests or across participant groups.	
METHODS Describe Participants	
3 The study population: The inclusion and exclusion criteria, setting and locations where the data were collected.	2. Were selection criteria clearly described?
Describe Recruitment	
4 Was recruitment based on presenting symptoms, results from previous tests, or the fact that the participants had received the index tests or the reference standard?	1. Was the spectrum of patients' representative of the patients who will receive the test in practice? 5. Did the whole sample or random selection of the sample receive verification using a reference standard of diagnosis?
5 Participant sampling: Was the study population a consecutive series of participants defined by the selection criteria in item 3 and 4? If not, specify how participants were further selected.	
6 Data collection: Was data collection planned before the index test and reference standard were performed (prospective study) or after (retrospective study)?	
TEST METHODS	
7 The reference standard and its rationale	7. Was the reference standard independent of the index test (i.e., the index test did not form part of the reference standard)? 8. Was the execution of the index test described in sufficient detail to permit its replication? 9. Was the execution of the reference standard described in sufficient detail to permit its replication?
8 Technical specifications of material and methods involved including how and when measurements were taken, and/or cite references for index tests and reference standard.	
9 Definition of and rationale for the units, cut-offs and/or categories of the results of the index tests and the reference standard.	
10 The number, training and expertise of the persons executing and reading the index tests and the reference standard.	
11 Whether or not the readers of the index tests and reference standard were blind (masked) to the results of the other tests and describe any other clinical information	10. Were the index test results interpreted without knowledge of the results of the reference test? 11. Were the reference standard results

available to the readers.	interpreted without knowledge of the results of the index test?
STATISTICAL METHODS	
12 Methods for calculating or comparing measures of diagnostic accuracy, and the statistical methods used to quantify uncertainty (e.g., 95% confidence intervals).	
13 Methods for calculating test reproducibility, if done.	
RESULTS REPORT: Participants	
14 When study was done, including beginning and ending dates of recruitment.	
15 Clinical and demographic characteristics of the study population (e.g., age, sex, spectrum of presenting symptoms, comorbidity, current treatments, recruitment centres).	
16 The number of participants satisfying the criteria for inclusion who did or did not undergo the index tests and/or the reference standard; describe why participants failed to receive either test (a flow diagram is strongly recommended).	
RESULTS: Test	
17 Time interval from the index tests to the reference standard, and any treatment administered between.	4. Is the period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?
18 Distribution of severity of disease (define criteria) in those with the target condition; other diagnoses in participants without the target condition.	
19 A cross-tabulation of the results of the index tests (including indeterminate and missing results) by the results of the reference standard; for continuous results, the distribution of the test results by the results of the reference standard.	14. Were withdrawals from the study explained?
20 Any adverse events from performing the index tests or the reference standard.	
ESTIMATES	
21 Estimates of diagnostic accuracy and measures of statistical uncertainty (e.g., 95% confidence intervals).	
22 How indeterminate results, missing responses and outliers of the index tests were handled.	13. Were uninterpretable/intermediate test results reported?
23 Estimates of variability of diagnostic accuracy between subgroups of participants, readers or centres, if done.	
24 Estimates of test reproducibility, typically imprecision (as CV) at 2 or 3 concentrations.	
DISCUSSION	
25 Discuss the clinical applicability of the	12. Were the same clinical data available

study findings.	when test results were interpreted as would be available when the test is used in practice?

Appendix 3(b): Justification for selection of checklist items for the evaluation of spatial visualisation testing literature

Item N°	QUADAS Checklist Item	Justification for Inclusion/ Exclusion
1	Was the spectrum of patients' representative of those who will receive the test in practice?	Include but wording modified to "Spectrum of participants who will receive the tests in this study"
2	Were the selection criteria clearly described?	Cochrane Handbook indicates that this item must be included
3	Is the reference standard test likely to classify the target condition correctly?	Include but wording modified to "Will the stated test(s) measure SVS?"
4	Is the time period between reference standard & index test short enough to be reasonably sure that the target condition did not change between the 2 tests?	Include but wording modified to "Is the test-retest time period short enough to be reasonably sure that any change between the 2 tests is due solely to the stated intervention?" (NB. Only included for those studies incorporating retesting)
5	Did the whole sample or a random selection of the sample receive verification using the reference standard?	Include but wording modified to "Did all participants receive the same reference standard SVS test(s)?"
6	Did patients receive the same reference standard regardless of the index test result?	Exclude. Cochrane Handbook indicates that this is only applicable if the index test is given before the reference standard and in the experimental setting of before and after testing the index test will always follow the reference test.
7	Was the reference standard test independent of the index test? (The index test was not part of the reference test)	Exclude. The index test in the before and after studies will always be independent of the reference test.
8	Was the execution of the index test described in sufficient detail to permit replication?	Include but combined with QUADAS item 9 with wording modified to "Was the execution of the reference standard

9	Was the execution of the reference standard test described in sufficient detail to permit replication?	and any subsequent retest described in sufficient detail to permit replication?”
10	Were the index test results interpreted without knowledge of the results of the reference standard?	Include for test-retest studies only
11	Were the reference standard results interpreted without knowledge of the results of the index test?	Include but wording modified to “Have the results been interpreted in a consistent manner?”
12	Were the same clinical data available when the test results were interpreted as would be available when the test is used in practice?	Exclude. Clinical data will not be collected. The spatial visualisation test score data will be experimental data.
13	Were uninterpretable results reported?	Include but wording modified to “Were the results presented in an understandable format”?
14	Were withdrawals from the study explained?	Include.

Appendix 3 (c) Modified QUADAS / STARD checklist for SVS performance studies

Study Identification (Author, year of publication, title,)				
Checklist completed:	Completed by:			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A

Appendix 3 (d) Completed QUADAS / STARD checklists for SVS performance studies

Study Identification (Author, year of publication, title,) Appleyard, R., & Coleman, L. (2010). Virtual environment for radiotherapy training (VERT) Final project report				
Checklist completed: 4.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	**Yes**	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	**No**	Unclear	N/A
3. Were the objectives of the study pre specified?	**Yes**	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	**Yes**	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	**No**	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	**Yes**	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	**N/A**
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	**No**	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	**Unclear**	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	**Unclear**	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	**Unclear**	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	**No**	Unclear	N/A
	4	4	3	1

Study Identification (Author, year of publication, title) Clem, D., Anderson, S., Donaldson, J., Hdeib, M. (2010). An exploratory study of spatial ability and student achievement in sonography					
Checklist completed: 4.2.2011	Completed by: AJW				
	Circle **ONE** option for each question				
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	9	0	2	1

Study Identification (Author, year of publication, title) Cohen, C.A., Hegarty, M. (2007). Sources of difficulty in imagining cross sections of 3D objects.				
Checklist completed: 4.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	9	2	0	1

Study Identification (Author, year of publication, title) Geiser, C., Lehmann, W., Eid, M. (2006). Separating “rotators” from “non-rotators” in the mental rotations test: a multigroup latent class analysis				
Checklist completed: 4.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant’s representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	9	1	1	1

Study Identification (Author, year of publication, title) Green, D., Appleyard, R. (2011). The influence of VERT™ characteristics on the development of skills in skin apposition techniques				
Checklist completed: 4.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	8	3	0	1

Study Identification (Author, year of publication, title) Hedman, L., Klingberg, T., Enochsson, L., Kjellin, A., Felländer-Tsai, L. (2007). Visual working memory influences the performance in virtual image-guided surgical intervention				
Checklist completed: 4.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	7	3	1	1

Study Identification (Author, year of publication, title) Hegarty, M., Waller, D. (2004). A dissociation between mental rotation and perspective taking spatial abilities				
Checklist completed: 4.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	7	2	2	1

Study Identification (Author, year of publication, title) Kaufman, S.B. (2007). Sex differences in mental rotation and spatial visualization ability: can they be accounted for by differences in working memory capacity?				
Checklist completed: 4.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	7	3	1	1

Study Identification (Author, year of publication, title) Keehner, M.M., Tendick, F., Meng, M.V., Anwar, H.P., Hegarty, M., Stoller, M.L., Duh, Q-Y. (2004). Spatial ability, experience and skill in laparoscopic surgery				
Checklist completed: 4.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	6	4	1	1

Study Identification (Author, year of publication, title) Keehner, M.M., Lippa, Y., Montello, D.R., Tendick, F., Hegarty, M. (2006). Learning a spatial skill for surgery: how the contributions of abilities change with practice				
Checklist completed: 4.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	9	2	0	1

Study Identification (Author, year of publication, title) Luursema, J-M., Buzink, S., Verwey, W.B., Jakimowicz, J-J. (2010). Visuo-spatial ability in colonoscopy simulator training				
Checklist completed: 4.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	5	5	1	1

Study Identification (Author, year of publication, title,) Parsons, T., Larson, P., Kratz, K., Thiebaux, M., Bluestein, B., Buckwalter, J.G., Rizzo, A.A. (2004). Sex differences in mental and spatial rotation in a virtual environment				
Checklist completed: 5.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	5	5	1	1

Study Identification (Author, year of publication, title) Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test: different versions and factors that affect performance				
Checklist completed: 5.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	6	4	1	1

Study Identification (Author, year of publication, title) Smoker, W.R.K., Berbaum, K.S., Luebke, N.H., Jacoby, C.G. (1984). Spatial perception testing in diagnostic radiology				
Checklist completed: 5.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	6	5	0	1

Study Identification (Author, year of publication, title) Waywell, L., Bogg, J. (1999). Spatial ability assessment: an aid to student selection for therapy radiography training.					
Checklist completed: 5.2.2011	Completed by: AJW				
	Circle **ONE** option for each question				
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?		Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?		Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?		Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?		Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?		Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?		Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?		Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?		Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?		Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?		Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?		Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?		Yes	No	Unclear	N/A
		8	3	0	1

Study Identification (Author, year of publication, title) Zacks, J.M., Mires, J., Tversky, B., Hazeltine, E. (2000). Mental spatial transformations of objects and perspective.					
Checklist completed: 5.2.2011	Completed by: AJW				
	Circle **ONE** option for each question				
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A	
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A	
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A	
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A	
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A	
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A	
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A	
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A	
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A	
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A	
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A	
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A	
	5	6	0	1	

Study Identification (Author, title, year of publication) Alias, M., Black, T.R., Gray, D.E. (2002). Effect of instruction on spatial visualisation ability in civil engineering students				
Checklist completed: 5.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	7	3	2	0

Study Identification (Author, title, year of publication) Gorska, R., Sorby, S.A., Leopold, C. (1998). Gender differences in visualization skills – an international perspective.				
Checklist completed: 5.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	6	6	0	0

Study Identification (Author, title, year of publication) Hegarty, M., Keehner, M., Khooshabeh, P. Montello, D.R. (2009). How spatial abilities enhance and are enhanced by dental education				
Checklist completed: 5.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	7	4	1	0

Study Identification (Author, title, year of publication) Hoyek, N., Collet, C., Rastello, O., Fargier, P., Thiriet, P., Guillot, A. (2009). Enhancement of mental rotation abilities and its effect on anatomy learning				
Checklist completed: 6.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	9	3	0	0

Study Identification (Author, title, year of publication) Jansen, P., Pietsch, S. (2010). Physical activity improves mental rotation performance				
Checklist completed: 6.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	8	3	1	0

Study Identification (Author, title, year of publication) Németh, B. (2007). Measurement of the development of spatial ability by Mental Cutting Test				
Checklist completed: 6.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	6	4	2	0

Study Identification (Author, title, year of publication) Rafi, A., Anuar, K., Samad, A., Hayati, M., Mahadzir, M. (2005). Improving spatial ability using a web based virtual environment (WbVE)				
Checklist completed: 6.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	8	2	2	0

Study Identification (Author, title, year of publication) Russell, C., Churches, A. (2010). What do we really want to know about spatial visualization skills among engineering students?				
Checklist completed: 6.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	8	3	1	0

Study Identification (Author, title, year of publication) Terlecki, M.S., Newcombe, N.S., Little, M. (2008). Durable and generalized effects of spatial experience on mental rotation: gender differences in growth patterns				
Checklist completed: 6.2.2011	Completed by: AJW			
	Circle **ONE** option for each question			
1. Participant spectrum: Was the spectrum of participant's representative of the participants who will receive the test in this study?	Yes	No	Unclear	N/A
2. Were the selection criteria described clearly?	Yes	No	Unclear	N/A
3. Were the objectives of the study pre specified?	Yes	No	Unclear	N/A
4. Reference test measurement: Will the stated reference test(s) measure spatial visualisation skill?	Yes	No	Unclear	N/A
5. Replication: Was the execution of the reference test(s) and retest(s) described in sufficient detail to permit its replication?	Yes	No	Unclear	N/A
6. Differential verification: Did all participants receive the same reference SVS test(s)?	Yes	No	Unclear	N/A
7. Test – retest time period: Is the test – retest time short enough to be reasonably sure that any change between the two tests is due solely to the stated intervention?	Yes	No	Unclear	N/A
8. Test review (Blinding): Were the index test results interpreted without knowledge of the results of the reference test(s)?	Yes	No	Unclear	N/A
9. Interpretation & review: Have the results been interpreted in a consistent manner?	Yes	No	Unclear	N/A
10. Interpretable results: Were the results presented in an acceptable format (not ambiguous)?	Yes	No	Unclear	N/A
11. Withdrawals: Clear report of what happened to all participants throughout the duration of the study?	Yes	No	Unclear	N/A
12. Validity / Reliability of tests: Was there any statistical analysis of the validity and reliability of the SVS test components?	Yes	No	Unclear	N/A
	10	2	0	0

Appendix 4

(a) Ethics confirmation UPR 16

(b) Pilot phase ethics information

FORM UPR16

Research Ethics Review Checklist

Please include this completed form as an appendix to your thesis (see the Research Degrees Operational Handbook for more information

Postgraduate Research Student (PGRS) Information		**Student ID:**	373155
PGRS Name:	Andrew James Williams		
Department:	SHSSW	**First Supervisor:**	Dr Jason Oakley
Start Date: (or progression date for Prof Doc students)		**February 2010**	

Study Mode and Route:	Part-time ☒	MPhil ☐	MD ☐
	Full-time ☐	PhD ☒	Professional Doctorate ☐

Title of Thesis:	The role of an immersive 3-D virtual reality environment in the development of the spatial visualisation skill of pre-registration therapeutic radiography students
Thesis Word Count: (excluding ancillary data)	83770

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University's Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study

Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

UKRIO Finished Research Checklist:
(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: http://www.ukrio.org/what-we-do/code-of-practice-for-research/)

a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame?	YES NO	☒ ☐
b) Have all contributions to knowledge been acknowledged?	YES NO	☒ ☐
c) Have you complied with all agreements relating to intellectual property, publication and authorship?	YES NO	☒ ☐
d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration?	YES NO	☒ ☐
e) Does your research comply with all legal, ethical, and contractual requirements?	YES NO	☒ ☐

Candidate Statement:

I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s)

Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC):	SHSSW 10-05 SFEC 2013-26

If you have *not* submitted your work for ethical review, and/or you have answered 'No' to one or more of questions a) to e), please explain below why this is so:

Signed *(PGRS)*:		**Date:** 26.5.2020

UPR16 – April 2018

Pilot Study

Dr Jeannette Bartholomew BSc PhD PGCE ILTM
Head of School

School of Health Sciences and Social Work
University of Portsmouth

School of Health Sciences and Social Work
University of Portsmouth
James Watson Hall (West)
2 King Richard 1st Road
Portsmouth PO1 2FR
United Kingdom

T: +44 (0)23 9284 4440
F: +44 (0)23 9285 4402

Andy Williams
School of Health Sciences and Social Work
University of Portsmouth

14 October 2010

Dear Andy

Application to SHSSW Research Ethics Committee: An evaluation of the effectiveness of simulated and virtual environments in the leaning and assessment of clinical skills for pre registration and radiography students (Ref: 10-05)

Thank you for submitting the above application to the SHSSW Research Ethics Committee.

Members have read your application and the Committee would like to give a favourable opinion with provision. Provision relates to the following points:

- Given the potential conflict of role as you as the lead researcher and a member of the course team, it is important that students do not feel coerced in any way to participate. However, your proposed method of recruitment via invitation packs being made available at the end of computer workshops should minimise the risk of this. It would be important to keep a face-to-face approach to recruitment between you and the student to a minimum.

- It should be made clear to potential participants that it is not yet known whether VERT will improve their skills in case they assume that it will and they take part on that basis.

- On the information sheet under 'What will happen to me...?' it states "Following this and at intervals during the course (usually annually) you will be asked to complete further on line questionnaires". Can you be more precise here? Students should know exactly how often they will be asked to complete these and also when the study will end so that they know exactly what they are signing up to at the outset.

- Will you have a system in place to support students who have difficulties using the VERT system?

- In the 'Complaints' section of the information sheet, the Head of School's contact details should be included as well as the Professional Lead for Radiography, as a more independent party.
- If you do use focus groups, it will be important to consider the precise recruitment procedure, focus group procedure, topic guide, location, time and facilitator. It would be advisable to provide more information on these issues in the future, nearer the time of the focus groups.

Additional more specific points include:

- Harold's telephone number is incorrect
- On the VERT questionnaire and some of the answer categories need thinking through more carefully e.g. 2-5 years and 5-10 years - if you want to answer 5 years which do you tick?

The Committee would appreciate a response regarding these points before a favourable opinion can be given.

Yours sincerely

Dr Rebecca Stores
Chair of SHSSW Research Ethics and Peer Review Committee

Information Sheet

Version 1.2: October 2010

An evaluation of the effectiveness of simulated and virtual environments in the learning and assessment of clinical skills for pre-registration radiography students

You are being invited to take part in a research study examining the use of the X-ray practice suite and the Virtual Environment for Radiotherapy Training (VERT™) platform. Before you decide whether to take part or not it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Talk to others about the study if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

The use of simulated and virtual environments in Radiography education is relatively new and we do not yet know how they will impact on skill development and improvement. The aim of this study will be an evaluation of how the simulated environments contribute to the development of clinical skills and how they may be used for the assessment of these skills.

Specifically we need to:

- determine the spatial ability of diagnostic and radiotherapy students
- determine whether these change over the duration of the course and to what extent the simulated environments contribute to that change
- develop clinical assessment packages which may support our current clinical assessment portfolio
- test these assessments for validity and reliability in comparison with current clinical assessments
- make recommendations for the integration of the simulated environments into future radiography course structures

Why have I been chosen?

During your radiography course you will participate in pre-placement tutorials and practical sessions in the diagnostic X-ray suite and / or VERT™ which will prepare you for your time in the clinical departments. These facilities are relatively new and we need to formally evaluate how they are being used. The study is open to all diagnostic and radiotherapy students joining the course in September 2010.

Do I have to take part?

Taking part in the research is entirely voluntary so it is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to withdraw at any time and without giving a reason. Your decision on whether to take part or not will have absolutely no impact on the running of the course or your participation in it.

What will happen to me if I take part and what do I have to do?

If you agree to participate, at the beginning of the study you will be asked to complete two spatial ability tests. You will be asked to complete further on line questionnaires, "paper and pencil" tests and surveys at the end of your second year and those of you using the VERT™ platform may be asked to participate in the assessment of clinical skills during your second and third year (the results of these will have absolutely no impact on your other work based learning marks.

Participation in voluntary focus groups may be required following clinical practice placements. Further information detailing how these will be set up and run will be provided for you nearer the time if they become necessary.

What are the other possible disadvantages and risks of taking part?

No other risk or inconvenience has been identified and you should not experience any discomfort while working in the simulated environments.

What are the possible benefits of taking part?

The results for your learning style and spatial ability can be made available for your personal use. You may find them useful when you are identifying your learning and development needs for your individual learning profiles, work based learning contracts and action plans.

What if there is a problem?

If you have any cause for complaint about any aspect relating to the way you have been approached or treated during the course of the study you should contact the Head of the School of Health Sciences & Social Work in the first instance:-

Dr Jeannette Bartholomew

Mail: jeannette.bartholomew@port.ac.uk

Telephone: 02392 844400

Alternatively you may contact the professional lead for Radiography:-

Mr Harold Clarke

Mail: Harold.clarke@port.ac.uk

Telephone: 0293 845391

Will my taking part in the study be kept confidential?

All information gathered during the study will remain confidential and will only be seen by the researchers. Any questionnaire that you complete will be identified solely with your student ID number which will facilitate correlation with previous or future results.

All paper copy questionnaires will be filed in a ring binder and stored in a locked filing cabinet in the Radiography academic team office. The results of on line electronic questionnaires and all statistical analysis using proprietary packages will be stored on an encrypted memory stick which will be stored in the same location. All data will be accessed solely by the researchers.

Any data which needs to be shown the researchers' directors of studies or project supervisors will be anonymised.

All data will be destroyed at the end of the study in line with the University of Portsmouth data protection policy and the Data Protection Act 1998

What will happen to the results of the research study?

The results of the research will provide evidence which will inform the outcomes and recommendations of the research project. Results appearing in any publication arising from the research will remain anonymous and you will not be identified individually unless you have consented to the release such information.

Paper copies of your individual performance scores and learning styles will be available to you and will be presented to you in a sealed envelope. Results from electronic surveys can be downloaded personally or mailed to you using your university email account.

Who is organising the research?

The research is being conducted by Mr Andrew Williams, Senior Lecturer in Radiography, School of Health Sciences and Social Work, University of Portsmouth.

It has been reviewed and approved by the SHSSW Research Ethics Committee

Contact Details:

For further information about the study or to discuss any concerns that you may still have please contact:-

Mr Andrew Williams

Mail: andrew.j.williams@port.ac.uk

Telephone: 01293 845994

Thank you for taking the time to read this information and for considering your participation in the study.

If you are happy to continue please complete and sign the attached consent form.

Yours Sincerely

CONSENT FORM

Title of Project: An evaluation of the effectiveness of simulated and virtual environments in the learning and assessment of clinical skills for pre-registration radiography students

Name of Researcher: Andrew Williams

Please initial box

I confirm that I have read and understood the information sheet dated 18 October 2010 (v1.2) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason ☐

I agree to audio recording of interviews ☐

I agree to take part in the above study. ☐

_________________________	________________	____________________
Name of Participant	Date	Signature
_________________________	________________	____________________
Name of Person taking consent	Date	Signature
(If different from researcher)		
_________________________	________________	____________________
Researcher	Date	Signature

Appendix 5

Experimental phase ethics information

Faculty of Science
University of Portsmouth
St Michael's Building
White Swan Road
PORTSMOUTH
PO1 2DT

Address
Date

FAVOURABLE OPINION

Protocol Title: The role of an immersive 3-D virtual reality environment (VERT™) in the development of spatial visualisation skill of pre-registration therapeutic radiography students

Date Reviewed: 19th Sept 2013

Dear Mr Williams,

Thank you for submitting your protocol for ethical review.

Your submission has been reviewed and I am pleased to inform you that your application has been given a favourable opinion by the Science Faculty Ethics Committee. The committee suggested it may be helpful to give greater emphasis to the point that it is the equipment/education that is the focus of the research, rather than the students' competence. The concern would be what would happen if it became clear that a radiography student had poor visualisation skills at the outset and they had not significantly improved. It would be helpful if you could amend your documentation to reflect this.

Please notify us in the future of any substantial amendments that may be required and send us a final study report.

Good luck with the study!

Yours sincerely,

John C

Dr John Crossland
School of Health Sciences and Social Work
Science Faculty Ethics Committee

CC -
Dr Chris Markham – Chair of SFEC
Dr Jim House – Vice Chair of SFEC
Jody Salt – Faculty Administrator

School of Health Sciences and Social Work

Radiography

Spatial Visualisation Skills Study

Information Sheet Version 2.3 (June 2013)

You are being invited to take part in a research study which will evaluate the use of the Virtual Environment for Radiotherapy Training (VERT™) platform and how it may be used to support the development of spatial visualisation skill. Before you decide whether to take part or not, it is important for you to understand why the research is being done and what it will involve.

Please take the time to read the following information carefully and talk to others about the study if you wish. If there is anything that you do not understand or if you need further information then please ask the research coordinator: Andy Williams. Take time to decide whether or not you wish to take part.

What is the purpose of the study?

Being able to visualise objects in 3 dimensions is an important skill for radiographers and the use of simulated and virtual environments in Radiography education is relatively new. The aim of this research is to evaluate how these environments can contribute to the development of these 3-D spatial visualisation skills.

Specifically we need to:

- determine the baseline spatial visualisation skill of radiography students
- determine whether this is different when compared with students enrolled on other undergraduate courses
- determine whether spatial visualisation skill can change over the duration of the course and to what extent the simulated and virtual environments have contributed to that change

Why have I been chosen?

During your radiography course you will participate in timetabled pre & post placement tutorials and practical sessions in the diagnostic X-ray suite and / or VERT™. These will prepare and support you for your time in the clinical departments. These facilities are relatively new and we need to formally evaluate how they contribute to visualisation skill development. The study is open to all first year radiography students joining a programme from September 2013.

Do I have to take part?

Taking part in the research is entirely voluntary, so it is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You will be free to withdraw at any time and without needing to give a reason. Your decision about whether to take part or not will have absolutely no impact on your participation in timetabled skills sessions, the running of the course or your participation in it.

What will happen to me if I take part and what do I have to do?

If you agree to participate, at the beginning of the study you will be asked to complete 2 online spatial ability tests and a spatial visualisation activity questionnaire. These should take no longer than 40 minutes in total. You will be asked to complete further tests when you have completed your clinical practice placements for this academic year and the cycle will be repeated in the same way during the next academic year.

What are the other possible disadvantages and risks of taking part?

No risk or inconvenience has been identified and you should not experience any discomfort while working in the simulated environments.

What are the possible benefits of taking part?

Your results can be made available to you for your personal use at the end of the study. You may find them useful when you are identifying your learning and development needs for your individual learning profile.

Will my taking part in the study be kept confidential?

All information gathered during the study will remain confidential. Any questionnaires that you complete will be identified solely with your student ID number which will facilitate correlation with previous or future results.

Any paper copy questionnaires will be filed in a ring binder and stored in a locked filing cabinet in the Radiography academic team office. The results of online electronic questionnaires and all statistical analysis using proprietary packages will be stored on an encrypted memory stick which will be stored in the same location. Any data which needs to be shown to the researchers' directors of studies or project supervisors will be anonymised.

All data will be destroyed at the end of the study in line with the University of Portsmouth data protection policy and the Data Protection Act 1998.

What if there is a problem?

If you have any cause for complaint about any aspect relating to the way you have been approached or treated during the course of the study you should contact the Head of the School of Health Sciences & Social Work in the first instance:-

Dr Jeannette Bartholomew

Mail: jeannette.bartholomew@port.ac.uk

Telephone: 02392 844400

Alternatively you may contact the Professional Lead for Radiography:-

Mr Harold Clarke

Mail: Harold.clarke@port.ac.uk

Telephone: 0293 845391

What will happen to the results of the research study?

The results of the research will provide evidence which will inform the outcomes and recommendations of the research project. Results appearing in any publication arising from the research will remain anonymous and you will not be identified individually unless you have consented to the release such information. Results of your individual performance scores will be made available to you at the end of the study should you so wish. Paper copies will be presented to you in a sealed envelope, while access to results from electronic tests and questionnaires can be provided so that they may be downloaded. Alternatively they may be mailed to you using your university email account.

Who is organising the research?

The research is being conducted by Mr Andrew Williams, Senior Lecturer in Radiography, School of Health Sciences and Social Work, University of Portsmouth as part of his Doctoral research. It has been reviewed and approved by the SHSSW Research Ethics Committee

Contact Details:

For further information about the study or to discuss any concerns that you may still have please contact:-

Mr Andrew Williams Mail: andrew.j.williams@port.ac.uk Telephone: 01293 845994

Thank you for taking the time to read this information and for considering your participation in the study. If you are happy to continue please complete and sign the attached consent form.

Radiography

CONSENT FORM

Title of Study: The role of an immersive 3-D virtual reality environment (VERT™) in the development of spatial visualisation skill of pre-registration therapeutic radiography students

Name of Researcher: Andrew Williams

REC Ref No: SFEC 2013-26

Please tick boxes

I confirm that I have read and understood the information sheet for the above study (version 2.3, June 2013). I have had the opportunity to consider the information, to ask questions and have these answered satisfactorily. ☐

I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without penalty ☐

I agree to take part in the above study. ☐

Name of Participant	Date	Signature

Name of Person receiving consent	Date	Signature
(If different from researcher)		
Researcher	Date	Signature

Appendix 6

(a) Administrator instructions

(b) Mental Rotation and Santa Barbara Solids Test Instruments

(c) Answer grid for study 1 online test

Appendix 6 (a) Administrator instructions for pilot study 1

Introduction to the session

Open with a thank you to all attendees for giving up their time

Outline the purpose of this part of the study – to determine whether there has been any change in student's ability to mentally visualise and transform an assortment of 3-D shapes. Information will be used to determine how the simulated and virtual environments (X-ray suite & VERT™) in radiography may be used to help students develop their spatial visualisation skills - an important part of a radiographers skill set

Tell participants not to open the booklet until instructed to do so and that all test instruments are included in the study booklet which is then handed out

Ask participants to record ID number on the cover page. Remind them that all answers will be anonymised but the ID number will allow collation with the previous results and to give individual feedback on results if required (they may wish to use in the future for their personal development plan)

Please ask participants to read the introduction on page 2

Point out that there are 2 sections to be completed

Point out that while this is not an exam, the tests should be done individually and there should be no talking while the instructions are being given or while tests are being completed

Tell participants that if they wish to leave at any point they are free to do so but to be aware that if they are going to leave the room they should do so quietly so as not to disturb other participants and that there will be no implications or repercussions regarding their place on the course or the opportunity to take part in further studies.

The investigators may however wish to follow up on their reasons for leaving.

Instructions for Test 1 (Redrawn Mental Rotation Test version A, Peters)

Ask the students to turn to page 3 of their booklet headed "Test Instrument 1"

Ask them to look at the 5 objects and satisfy themselves that they are the same shape but rotated around the vertical axis. You can demonstrate this by rotating your extended hand.

Then ask them to look at the next 2 objects and point out that they are identical but different to the first 5. They should satisfy themselves that this is the case.

Follow this by asking them to look at the next set of 5 objects – you should point out to them that the image on the left hand side is known as a target figure. The other 4 are known as stimulus figures. The stimulus figures are rotated versions of the target. Two of them are correct. Participants should identify both correct figures by putting an X through each of them.

When they have done this, tell them that the correct answer is the **first** and **third** object

Ask them to move on to the next three examples on page 4 –

Correct choices 2: second & third, 3: first & fourth, 4: first & third

Check that there are no questions at this point before moving on to the test

Tell participants that there are ***12 test items on 2 pages***

Read the following instruction:

"We are now ready to move on to the test itself, there are 12 test (the target) figures and 4 associated criterion figures for each of the target figures. Remember that there are 2 and only 2 correct alternatives for each test (target) figure. You should mark the correct criterion figures with a large X. You will score 1 point for each correct pair you identify.

You have 4 minutes to complete this section and you may start now"

When the timer indicates 4 minutes participants should be instructed to stop writing regardless of whether they have finished all test objects

Instructions for Test 2 (Santa Barbara Solids Test)

When all students have completed test 1 you should move on to the second test.

Ask participants to turn to test instrument 2 in their booklet and you read the instructions on page 9 & 10 while they follow them.

Then ask them to look at the sample problem on the following page (p. 11) and satisfy themselves that "C" is the correct answer

Ask them if there are any questions relating to test 2 at this point

Read the following test instructions to them and then start the timer

"Circle the cross section you would see when the grey cutting plane slices the object. Imagine that you are facing the cutting planes head on, as if you were looking in a mirror.

Make your choice based on the shapes of the possible answers, not their sizes.

You have 5 minutes in which to complete the test. You may begin"

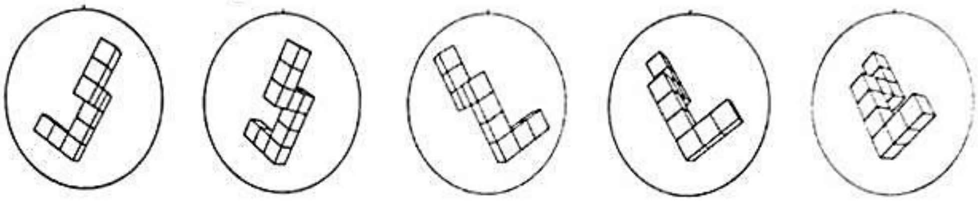
When the timer indicates 5 minutes participants should be instructed to stop writing regardless of whether they have finished all test objects.

TEST ENDS

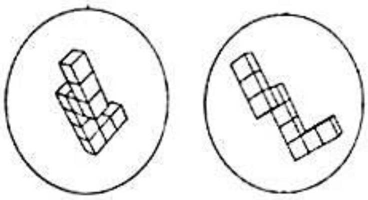
Appendix 6 (b) Vandenberg & Kuse 20 item Mental Rotation Test and scoring key

Mental Rotation Test

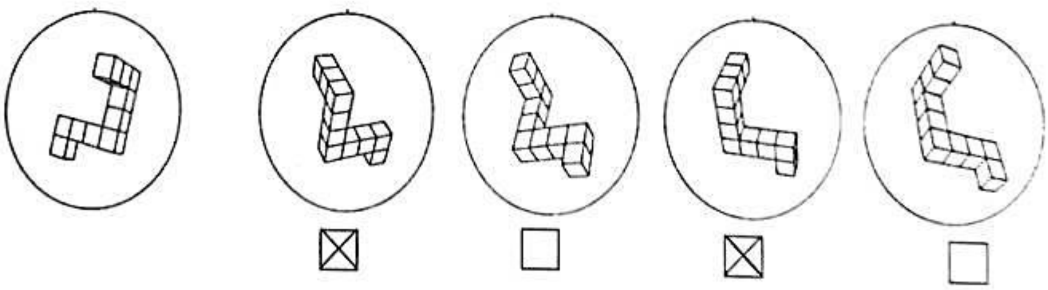
This is a test of your ability to look at a drawing of a given object and find the same object within a set of dissimilar objects. The only difference between the original object and the chosen object will be that they are presented at different angles. An illustration of this principle is given below where the same single object is given in five different positions. Look at each of them to satisfy yourself that they are only presented at different angles from one another.



Below are two drawings of new objects. They cannot be made to match the above five drawings. Please note that you may not turn over the objects. Satisfy yourself that they are different from the above.



Now let's do some sample problems. For each problem there is a primary object on the far left. You are to determine which two of four objects to the right are the same object given on the far left. In each problem always <u>two</u> of the four drawings are the same object as the one on the left. You are to put Xs in the boxes below the correct ones, and leave the incorrect ones blank. The first sample problem is done for you.

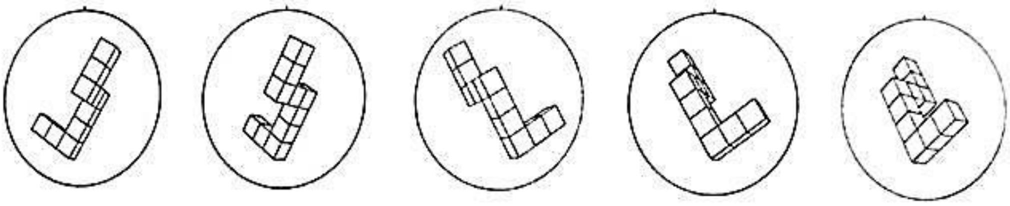


Proceed to the next page

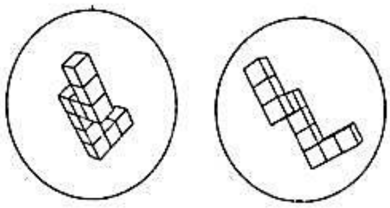
1

Mental Rotation Test

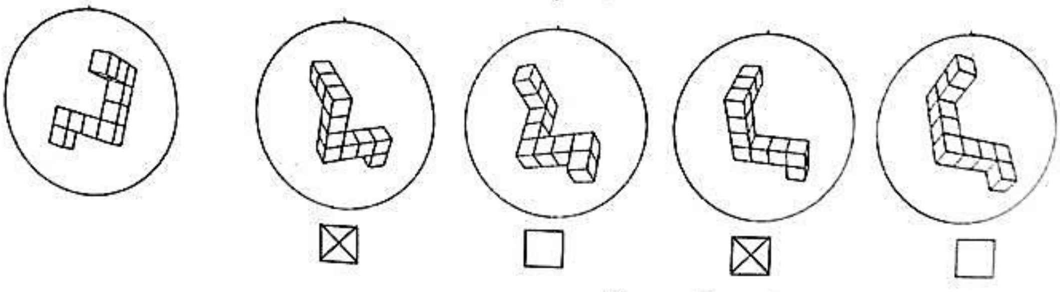
This is a test of your ability to look at a drawing of a given object and find the same object within a set of dissimilar objects. The only difference between the original object and the chosen object will be that they are presented at different angles. An illustration of this principle is given below where the same single object is given in five different positions. Look at each of them to satisfy yourself that they are only presented at different angles from one another.



Below are two drawings of new objects. They cannot be made to match the above five drawings. Please note that you may not turn over the objects. Satisfy yourself that they are different from the above.

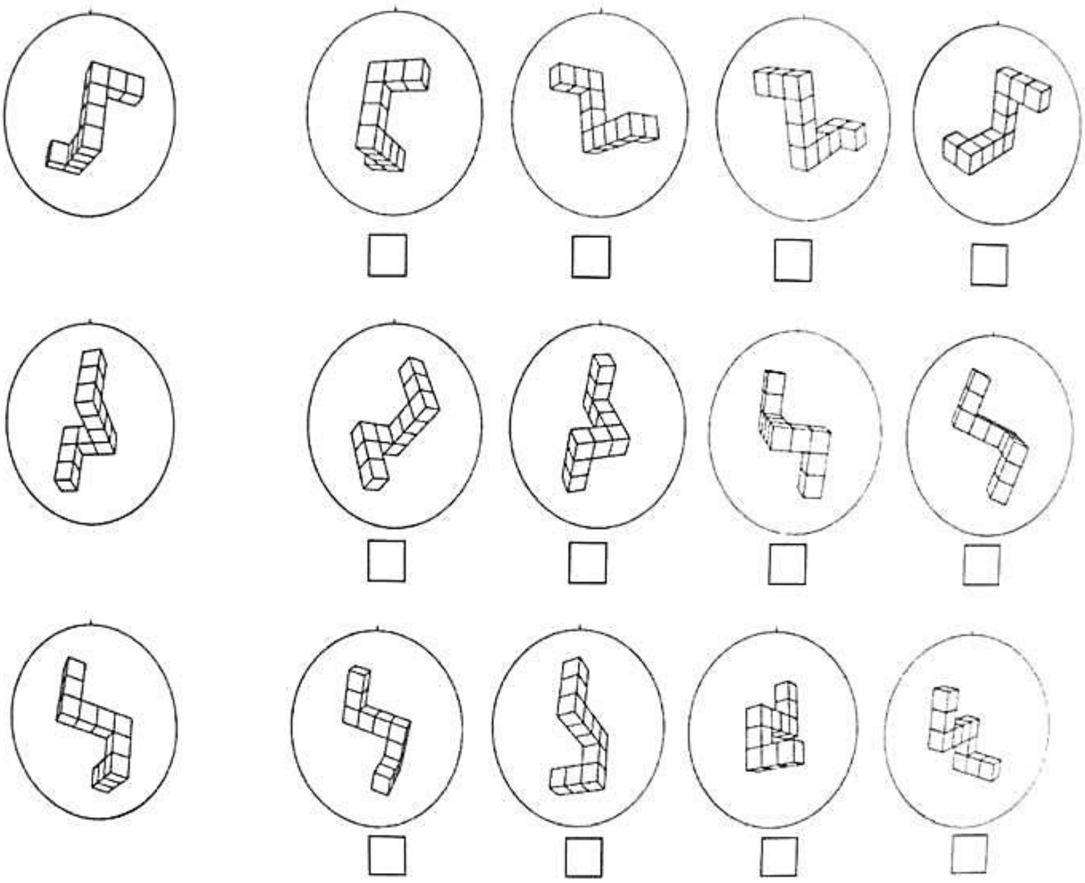


Now let's do some sample problems. For each problem there is a primary object on the far left. You are to determine which two of four objects to the right are the same object given on the far left. In each problem always <u>two</u> of the four drawings are the same object as the one on the left. You are to put Xs in the boxes below the correct ones, and leave the incorrect ones blank. The first sample problem is done for you.



Proceed to the next page

Do the rest of the sample problems yourself. Which two drawings of the four on the right show the same object as the one on the left? There are always two and only two correct answers for each problem. Put an X under the two correct drawings.

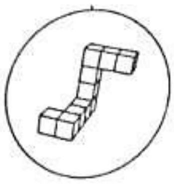


Answers: (1) first and second drawings are correct
(2) first and third drawings are correct
(3) second and third drawings are correct

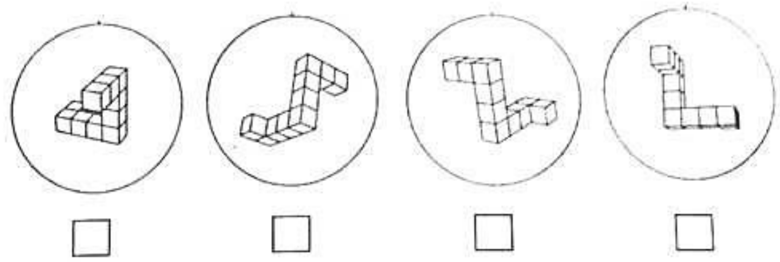
This test has two parts. You will have 3 minutes for each of the two parts. Each part has two pages. When you have finished Part I, STOP. Please do not go on to Part II until you are asked to do so. Remember: There are always two and only two correct answers for each item.

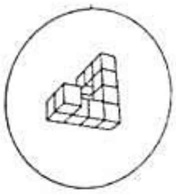
Work as quickly as you can without sacrificing accuracy. Your score on this test will reflect both the correct and incorrect responses. Therefore, it will not be to your advantage to guess unless you have some idea which choice is correct.

DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO

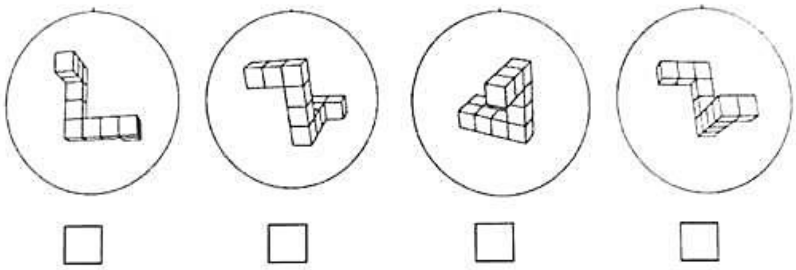


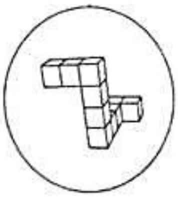
1.



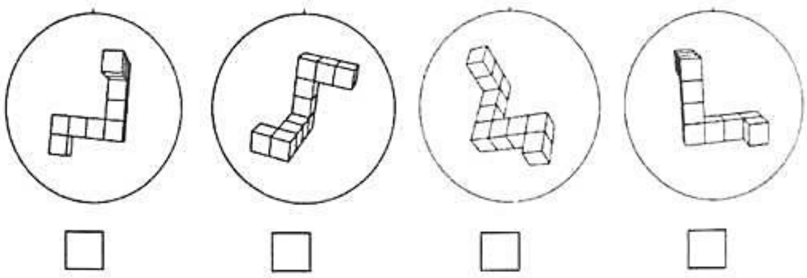


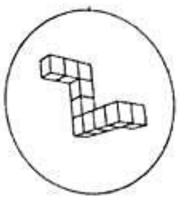
2.



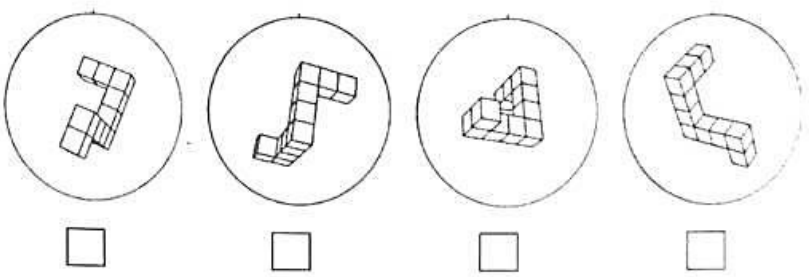


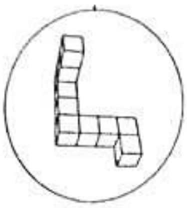
3.



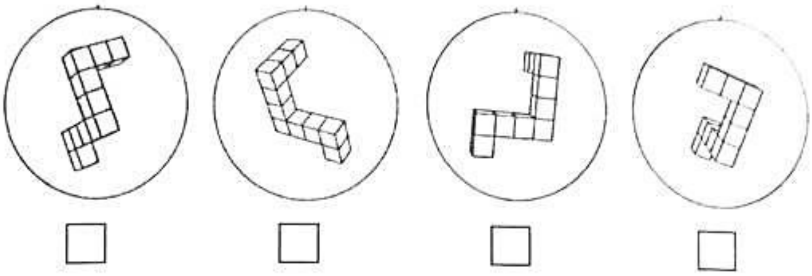


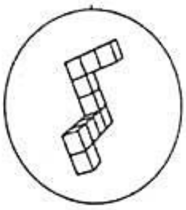
4.



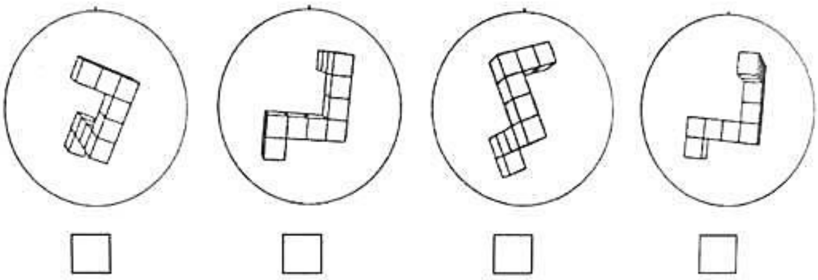


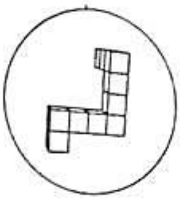
5.



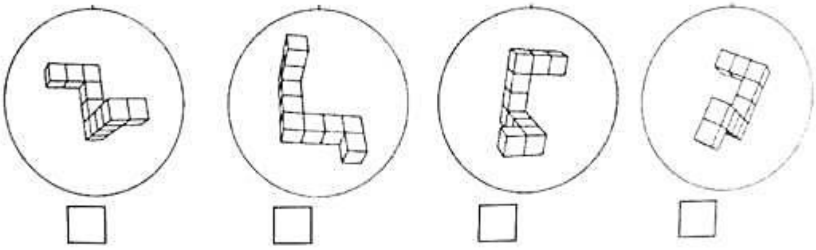


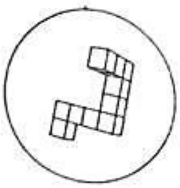
6.



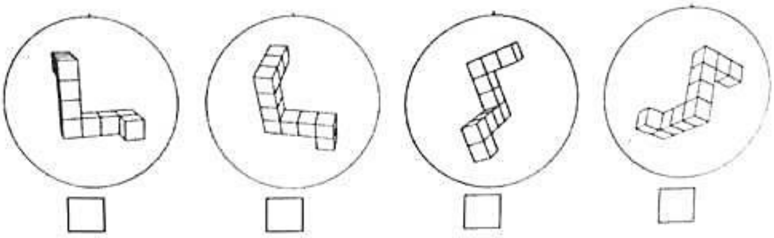


7.

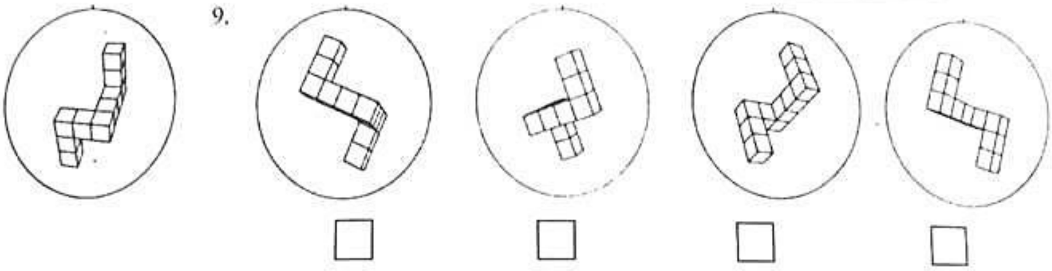




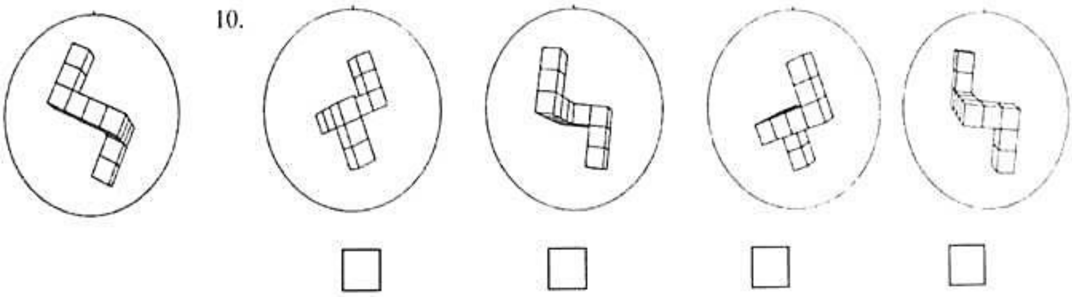
8.



9.

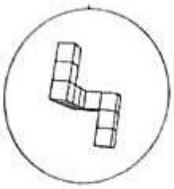


10.

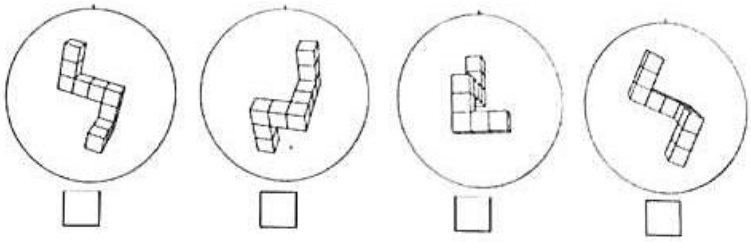


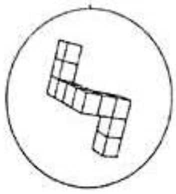
DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO.

STOP

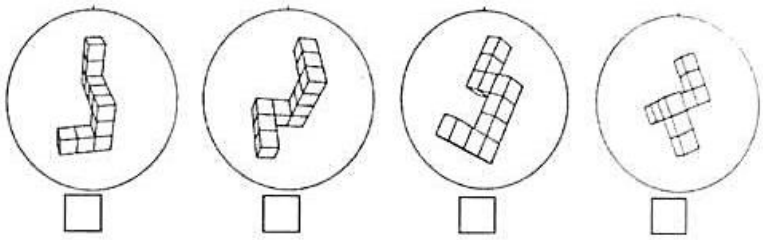


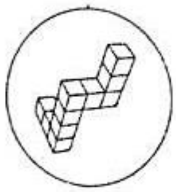
11.



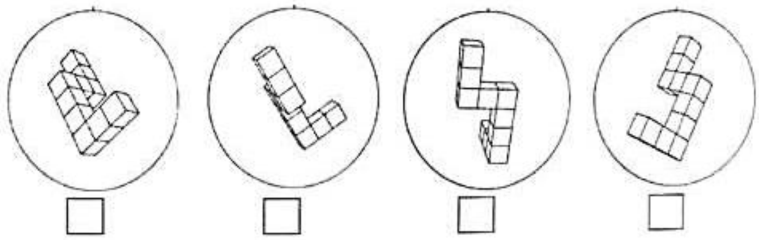


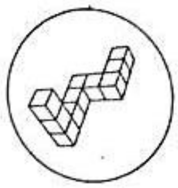
12.



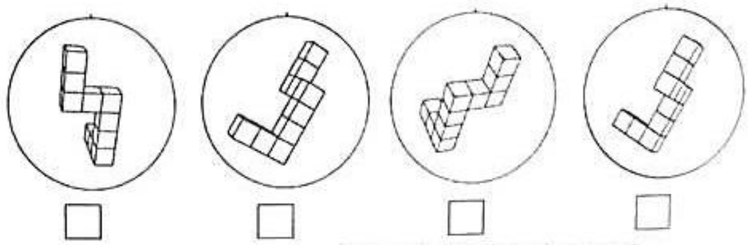


13.



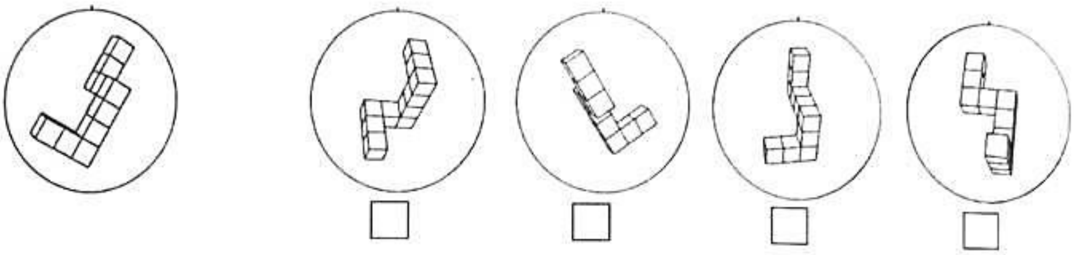


14.

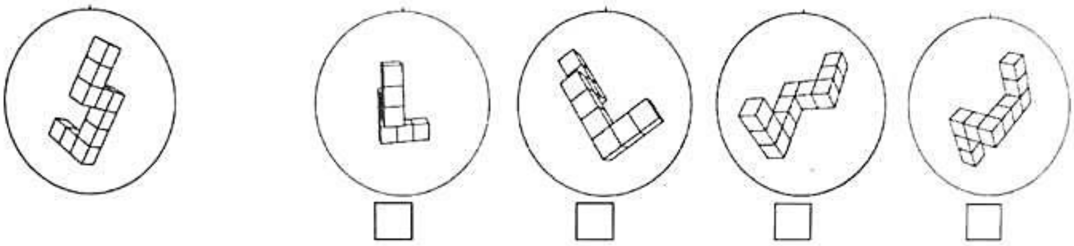


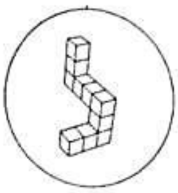
Proceed to the next page

15.

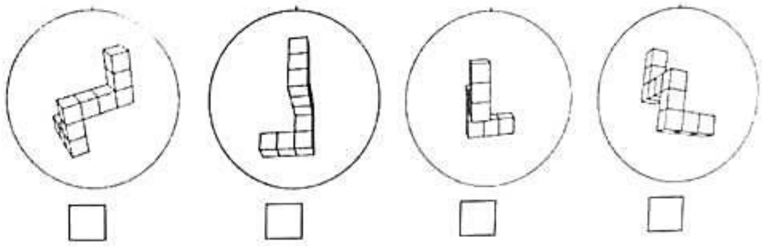


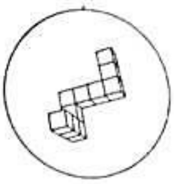
16.



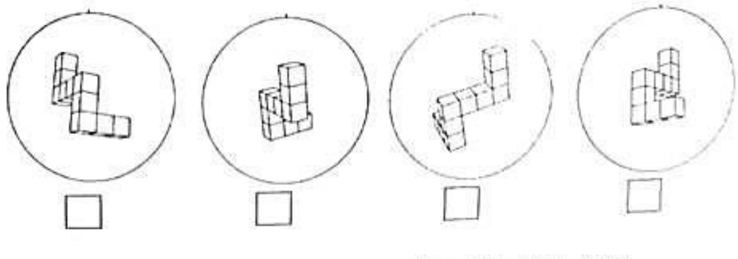


17.

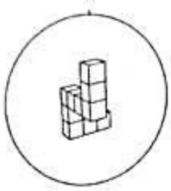




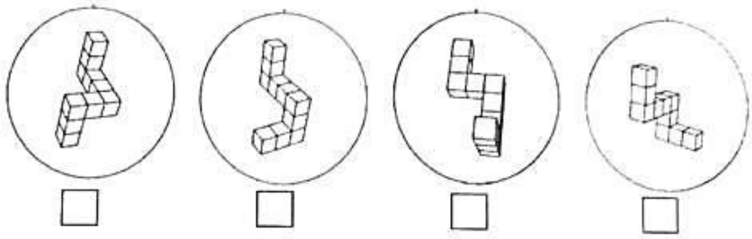
18.

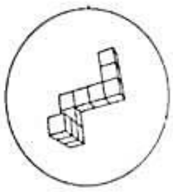


Proceed to the next page

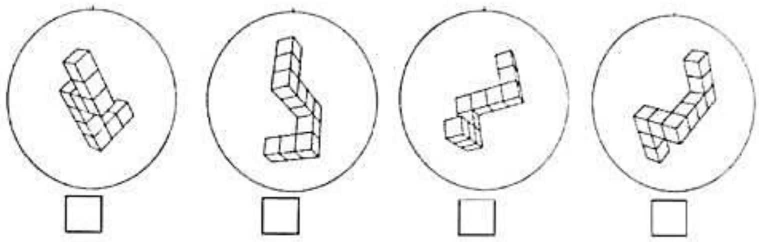


19.





20.



END

Vandenberg & Kuse 20 Item Mental Rotation Test Scoring Key

Both object selections must be correct to gain1 point for each question

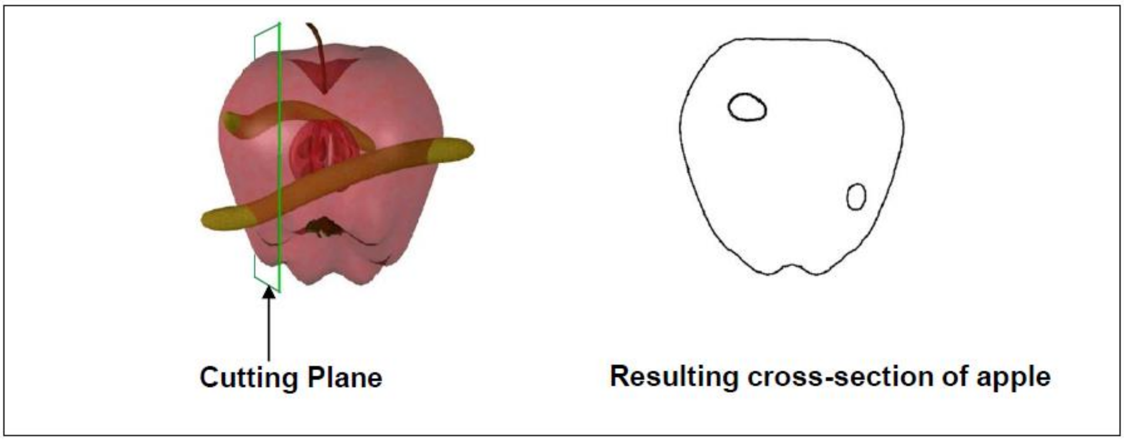
Question	Correct Objects
1	1, 3
2	1, 4
3	2, 4
4	2, 3
5	1, 3
6	1, 4
7	2, 4
8	2, 3
9	2, 4
10	1, 4
11	2, 4
12	2, 4
13	2, 4
14	1, 4
15	2, 4
16	2, 3
17	1, 3
18	1, 4
19	2, 4
20	2, 3

Cross Section Test

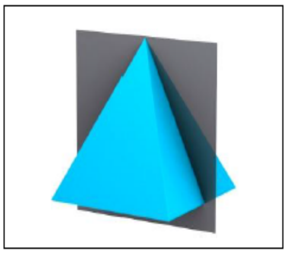
This is a test about **cross-sections**. A cross-section is the 2D shape that results when a cutting plane intersects an object.

You've seen many examples of cross-sections in everyday life. For example, when you slice an apple from top to bottom, the resulting cut surface is a cross-section of the apple.

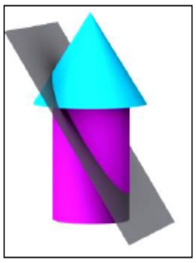
The picture below shows an apple with some worms inside. Note that the cross section on the right shows both the apple and the shapes and locations of the sliced worms inside the apple.



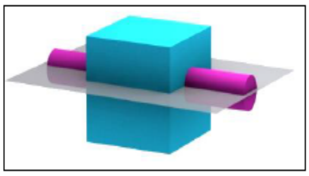
In this multiple choice test, you will be asked to identify the cross sections of three types of figures:



Single object



Attached objects

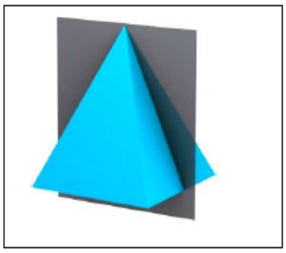


Nested objects
(one object is inside another)

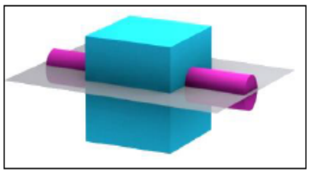
Here are some important things to remember:

- All figures are solid (not hollow) objects.
- The objects are about 6-8 inches tall. Imagine that they are on the table in front of you.
- Attached figures are “glued together” at their edges.
- Nested objects consist of one object inside another. In the nested object above, the cylinder extends all the way through the cube. If you sliced this figure, you would see the cylinder inside the cube.

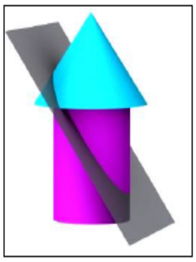
The cutting planes, shown in grey, will have different orientations.



Vertical Plane



Horizontal Plane

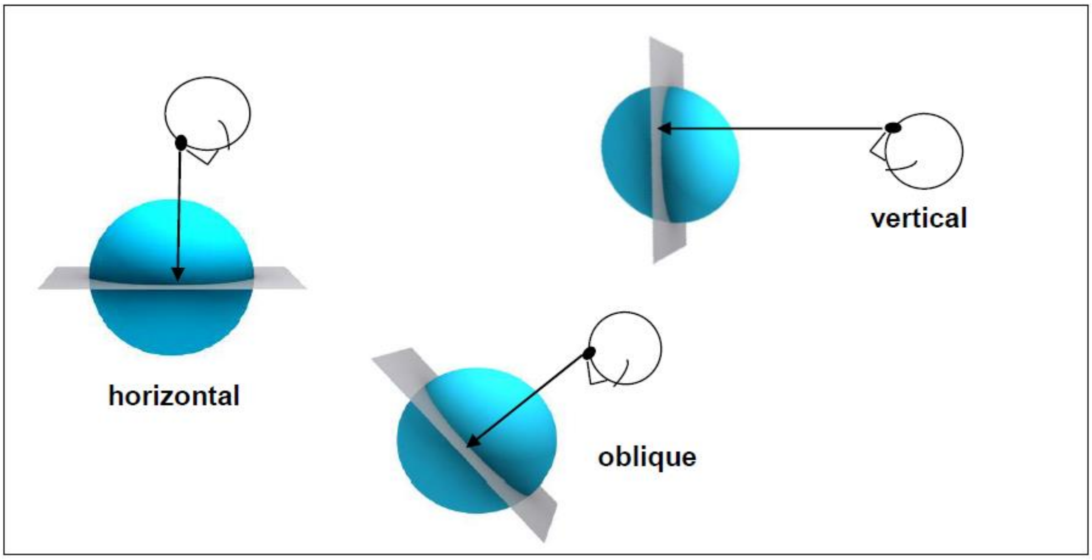


Oblique Plane

You will see three types of cutting planes: horizontal, vertical, and oblique.

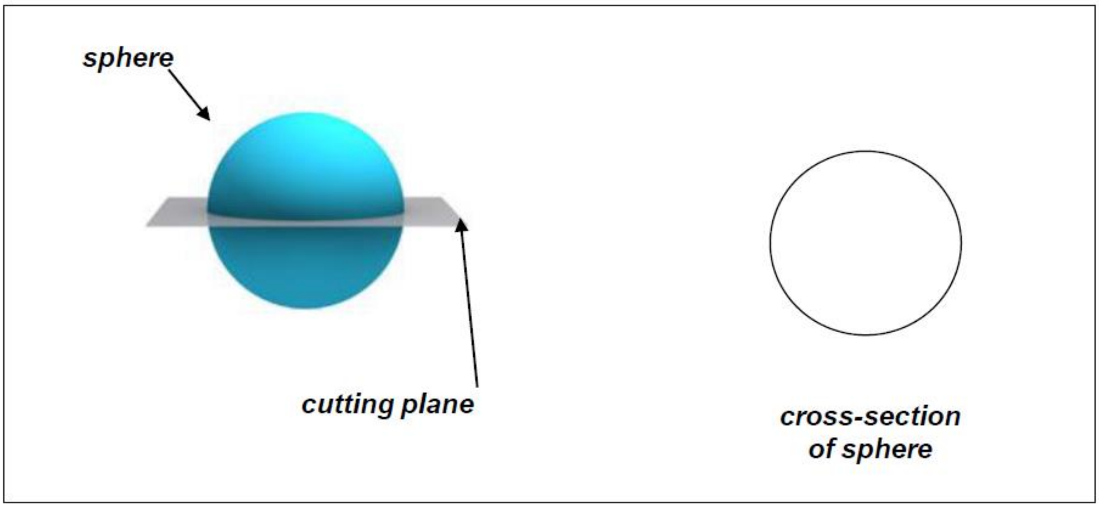
For each type of cutting plane, try to imagine the cross section that would result if you faced the cutting plane head-on, as if you were looking at your reflection in a mirror.

.

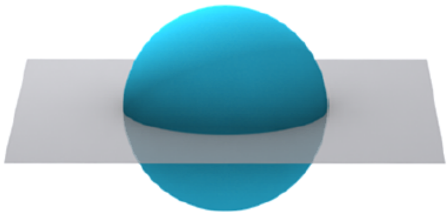


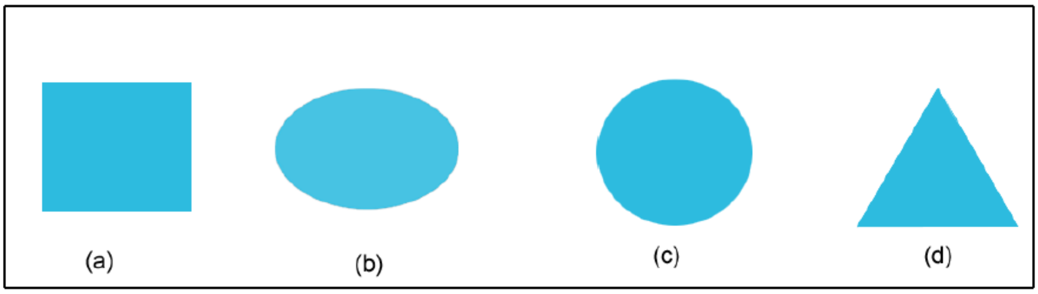
You should also assume that the objects are 6-8 inches tall, and that they are sitting on the desk in front of you.

In the example below, the cutting plane would produce the cross section on the right.



Sample Problem

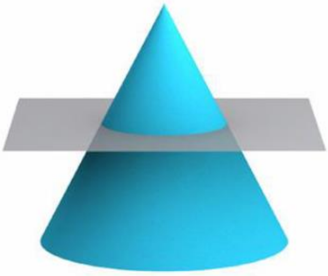


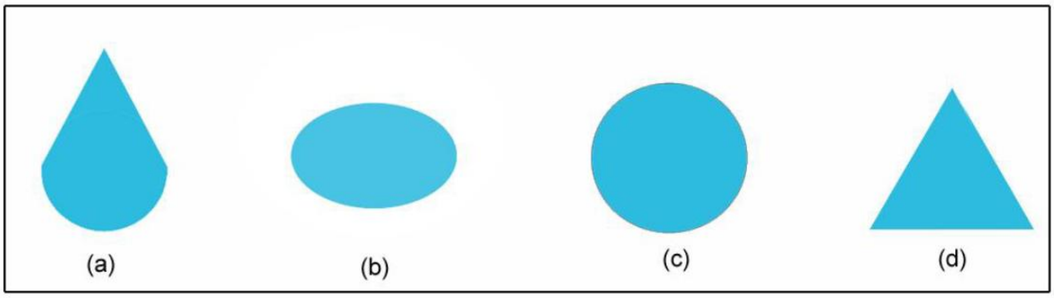


Instructions:

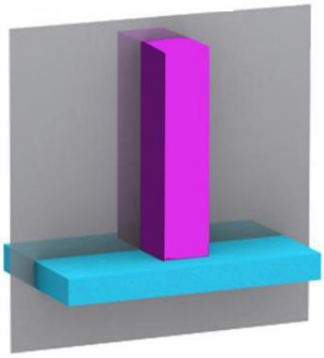
Circle the cross-section you would see when the grey cutting plane slices the object. Imagine that you are facing the cutting plane head-on, as if you were looking in a mirror. Make your choice based on the shapes of the possible answers, not their sizes.

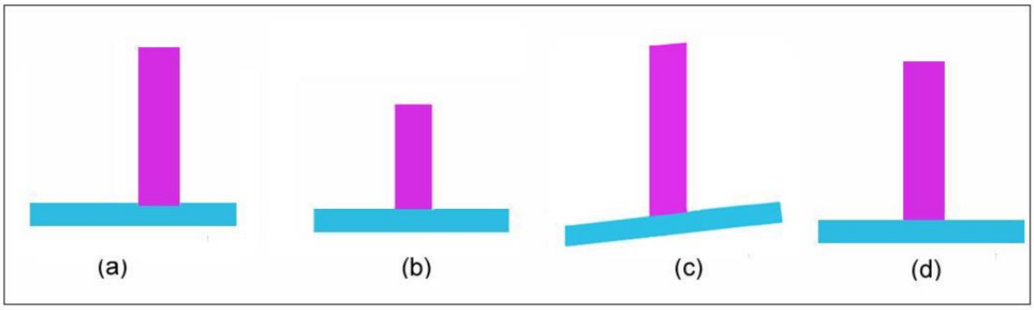
Problem 1



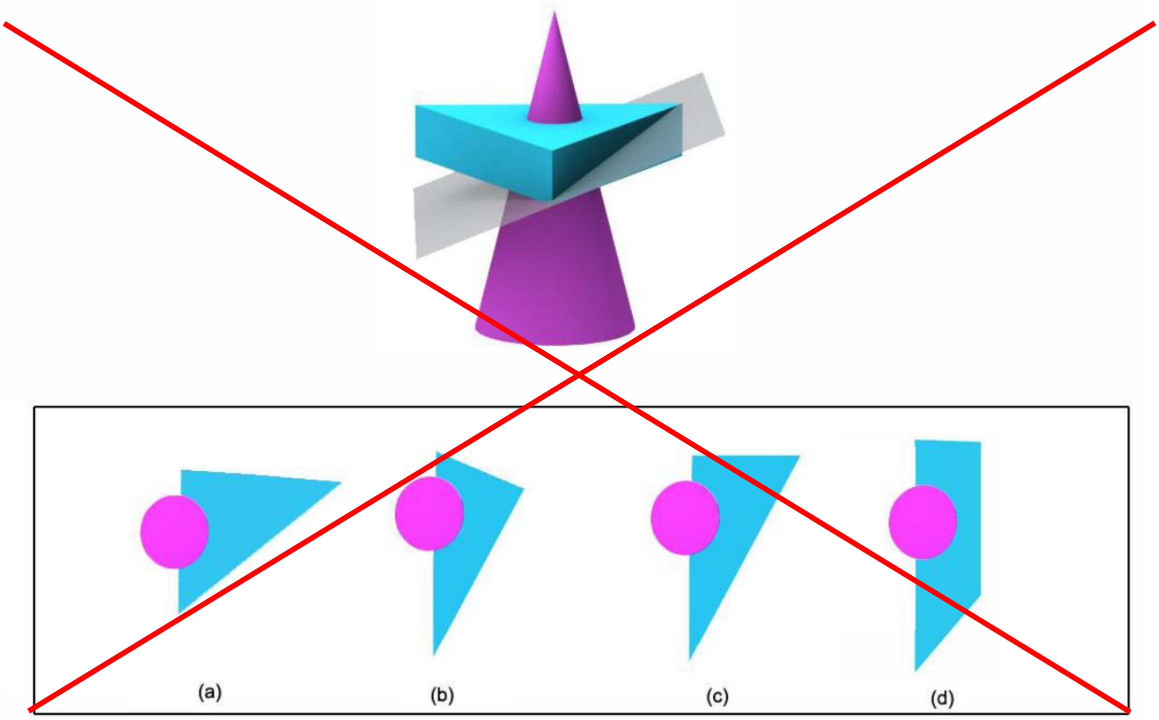


Problem 2

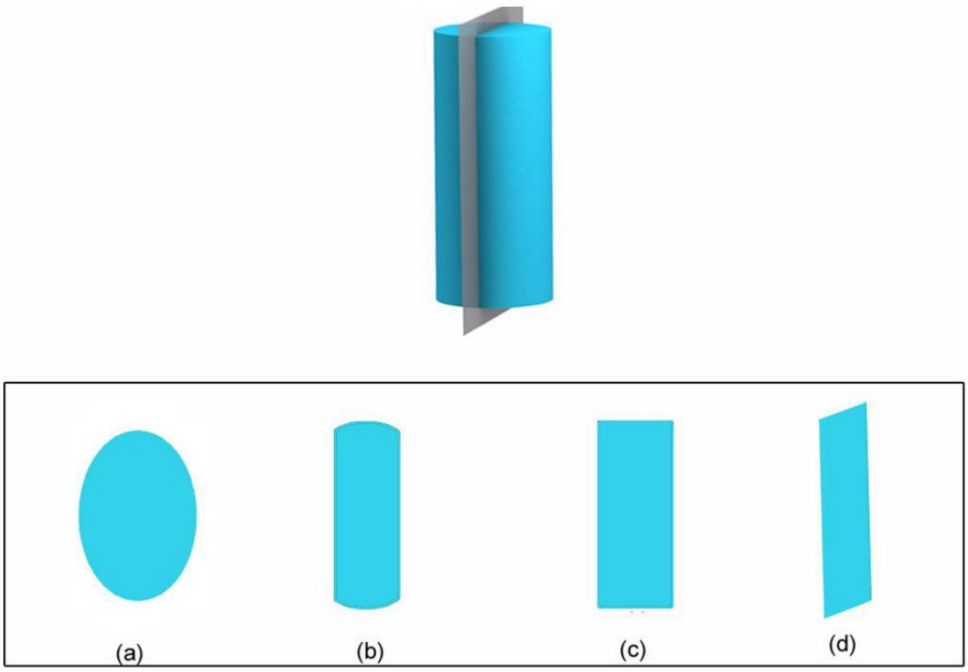




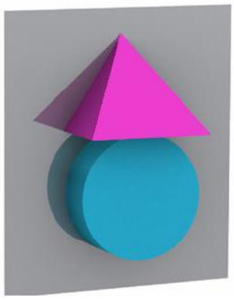
Problem 3

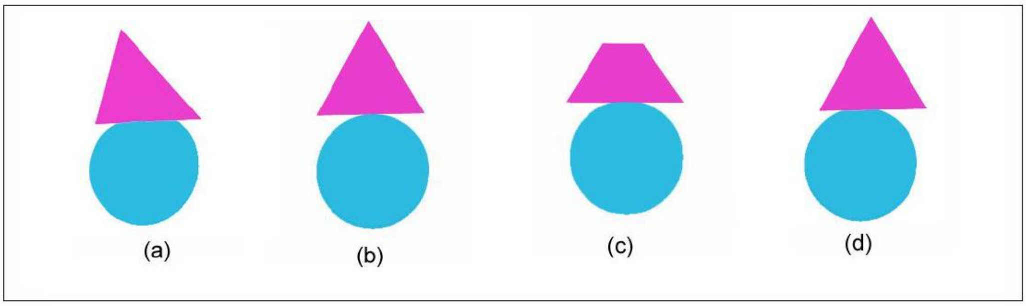


Problem 4

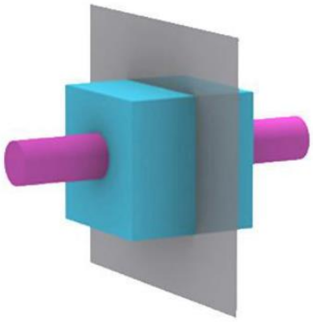


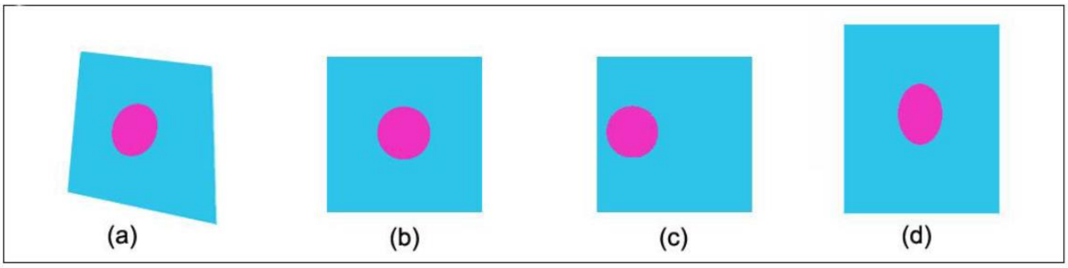
Problem 5



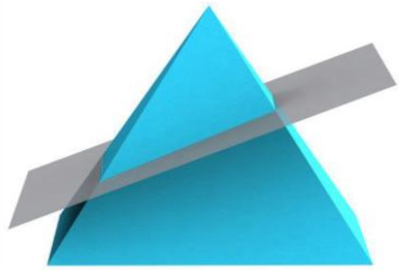


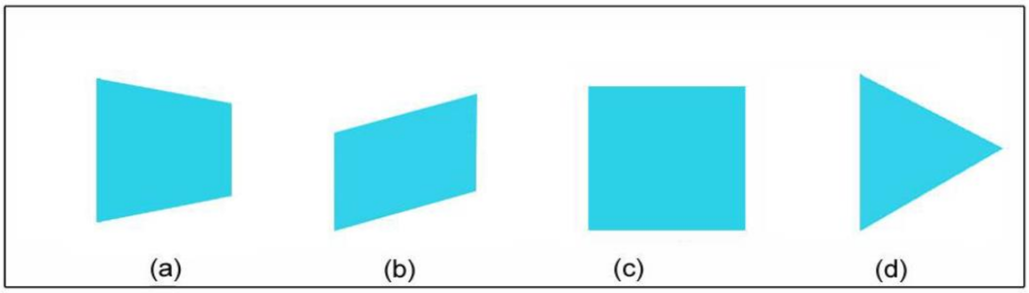
Problem 6



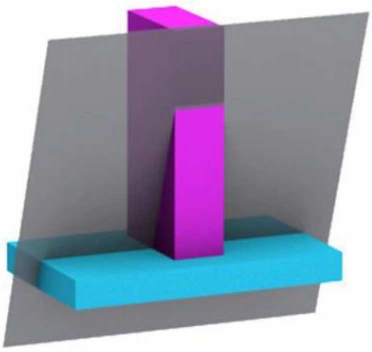


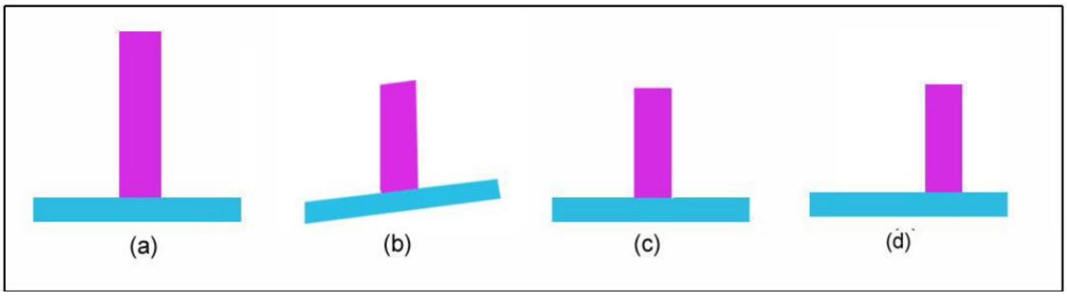
Problem 7



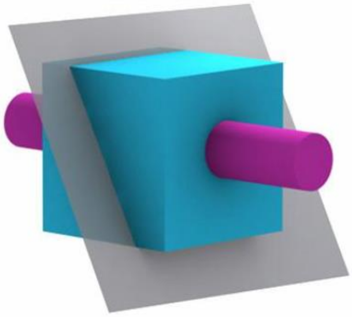


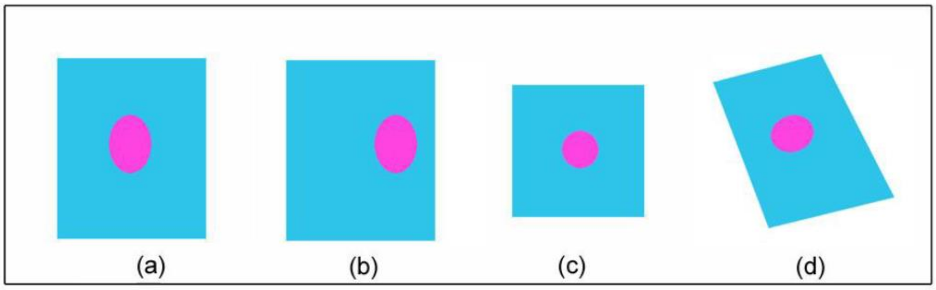
Problem 8



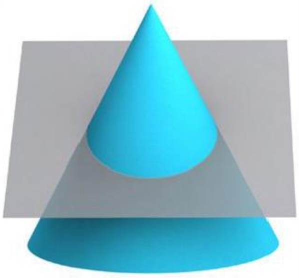


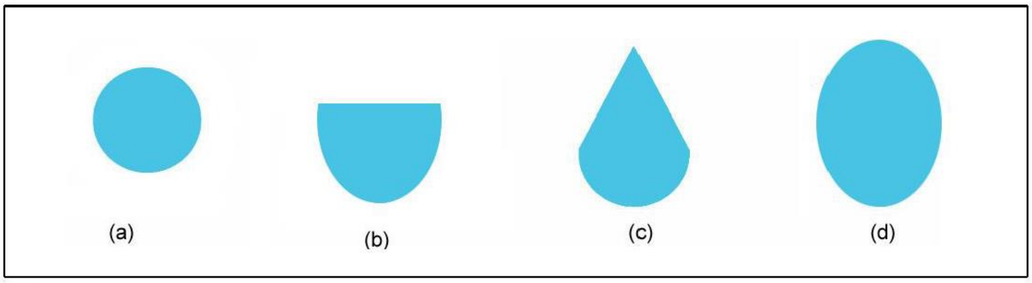
Problem 9



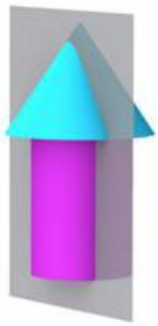


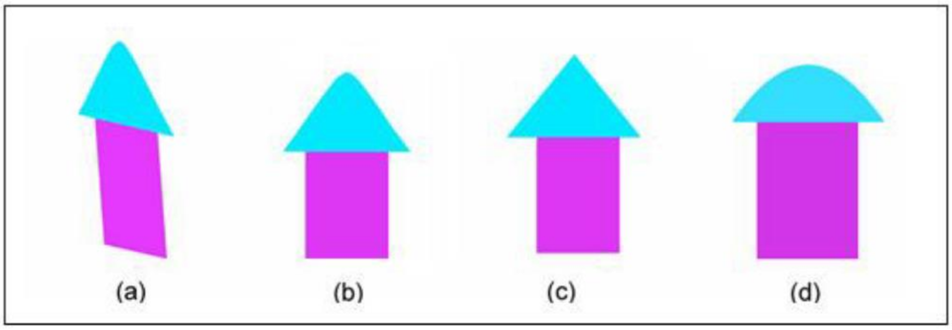
Problem 10



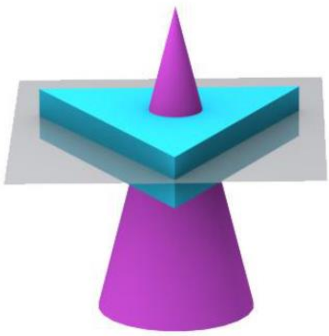


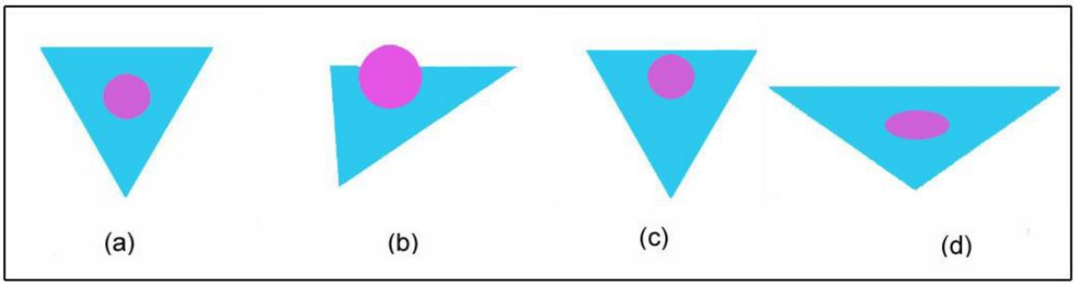
Problem 11



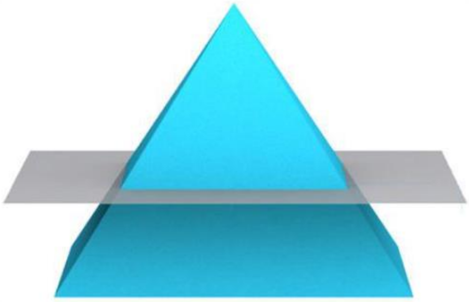


Problem 12



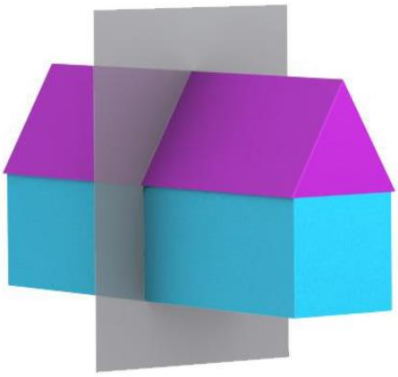


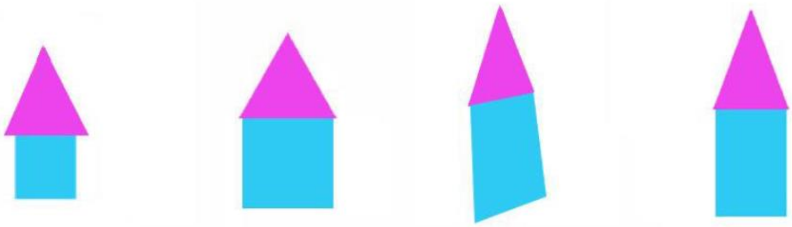
Problem 13



(a) (b) (c) (d)

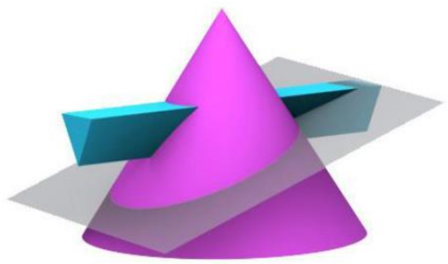
Problem 14

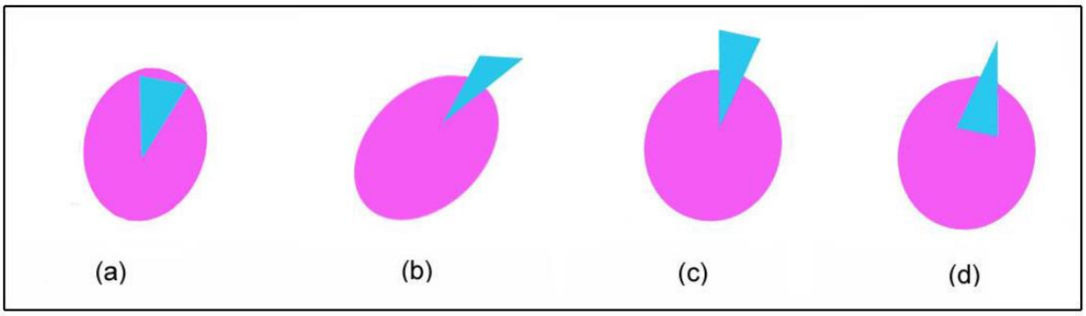




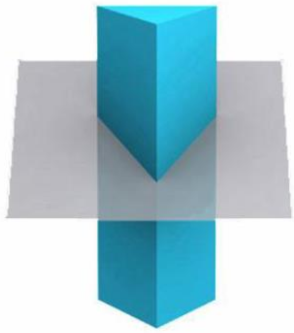
(a) (b) (c) (d)

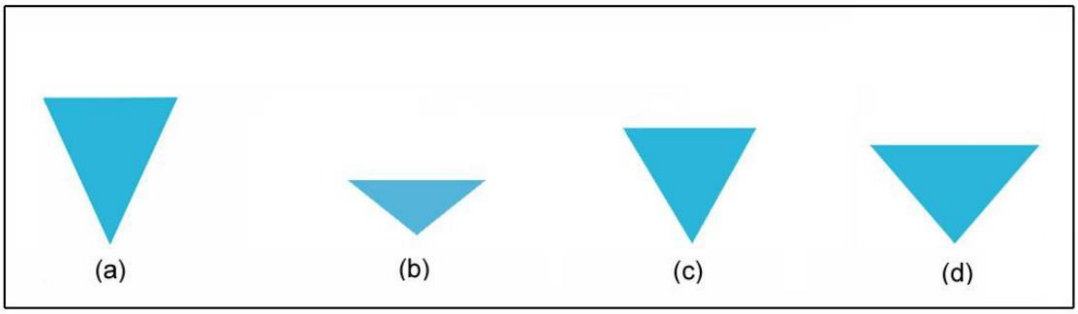
Problem 15



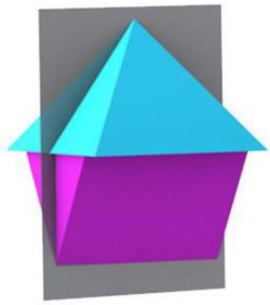


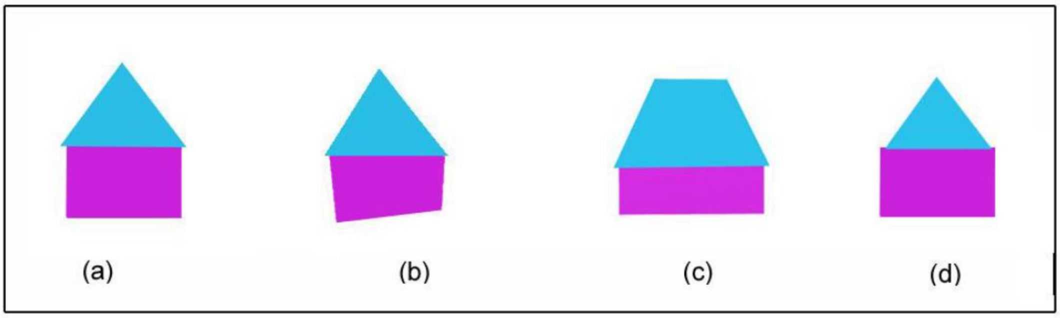
Problem 16



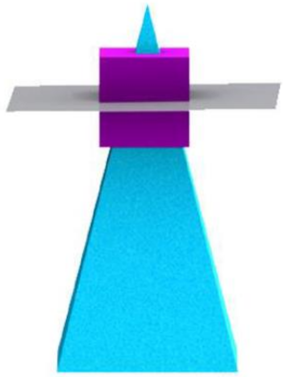


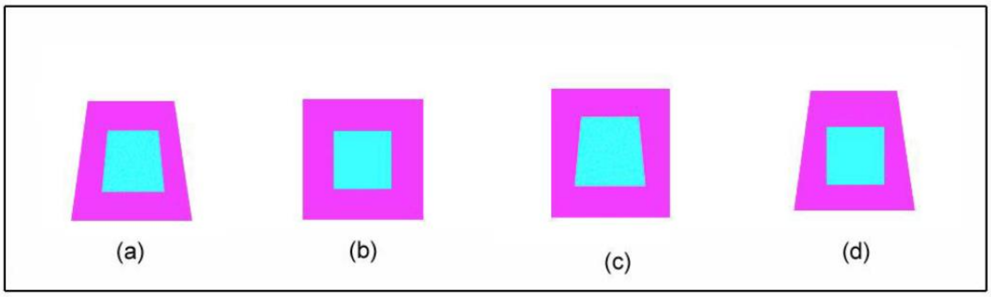
Problem 17



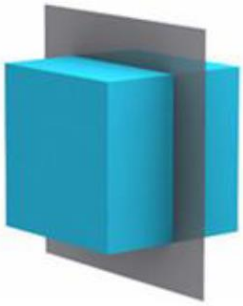


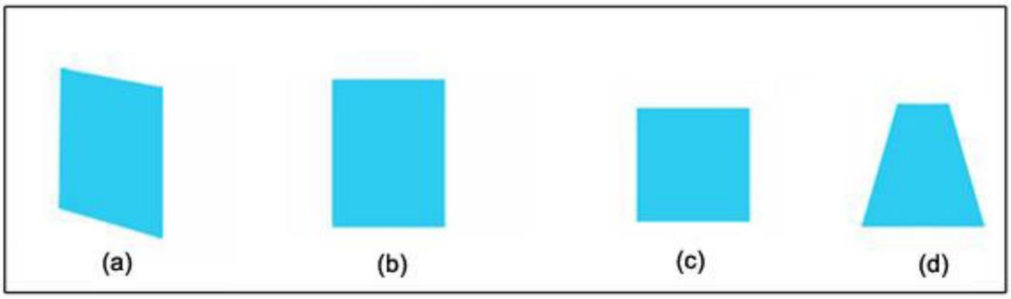
Problem 18



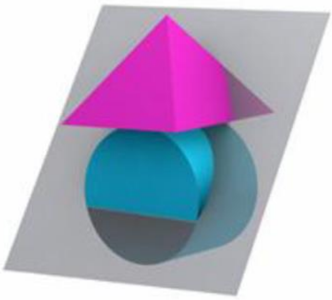


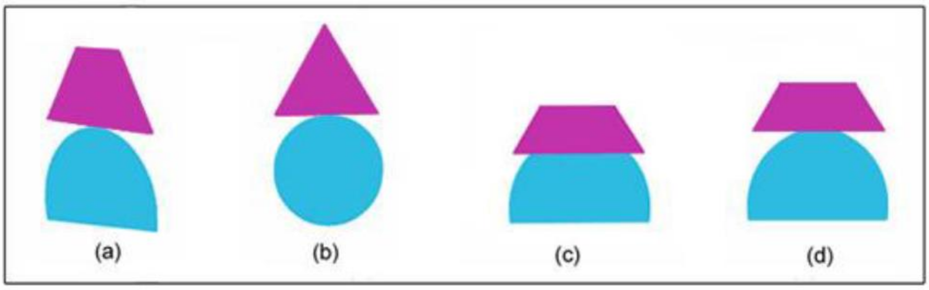
Problem 19



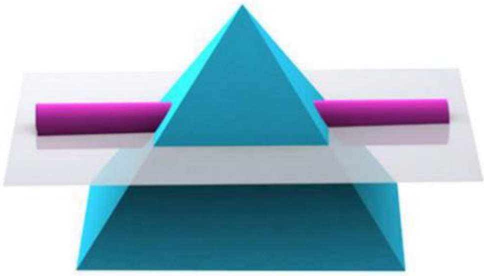


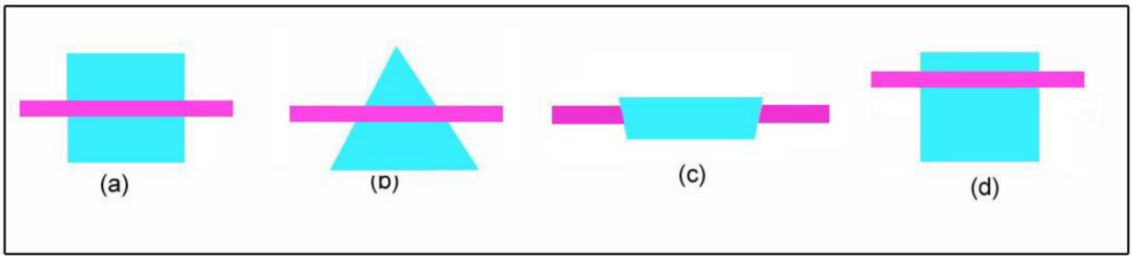
Problem 20



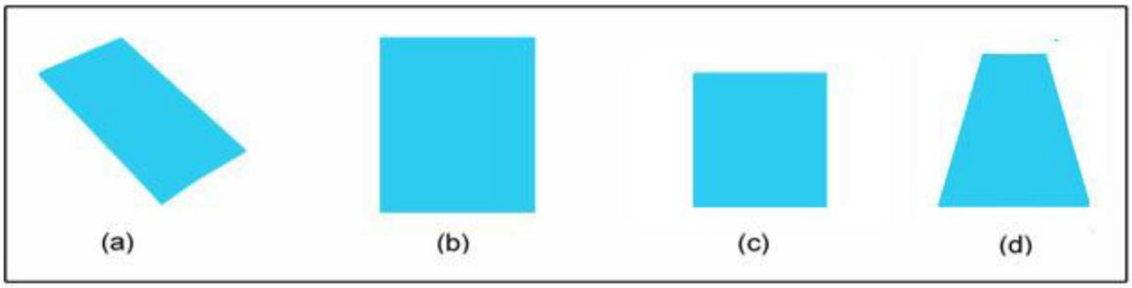


Problem 2 1

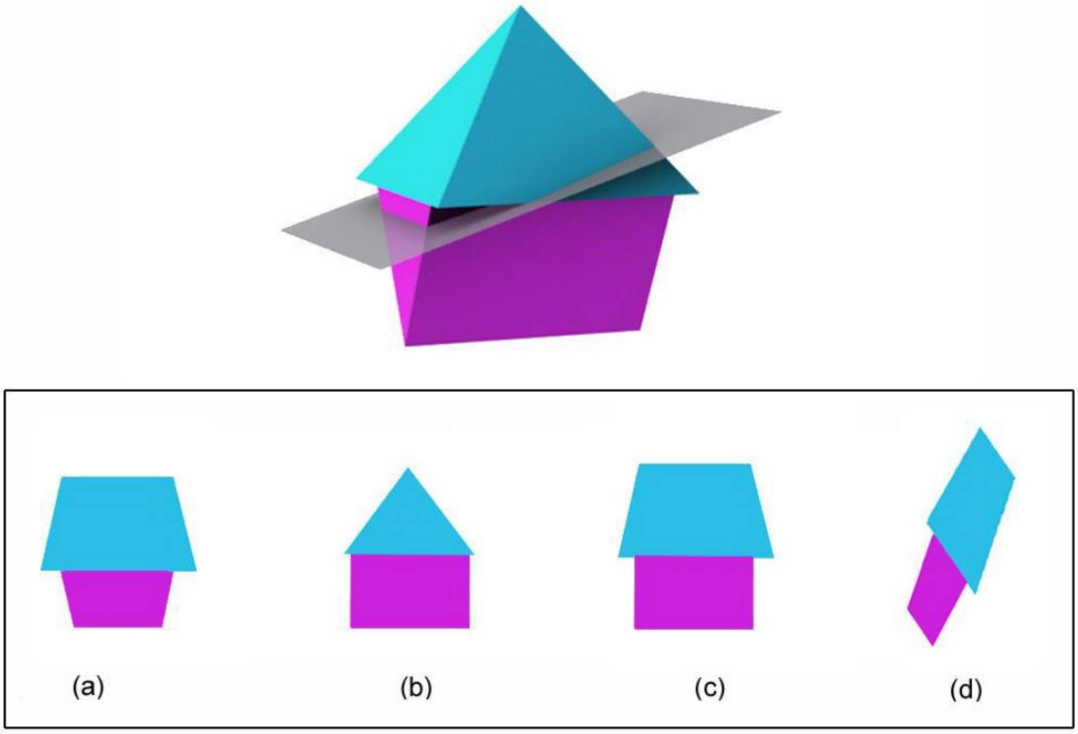




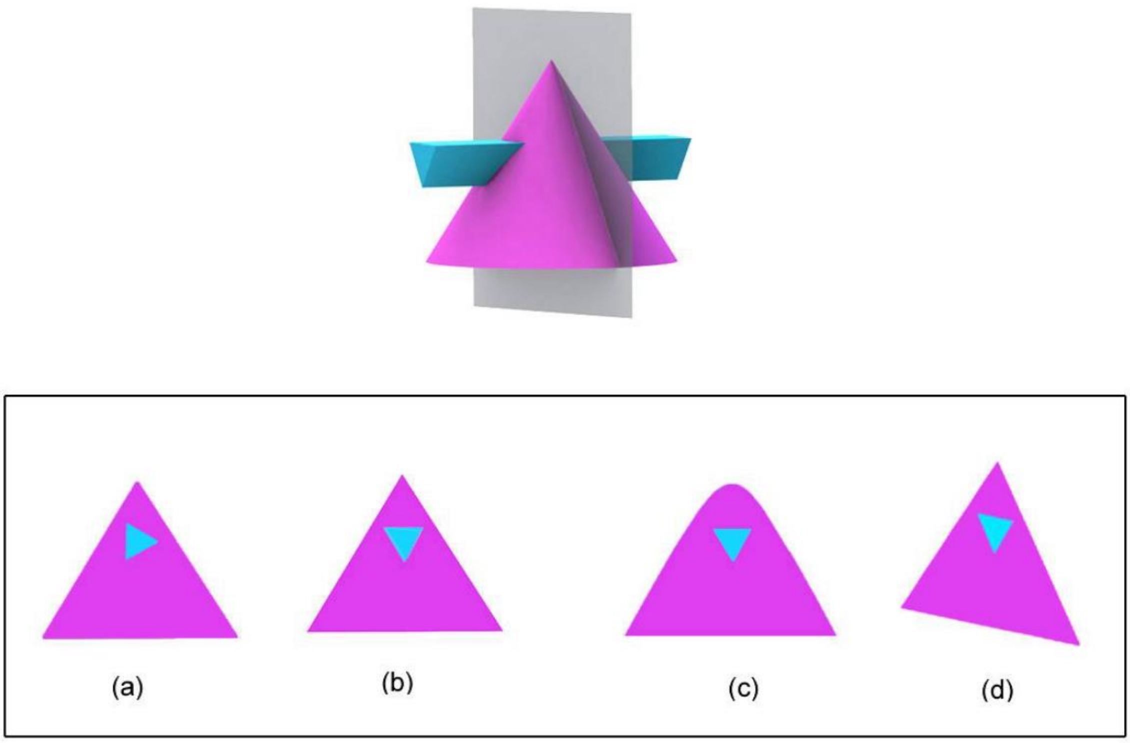
Problem 22



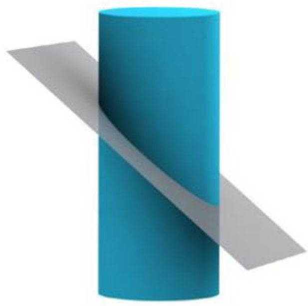
Problem 23

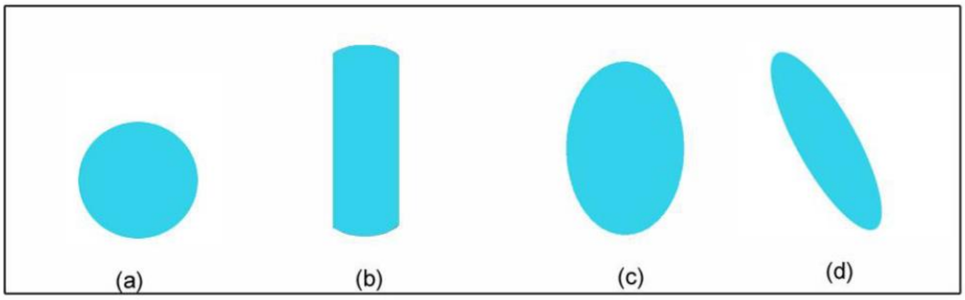


Problem 24

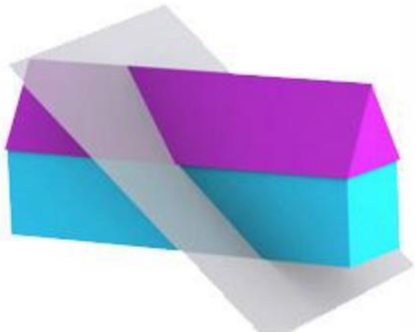


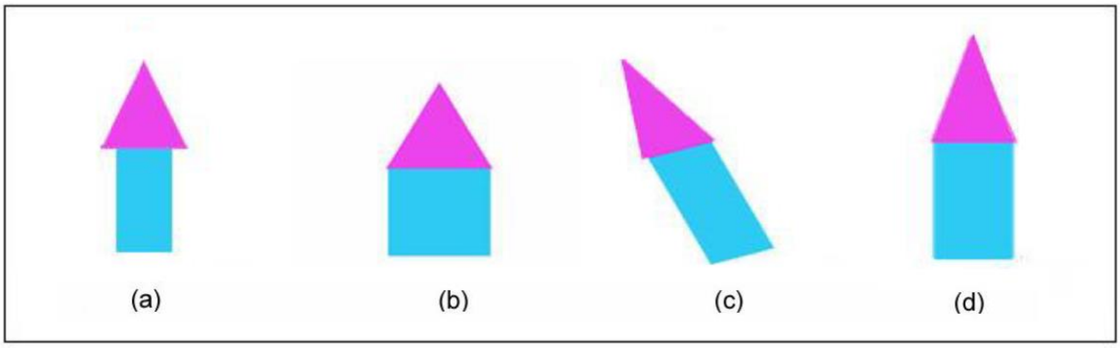
Problem 25



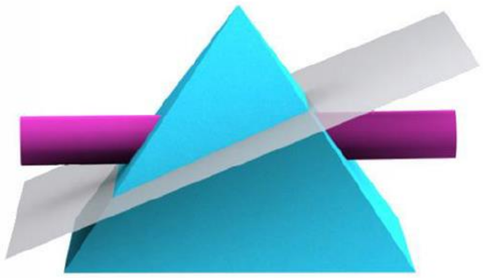


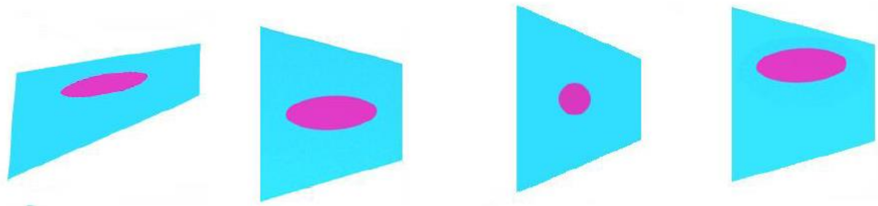
Problem 26





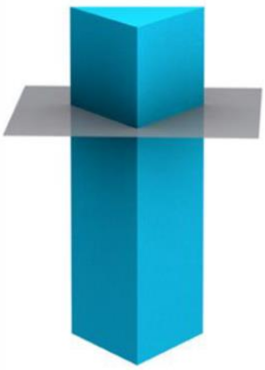
Problem 27





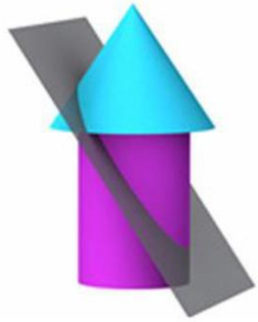
(a) (b) (c) (d)

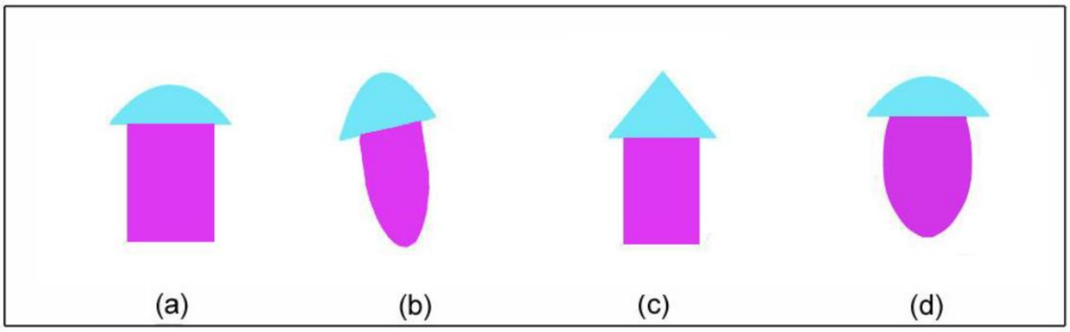
Problem 28



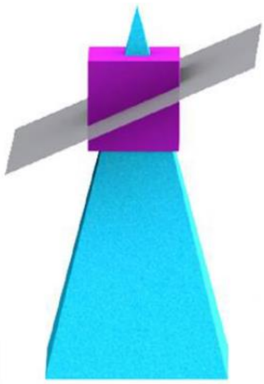
(a) (b) (c) (d)

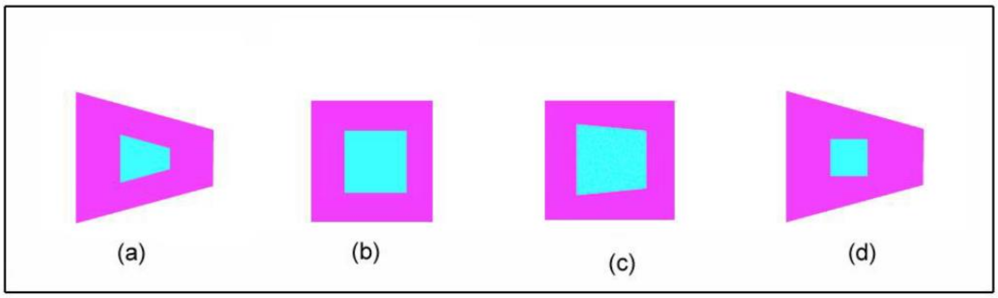
Problem 29





Problem 30





End of exercise

Santa Barbara Solids Test Answer Key

Object (Problem N^{o})	Correct Answer	Egocentric Distractor (Foil)
1	C	B
2	D	C
3	Object Withdrawn by Developers	
4	C	D
5	B	A
6	B	A
7	A	B
8	C	B
9	A	D
10	D	B
11	B	A
12	A	D
13	B	C
14	B	C
15	C	B
16	A	C
17	A	B
18	B	A
19	C	A
20	D	A
21	A	C
22	B	A
23	A	D
24	B	D
25	D	C
26	C	A
27	A	D
28	D	A
29	C	B
30	B	D

Appendix 6 (c) Answer grid for Study 1 online test

POWERPONT® MENTAL ROTATIONS TEST (VERSION A)

For each question, please circle the **TWO** letters of your choice:

Question Number	**Answer Choices**			
1	**A**	**B**	**C**	**D**
2	**A**	**B**	**C**	**D**
3	**A**	**B**	**C**	**D**
4	**A**	**B**	**C**	**D**
5	**A**	**B**	**C**	**D**
6	**A**	**B**	**C**	**D**
7	**A**	**B**	**C**	**D**
8	**A**	**B**	**C**	**D**
9	**A**	**B**	**C**	**D**
10	**A**	**B**	**C**	**D**
11	**A**	**B**	**C**	**D**
12	**A**	**B**	**C**	**D**

POWERPOINT® SANTA BARBARA SOLIDS TEST (Cross sections test)

For each question, please circle **ONE** letter of your choice:

Question Number	Answer Choices			
1	A	B	C	D
2	A	B	C	D
3	A	B	C	D
4	A	B	C	D
5	A	B	C	D
6	A	B	C	D
7	A	B	C	D
8	A	B	C	D
9	A	B	C	D
10	A	B	C	D
11	A	B	C	D
12	A	B	C	D
13	A	B	C	D
14	A	B	C	D
15	A	B	C	D

Appendix 7

Study 1 Demographics Questionnaire

DEMOGRAPHICS QUESTIONNAIRE (April 2011)

Student ID N° _______________

Date_____________

Gender:

Female ☐ Male ☐

Age ☐

Are you left or right handed?

Left ☐ Right ☐

In relation to using computer technology, how would you describe yourself?

Very confident	Confident	Not very confident	Far from confident
☐	☐	☐	☐

Do you play computer games?

Yes ☐ No ☐

If Yes, how oftern do you play?

Daily	Weekly	Monthly	Less than monthly
☐	☐	☐	☐

How long (approximately) have you been playing computer games?

< 6 months	1 year	2 – 5 years	6 – 10 years	>10 years
☐	☐	☐	☐	☐

What type of computer games do you play (or have played) ? (Please circle all that apply)

3 D first person action	City-building games
Adventure	Arcade
Educational	Maze
Music	Pinball
Platform	Puzzle
Stealth	Fighting
First-person shooter	Role-playing
Multiplayer Online Games	Simulators (eg Flight, Racing)
Sports	Military / Space Strategy
Strategy wargames	
Other (please specify)	

Appendix 8

Usability questionnaire & free text responses for online test instruments

Questionmark Perception Usability Questionnaire

Date: 27th April 2012

The PC based test instructions were clear

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
☐	☐	☐	☐	☐

The PC based test objects were easy to see

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
☐	☐	☐	☐	☐

I preferred the PC based mental rotation test compared to the paper test

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
☐	☐	☐	☐	☐

I preferred the PC based cutting planes test compared to the paper test

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
☐	☐	☐	☐	☐

If you would like to add any comments to support your answers please do so overleaf

Free text comments with coding:

Microsoft PowerPoint free text comments (June 2011)

1. The computer images are sharper but the lines (at?) the images are still incomplete which my eye finds distracting and confusing. Particularly if a line is going away or coming towards me in test 1.
2. I preferred the mental rotation on the PC, I preferred the cutting plane test on paper
3. Found the slides hard
4. Thank you!

QuestionMark Perception free text comments (April 2012)

1. The larger screen & clearer images made the PC based test better than the paper version
2. The test seemed much easier on the PC than on paper – I didn't struggle as much with the PC test
3. I'd like to retake this test using a PC but having a blank piece of paper to draw the images & draw how I think they would look rotated to help me choose my answer
4. Had to scroll down to see some of the images ie too big
5. I felt it was easier to imagine the objects with the paper diagrams than using the computer version
6. The size of the object was fine
7. The fact that I could see the clock made me more stressed
8. I found the fact that the target was further from the figures (answer options) disturbing
9. Cutting plane test: having to scroll to see the examples was disturbing
10. It took a while for me to see the rotational ones on the computer but once I had the hang of it it was much easier for me than the paper one
11. I did find the ticking clock disturbing as it was (I feel)pushing me to go faster it would have been better to have just the minutes and seconds in say increments of 30 or 15 seconds
12. The images in the cutting exercise was too big I had to zoom out, which was distracting, would have been better to have the image of the whole shape bigger than the cross section shapes smaller

Analysis phases:

Phase 1 - Initial Reading – identify broad themes

Phase 2 - Second Reading –sub themes and experiences

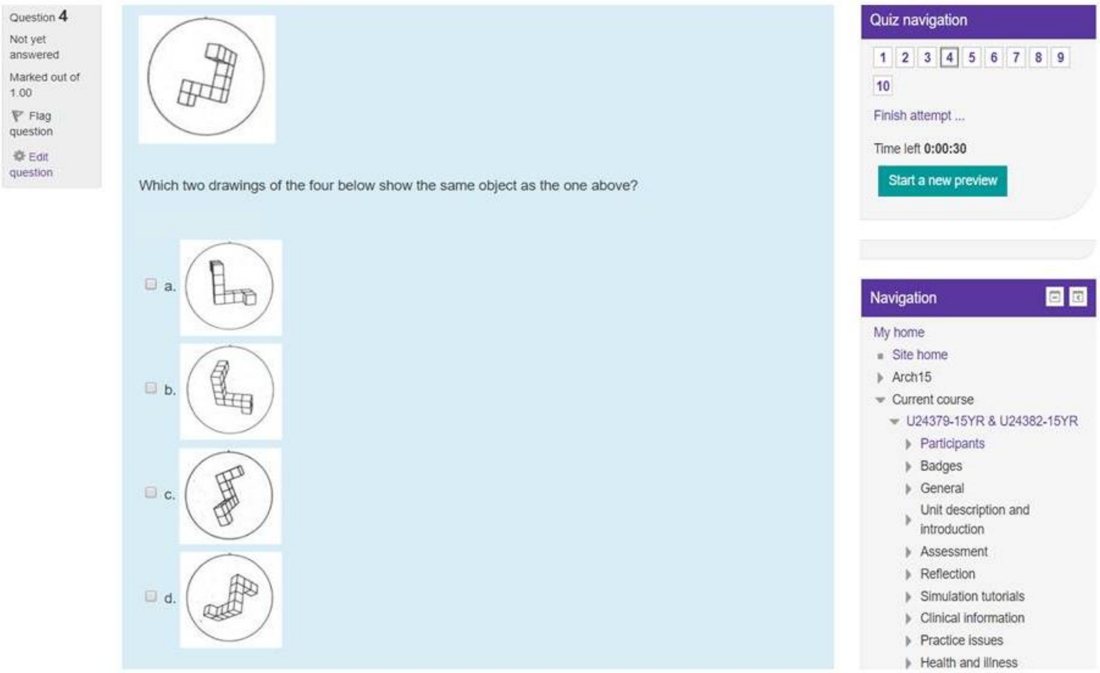
Phase 3 – Review and linking of phase 1 themes with phase 2 observations

Summary of code words and themes				
Phase 1 Broad Themes	PC	Paper	Images	Clock
Phase 2 Linked Themes	Preferred mental rotation	Preferred cutting plane	Sharper	Stressed
	Clearer	Easier	Incomplete	Ticking
	Better		Distracting and confusing	Disturbing
	Much easier		Found slides hard	Pushing me to go faster
	Didn't struggle as much		Cutting exercise too big	
	Rotational ones- Took a while to see		Distracting	
	Easier than paper		Whole shape bigger / cross section smaller	
Phase 3 Summary			Clarity, display, layout, size	Disturbing/ stressing and feeling rushed

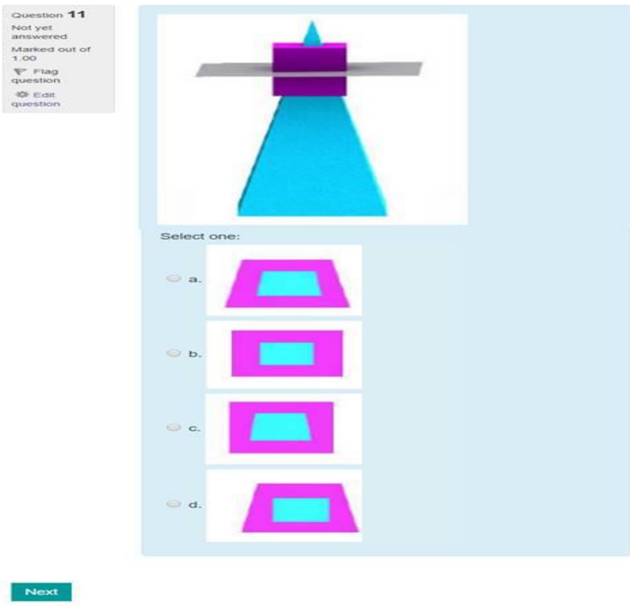
Appendix 9

Study 4 Moodle Quiz Screen Shots

Mental Rotation Test Object



Santa Barbara Solids Test Item 18



Appendix 10

Study 6 Demographic, preferred hand and spatial activities survey

Previewing Questionnaire

SPATIAL VISUALISATION SKILL STUDY DEMOGRAPHICS QUESTIONNAIRE

Page 1

1 Are you

○ **Male** ○ **Female** ● **No answer**

2 **State you age**

Page 2

3 Please indicate the category that best describes which hand you use for each of the activities listed below

		Always left	Usually Left	No Preference	Usually Right	Always Right
Writing	●	○	○	○	○	○
Throwing	●	○	○	○	○	○
Using scissors	●	○	○	○	○	○
Holding Hairbrush	●	○	○	○	○	○
Cutting with Knife (without a fork)	●	○	○	○	○	○
Holding a spoon	●	○	○	○	○	○

Striking a match	(•)	( )	( )	( )	( )	( )
Using a computer mouse	(•)	( )	( )	( )	( )	( )

Navigation

Administration

Page 3

4 Do you play computer games?

If you selected YES then please answer questions 5-8 on the next page. If you answered NO please skip to question 9.

( ) **Yes** ( ) **No** (•) **No answer**

Page 4

5 How often do you play?

( ) **Daily** ( ) **Weekly** ( ) **Monthly** ( ) **Occasionaly** (•) **No answer**

6 How long ago did you start playing?

( ) **< 6 months** ( ) **1 year** ( ) **2-5 years** ( ) **6-10 years** ( ) **>10 years**
(•) **No answer**

7 What tyoes of games do you play? Please select all that apply

- [ ] **First person action**
- [ ] **Adventure**
- [ ] **First-person shooter**
- [ ] **Puzzle**
- [ ] **Role-playing**
- [ ] **Maze**
- [ ] **Arcade**
- [ ] **Multiplayer Online**
- [ ] **Simulators (eg Flight, Racing)**
- [ ] **City-building games**
- [ ] **Educational**
- [ ] **Pinball**
- [ ] **Strategy wargames**

8 **Others please state:**

Page 5

9 Please indicate your participation in the following activities:

Navigation

Administration

		Never	In the past but not now	Less than once a week	Once per week	More than once a week
Team sport						
Individual sport						
Drawing in 3D perspective						
Mechanical/Technical drawing						
Arts and Crafts						
Juggling/ Baton twirling						

Page 6

10 Which types of toy did you play with when you were youger? (please select all that apply)

- **Action figures**
- **Arts / Crafts**
- **Construction toys (eg building bricks)**
- **Model building**
- **Puzzles**
- **Dolls / Puppets**
- **Electronic hand held games**
- **Board games**

11 **Other please state**

Submit preview Reset