BODY IMAGE DISTORTION IN PHOTOGRAPHY

Bernard Harper

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy.

June 2006

Department of Psychology
University of Liverpool

"Then too there is the oft-repeated statement that women in photographs appear shorter and stouter than in real life. Frankly I do not see it that way, but perhaps those who are not trained to analytic vision are so deceived. *At any rate I have had others speak repeatedly of the fact that stereo does not have this effect. I have had several women say about a stereogram, "Oh! I like that. It doesn't make me look as fat as I usually do in a photograph!"*

Three Dimensional Photography: The Principles of Stereography by Herbert C. McKay (1953) p. 63.

Acknowledgements

I would like to thank Richard Latto for his patience and endless support from beginning to end of this work. I am also grateful to Nick Lodge and David Harrison from the Independent Television Commission for their willingness to support speculative studies in a completely novel area of research. All of the computer generated images in this thesis were designed and built by Philip Berridge. They were innovative and technically brilliant creations that were vital to the success of these experiments. Special thanks go to Rachel Perry. Her ability to recruit models for the swimsuit photography made this task much easier than was first envisaged. Gabrielle Salfati not only provided moral support but also acted as fine model too. Stacey Conchie was another model who also a tirelessly supportive colleague. Rob Black, Laurence Argyle and Christina Groom gave great assistance too in the running of experiments, working with the equipment or just being there when I needed them. Special thanks must also go to David Brooks for teaching me photography and making me believe that there were still innovative experiments to be done in this field. I would also like to thank John Manning and David Brodie for encouraging me to conduct useful experiments, rather than just take photographs. I would finally like to thank all of the models and participants who happily gave their time and support to this work. I hope they agree it was worth it.

CONTENTS

List of Figures

List of Tables

Abstract.

This thesis investigates the theory that photography is, in terms of body image perception, an intrinsically distorting and often fattening medium. In the professional practice of photography, film and television, there is a widely held belief that the camera "adds 10lbs" to the portrayed weight of actors and presenters. The primary questions addressed here relate to the true extent of the fattening effect, to what perceptual mechanisms it can be ascribed and if it can be counteracted in common practice.

Current theories in the perception of photographic images rarely, if ever discuss the medium's perceptual accuracy in recording the original scene. It is assumed by many users that most photographs convey essentially the same information they would have seen had they been present when they were taken. Further, it is generally accepted that photographs are an accurate, veridical and scientific method of record and their content should be trusted unless there is evidence of a technical failure, editing or deliberate tampering. This thesis investigates whether this level of trust is appropriate, specifically by examining the reliability of photography in relation to reproducing the face and form of human subjects.

Body Image Distortion (B.I.D.) is a term normally used to describe the primary diagnostic symptom of the slimming disease, anorexia nervosa. However, it is demonstrated here that people viewing 2D photographic portraits often make very significant overestimations of size when comparing otherwise identical stereoscopic images. The conclusion is that losing stereoscopic information in conventional 2D photography will cause distortions of perceived body image, and that this is often seen as a distinct flattening and fattening effect. A second fattening effect was also identified in the use of telephoto lenses. It is demonstrated, using psychophysical experiments and geometry that these 2D images cannot convey the same spatial or volumetric information that normal human orthostereoscopic perception will give. The evidence gathered suggests that the Human Visual System requires images to be *orthostereoscopic*, and be captured using two cameras that mimic as closely as possible the natural vergences, angle of view, depth of field, magnification, brightness, contrast and colour to reproduce scenes as accurately as possible.

The experiments reported use three different size estimation methodologies: stereoscopic versus monocular comparisons of human and virtual targets, bodyweight estimations in portraits taken at differing camera to subject distances and synoptic versus direct viewing comparisons. The three techniques were used because photographic images are typically made without disparity and accommodation/ vergence information, but with magnifications that are greater than found with direct viewing of a target. By separately analysing the effects of disparity, magnification and accommodation/vergence the reported experiments show how changes in each condition can effect size estimation in photographs. The data suggest that photographs made without orthostereoscopic information will lead to predictably distorted perception and that conventional 2D imaging will almost always cause a significant flattening and fattening effect. In addition, it is argued that the conveyed jaw size, in relation to neck width is an important factor in body-weight perception and this will lead to sexually dimorphic perception: disproportionately larger estimations of bodyweight are made for female faces than male faces under the same photographic conditions.

CHAPTER ONE

Introduction

1.1 Image distortions in photography

Photography, as a professional discipline, is over one hundred and seventy years old (1)[1] . And from its inception, people have been enthralled by the immediacy and impact of the photographic image. The seemingly perfect photographic perspectives, textures and contours of even the earliest images were somehow more natural and accurate than the finest representations an artist could reproduce on canvass. It was obvious to almost everyone in the Victorian era that photography was the wonder of the age (Rosenblum, 1984) and that life-like painted representations had had their day. This was despite the fact that painters had produced viewing devices, colour pallets and application techniques specifically intended to reproduce the reality of direct vision and the "delights of seeing" with uncanny accuracy. Yet photography, even without colour reproduction, modern panchromatic film sensitivity or aberration corrected optics, easily convinced the public of its supremacy. The fact that the film emulsions available were so insensitive to light (portrait exposures in a daylight studio were often over one minute long) seemed to matter little to the paying customers. Head clamps had to be used to stop portrait subjects moving during the exposure. The blue sensitive emulsions had very high contrast, rendering the fine lines and underlying venous structure of some adult faces as pockmarked and aged. And when people complained that their portraits were unflattering, the photographer could counter with the phrase "But everyone knows that the camera *never* lies."

[1] All World Wide Web derived references will be numbered in this way and listed by this number in the bibliography.

Photographers should be aware that the camera *always* lies to some degree. Truly accurate and lifelike reproduction is almost impossible to achieve, even with the most advanced full colour holograms (Marks 1999). For photographers to mislead to a portrait subject and convince them that the fault lay with themselves and not with the technology was unfortunate. However, the technical knowledge available at the time was so limited that the only route available to professional photographers was to continue with the deception and hope that scientific research would somehow catch up with their assertion. Sadly, they would need to wait for a very long time. This thesis will show that one and a half centuries of continuous study and development later, we are still far from this goal.

The immediacy, colour and depth of reality are so complex that producing a non-distorting imaging technology is still a very distant dream (2). Yet in the almost universal activity of photographic portraiture, many people still seem to trust the integrity of the mediated image. If people look overweight in photographs, it is still reasonable for most people to presume that they are overweight in reality. This thesis details seven pilot studies and eight experiments that, in principle, could have been conducted at any time in the scientific history of photography. It is open to speculation as to why the photographic and film industries were satisfied that photography was non-distorting, and did not systematically investigate the fattening effects of the photographic image. It has been the implicit assumption of users of photography that the body images portrayed in photographic portraits are essentially truthful. Yet the experiments reported here challenge that assumption and hopefully offer a new understanding of how the human visual system interprets two-dimensional and three-dimensional mediated and directly viewed scenes.

It may be that if the pioneers of commercial photography had made a concerted effort to reproduce the fidelity of human vision in a photomechanical technology, then the distorting effects of photography would be more widely known than they are today. However, most of the important photographic research has been funded by large photographic companies whose short-term product-oriented goals may have skewed fundamental research in this area. While the author has found no evidence to the contrary, it does seem unlikely that some of the more obvious perceptual effects reported here have not been observed by photographic theorists or researchers in a position to be able to investigate them. Throughout most of the 20^{th} century, multi-national corporations such Kodak, Polaroid, Fuji and Agfa have filed countless thousands of patents that attest to their intensive photographic research and development programs. Of the Top 100 private sector patent recipients in 2002 (3) photographic and imaging-related companies predominate. Eastman Kodak owns the most photographic patents of any company and is eighth of the all-time list of patent owning companies. (4)

However, most of the patents filed relate to the protection of highly commercial processes and products. Despite searching the patent office lists and extensively on the web, it is surprising to find that few of their patents relate to human-factors research and none could be found that derive from basic research into visual perception. Though Kodak has been the major publisher of technical papers in photographic theory and practice, they seem to have done almost have no work in quantifying the differences between direct vision and the lifelike reproduction of people, objects and scenes, with the notable exception of colour fidelity.

1.2 Photographic fidelity and accurate colour reproduction

Colour discrimination is an important part of human existence. Whether it is choosing between unripe or ready fruits, deadly or tasty mushrooms, live or neutral wires, stop or go lights, toxic or safe substances, colour perception is understood to be fundamental aspect of most peoples everyday lives. Yet the RGB colour space that is most commonly used to artificially represent natural colours can only show approximately ten percent of the over one million brightness and colour combinations that most people can discriminate (Evans 1959). A simple but surprising example of this inadequacy sometimes referred to as "metameric failure" (5) can be found with the photographs found on almost any packet of flower seeds. People often buy products using the colour photograph on the packaging as a guide. They expect and even rely on the fact that the photograph will show an accurate colour reproduction of the product, in this case an adult plant. For obvious commercial reasons, it is vital that this image is a good match. Colour photography however is very poor at providing adequate representation of subtle colourations and natural hues. Photographers call this phenomenon "subject colour failure" and accept it as a fundamental limitation of their medium. Seed growers overcome the problem by using hand colouring over a monochrome rendition of the adult plant. They compare each colour on the adult plant against a physical colour chart, often from the Pantone printing ink colour space. These colour values are conveyed to the printers who then ensure that the correct dyes are printed over the monochrome photograph to ensure a good match. Commercial colour printers rarely use the RGB colour space, as they have understood its limitations for decades. Most photomechanical reproduction systems are based on the CMYK colour space, yet can have up to sixteen colour layers in the final print. Even then, it is known that there are some combinations or types of colours that the best systems still cannot reproduce.

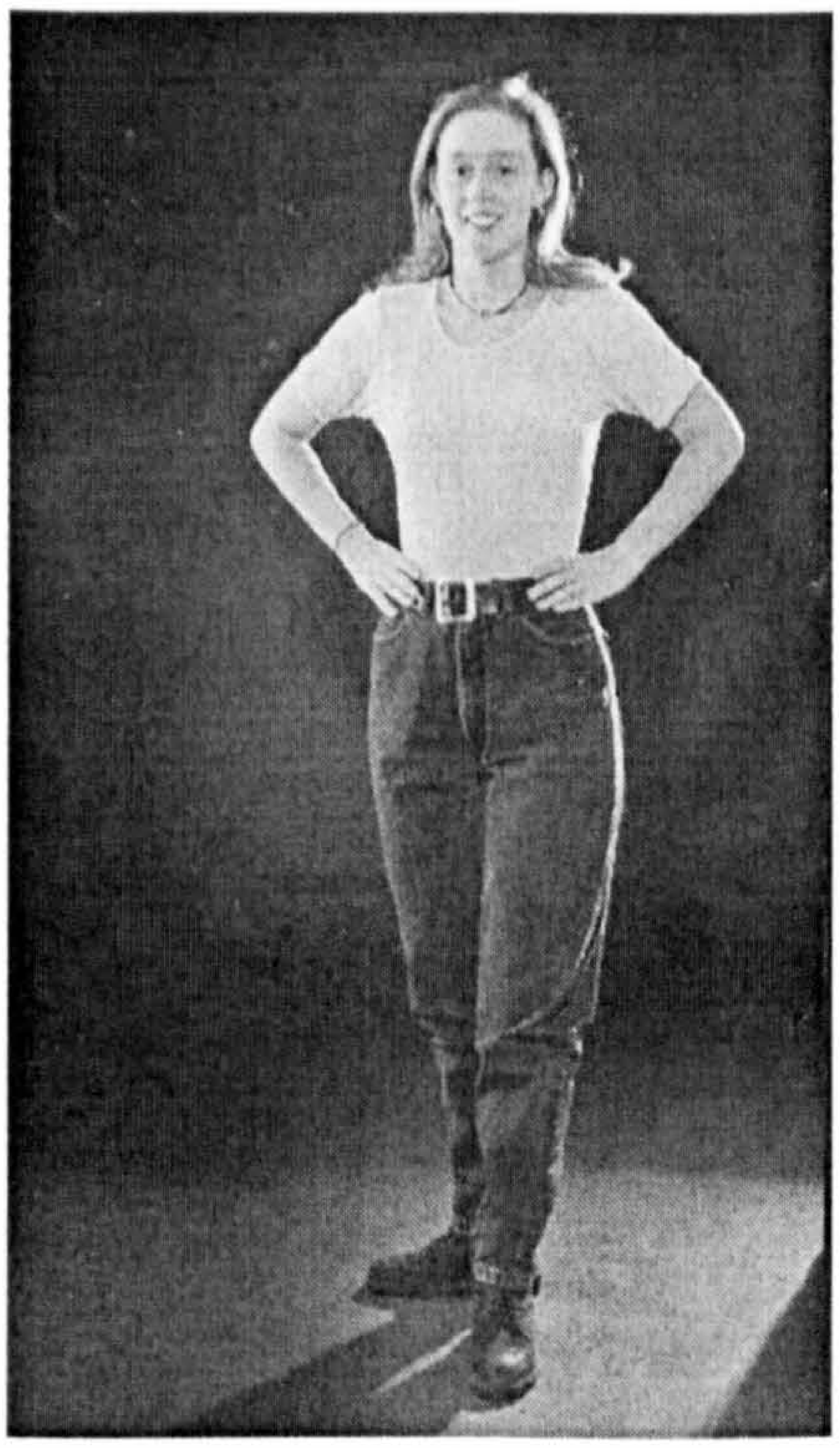
The Internet uses the restricted sRGB colour gamut to convey its photographic images. So it is hardly surprising that the selling of flowers, clothing, paintings and other artworks seems to have been commercially hindered by this lack of image fidelity. It is likely that eventually all visual media will use a colour space that is derived from a more fundamental understanding of the human visual system. It would then be possible for people to accurately sense remote stimuli that previously could only be experienced by direct contact. For instance, doctors could diagnose many illnesses without physically being there, if imaging systems could convey the information they require. If they could do this remotely, then doctors could be telepresent almost anywhere and the time saved between accidents and treatment could save many lives. It is generally accepted in advanced imaging research that no current electronic imaging system can achieve these levels of fidelity (6). However, some photographic film-based technologies do offer levels of colour reproduction that can sometimes give the impression of life-like image quality. Transparency films, such as those provided by the Kodachrome, Ektachrome or Fujichrome film technologies have good colour reproduction and as such offer the best opportunity for photographers to produce high quality imaging in a single stage without the need for colour correction later. The colour photography conducted by the author in the stereoscopic experiments reported here used Fujichrome 50 transparency film, which had combinations of high resolution, contrast reproduction and colour fidelity that were found to be satisfactory in side-by-side comparisons between the original and reproduced colours.

1.3 Perception and photographic practice

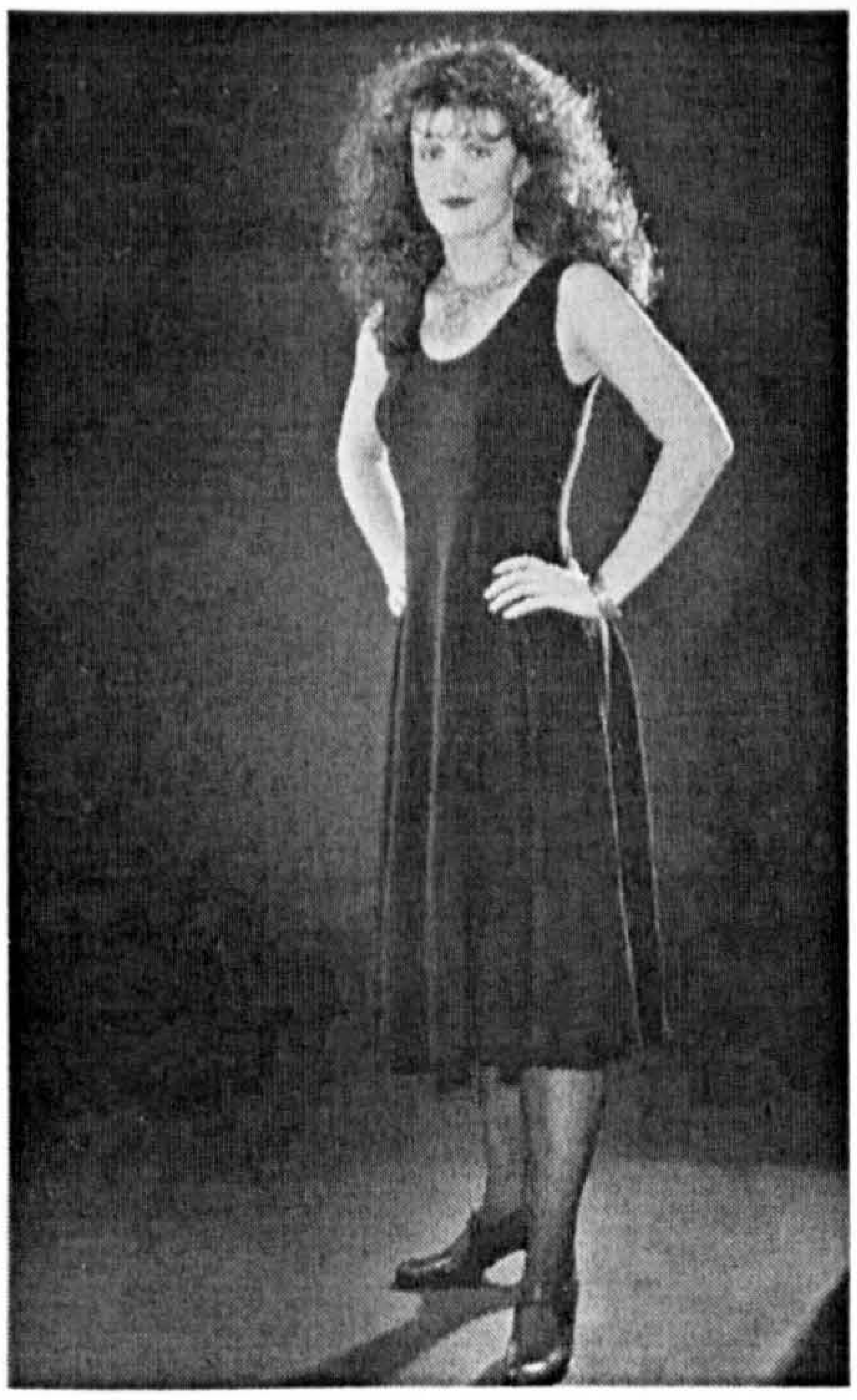
For working photographers and photographic technical authors, the lack of basic perception research in their field has led them to finding their own solutions to a range of imaging problems. In the autumn of 1994, while working as a staff photographer and photography lecturer in the University of Liverpool, the author was asked to undertake an unusual photographic assignment. Professor David Brodie of the Department of Movement Science required a series of photographs of a female student of his who would be dressed and made-up to portray three different roles or characters. The images were for an experiment on Body Image Distortion (B.I.D) to compare the responses of "normal" and anorexic women to role models, where the persons portrayed had the same body in each image. At that time he and Professor Peter Slade had completed a review of anorexia nervosa, bulimia nervosa and B.I.D. papers and had concluded that the primary diagnostic characteristic of anorexia nervosa Body Image Distortion (BID) was related to poor conceptualisations of *body concept* rather than failure of *body percept* due to flaws of vision or perception .

As no one had found any perceptual mechanism that could cause body image distortion, the conclusion was that it must due to some kind of learning, personality or mental disorder (Slade & Brodie 1994). His proposed experiment hinged on the belief that if body concept was disordered in anorexics, portraying the same body in different roles would lead to weight estimations that were different from the control group. The suggested roles were a student, a cook and a model. It was pointed out that while the first two roles were easy to portray and photograph, finding someone who could convincingly pass for a model was highly unlikely amongst the 100 female students in his department. It was argued that models are unusual people because they need to look unusually young, healthy and yet extremely slim to counteract the fattening and distorting effects of photography. He found this to be a highly implausible belief as he had never heard of it and insisted that the photography be completed as planned . The body image experiment photographs were taken, but never used (images reproduced overleaf). While the "model" was perfectly convincing as a student and as a cook, it quickly became apparent to that neither she, nor any of her contemporaries could ever pass for a real model. Professor Brodie encouraged the author to devise experiments to explore the idea that photography could indeed be an innately fattening medium. This process is described in detail in the reports of the first pilot studies (in appendix section H) that were conducted prior to the experiments reported in chapters 5,6 &7.

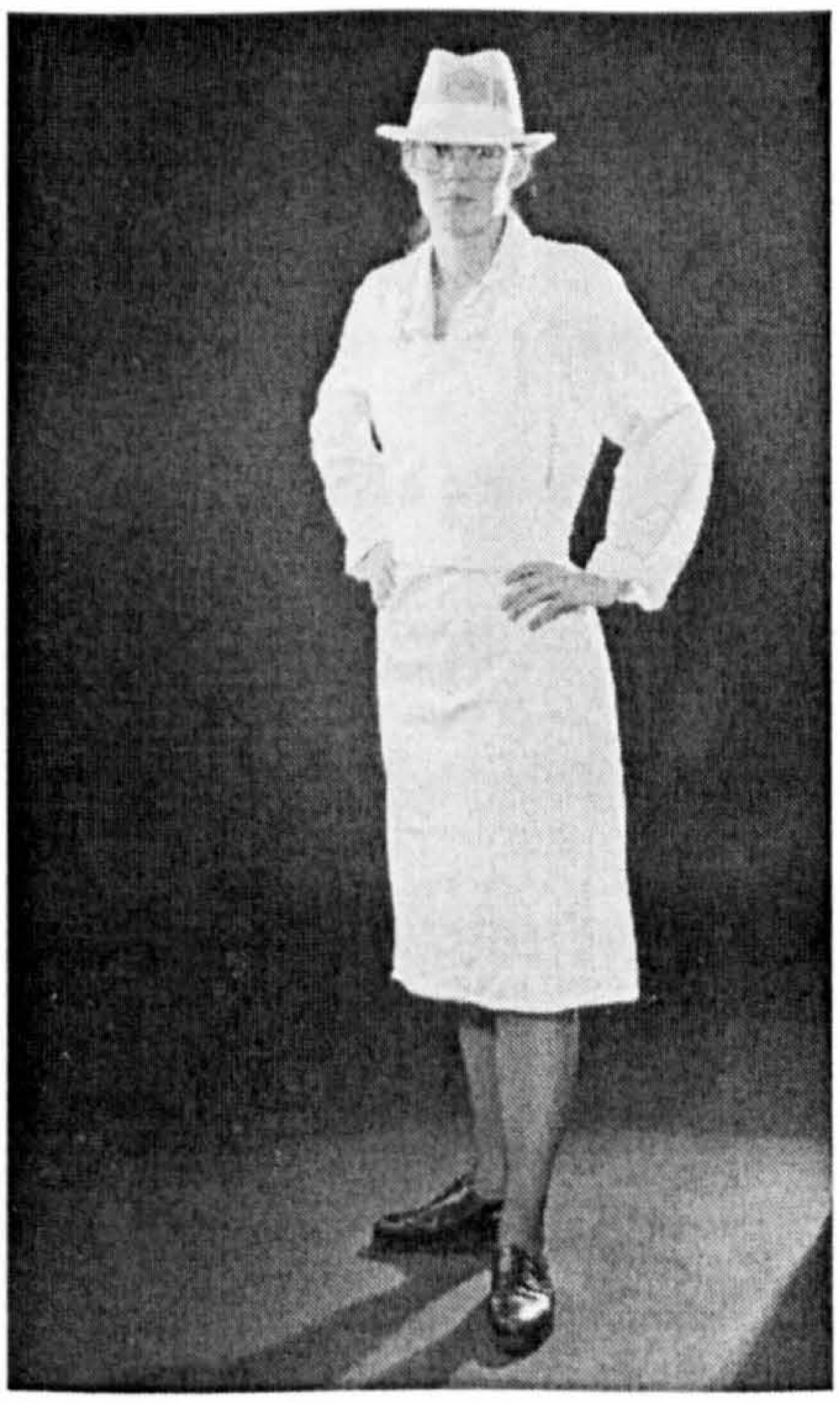
During the same period, the author was teaching photography to adult education students. His theories on why photographers have an obligation to photograph people with as little distortion as possible were seen by some students as controversial. The main theory presented was that the *flattening and fattening effect* produced when conveying the three-dimensional form of a person on to the two-dimensional medium of a photograph could be very *unflattering and unrepresentative.* The students questioned the validity of this approach because none of these theories were in any of their reference books. To counter this point, the author said "My photography tutor told us on the first day of our studies that he had some good news and some bad news for us. The bad news was that we had to buy two very expensive reference books. The good news was that nothing new had been discovered in photography for the last fifty years, so it was unlikely that we would ever need to replace them." What he meant was that almost everything that was relevant to photographers working lives had already been discovered, so the books he had used as a student would still be adequate for the task of teaching professional photography to us. This seemed to be an oversimplification and photography must have some secrets that it had yet to reveal. The experiments reported here are clear challenges to the notion that everything important in the science of photography was discovered long ago.



"Sue. Student aged 21"



"Sarah. Model, aged 23"



"Sylvia. Cook aged 22"

Figure 1.1 Images for a proposed study with anorexia patients to find out if a portrayed role changes estimated body size of the same model.

1.4 Body Image and the Media.

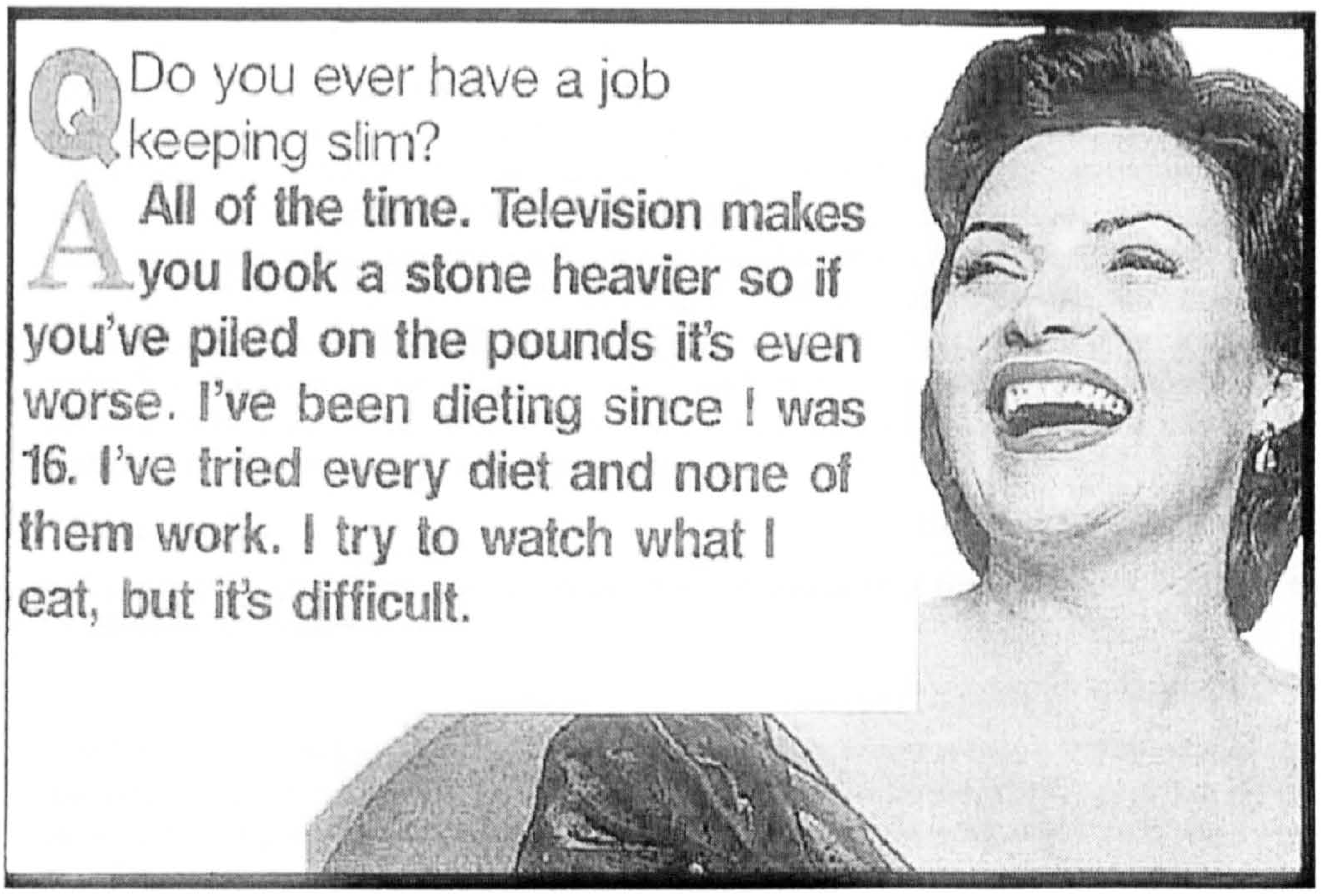


Figure 1.2 Quotation from TV presenter Loraine Kelly suggesting that television cameras can make one look a "stone"(14 pounds) heavier on TV than in reality.

The quotation above is taken from the Sunday Mirror Personal Magazine (1998, September 20 p3) and was made by TV presenter Loraine Kelly. She repeats the often made, yet anecdotal claim that television can make one look "a stone heavier" (14 pounds, 6.3 kilograms) on TV than in reality. A variation on the statement is below. But here it relates to still photography and proposes a fattening effect that is lower than in television. The quotation is taken from The Guardian newspapers "Notes and Queries" Stage and Screen section. This is a weekly question and answer feature that is written and answered by the readers. The question was asked in the newspaper on the 24 February 2001. However, the replies below were only ever reproduced on the guardian website (7) and include a contribution by this author.

<u>STAGE AND SCREEN</u>

Why is it that, even with today's complex technology, the camera is still said to 'put 10lbs on you'? And will this problem ever be solved?

L Gunby, Sheffield

- The problem is not the camera. The problem is the fact that you regard putting on ten pounds as a problem. We don't need a better camera, we just need a less weight obsessed society.

Colin Cuthbert, Hounslow UK

- We need a better camera! Photographers since the 1850s have used the expression "The camera never lies" to cover up the fact that it is an innately distorting medium. Only the most naive of photographers believe that it is a truly accurate method of visual record. The scenes that photographers perceive with their direct vision are always different to the final 2D photograph. The art of photography is to understand the distortions of depth, colour, contrast, brightness and magnification in every photograph and use them to make involving pictures. Experienced portrait photographers know that the camera can be "flattening and fattening" and make adjustments to the lighting, pose and angle of view to compensate. Or, they hire very skinny models and let nature make the compensation for them! It is only when we have developed a camera that can accurately simulate human vision that our friends will stop telling us to "put that damn thing away" and trust that anyone can take pleasing photos for the family album.

  Bernard Harper, Liverpool U.K.

- If it's true then shouldn't someone get the female actors from Friends, Frasier and Ally McBeal to hospital?

  Victoria Grant, London UK

- The fatness is created, as another answer stated, when a three dimensional object is reproduced on a flat plane eg. photo. I do work in 3d photography and if a 3d photo is taken correctly a person appears to be normal size because you loose the constraints of a flat medium.

  John Penna, Epsom UK

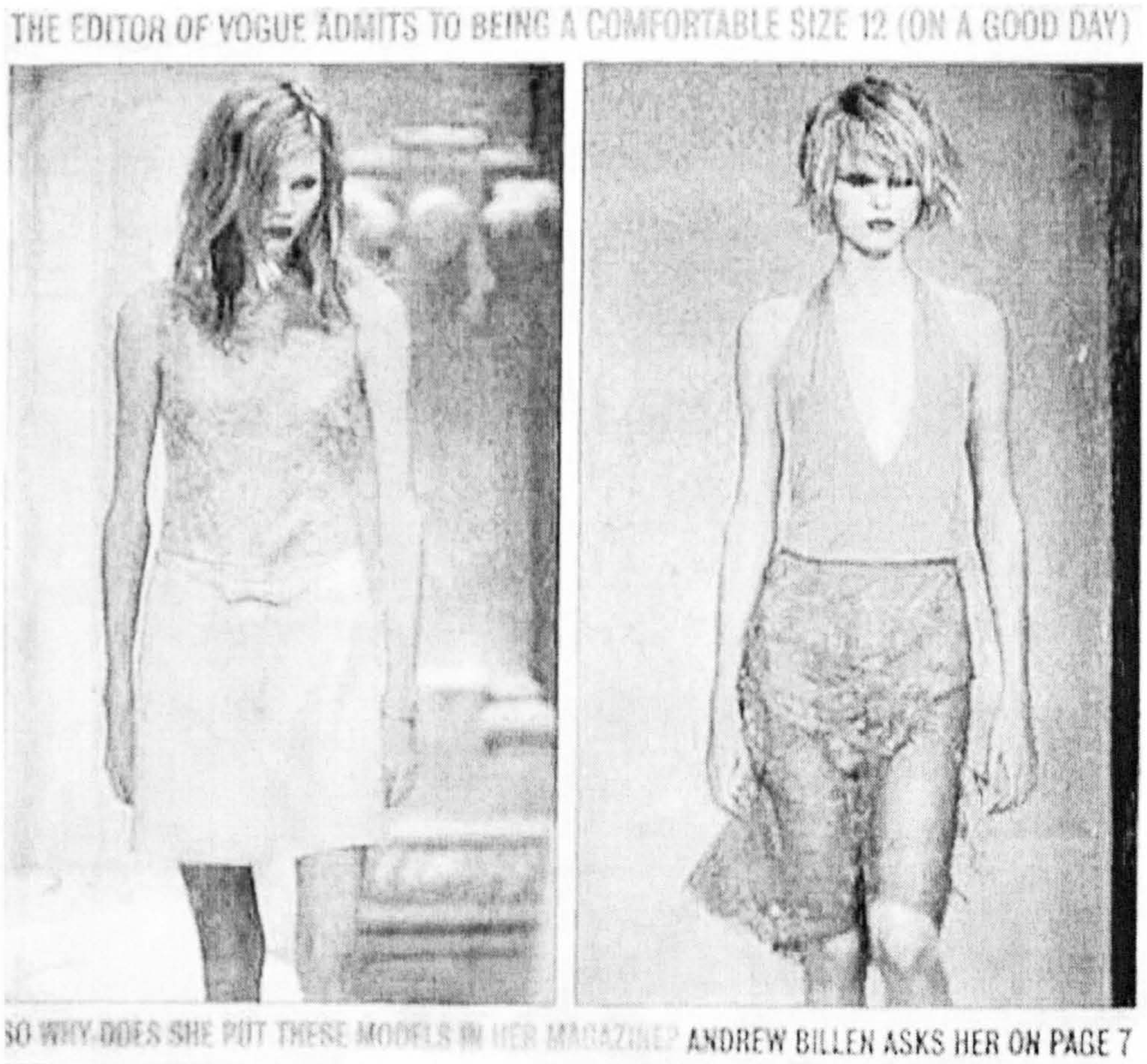


Figure 1.3 Cover photograph from the Observer newspaper (September 8 1999)

Figure 1.3 shows a typical example of "long lens" photography used in the reporting of fashion shows. Press and fashion photographers need to use powerful telephoto lenses to give close-up images of catwalk models from special platforms that are often 10 to 20 metres away. A typical lens used would be a 300mm of a fixed focal length, though lenses of greater than 500mm are required to give "head and shoulders" shots from these distances. Such lenses allow a camera to be six times further away from the subject while giving the same magnification as a standard lens. These images are very important to the fashion industry as they are effectively free advertising that is disseminated throughout the publishing world. Every effort is made to ensure that the viewpoints, backgrounds and lighting are optimised to give the best possible photographs of the products. Models too are chosen specifically to look good under these conditions. Any model who gave an appearance of accentuated body fat and did not flatter the designs or fabrics is unlikely to be chosen again for catwalk modelling. Historically, this has driven model choice to become slimmer than average and often having a body mass index well below 21 BMI considered to be the ideal (12).


TITANIC STAR IS SO IN LOVE ... AND SO HAPPY WITH HER NATURAL CURVES

Kate bonds on Bondi with her dream man

PICTURE EXCLUSIVE

Figure 1.4 Kate Winslet "paparazzi" honeymoon photographs (December 1998 Sunday Mirror).

Figure 1.4 shows photographs taken by unnamed "paparazzi" photographers and were published in the Sunday Mirror newspaper. They feature actress "titanic" Kate Winslet on honeymoon with Jim Threapleton in November 1998. These images are typical of the type of "long lens" photography used by the Paparazzi that allow them to take candid photographs from a great distance without the subjects being aware that they are photographed. Lenses of 500-1000mm focal length can give camera to subject distances in excess of 50 to 100 metres and can effectively render the photographers as invisible to their subjects. Kate Winslet was referred to as "titanic" Kate by this newspaper because the text made it clear that she was "too fat" to be a leading actress and used her starring role in the film Titanic to allow them to insult the actress every time she was written about. Many newspapers and magazines had made the observation that Ms Winslet was too large and suggested that that her career would be enhanced if she lost weight. However, in a interview in GQ magazine (February 2003 UK edition), the actress claimed that throughout her career she had never been over her ideal bodyweight. As a young actress however she had been

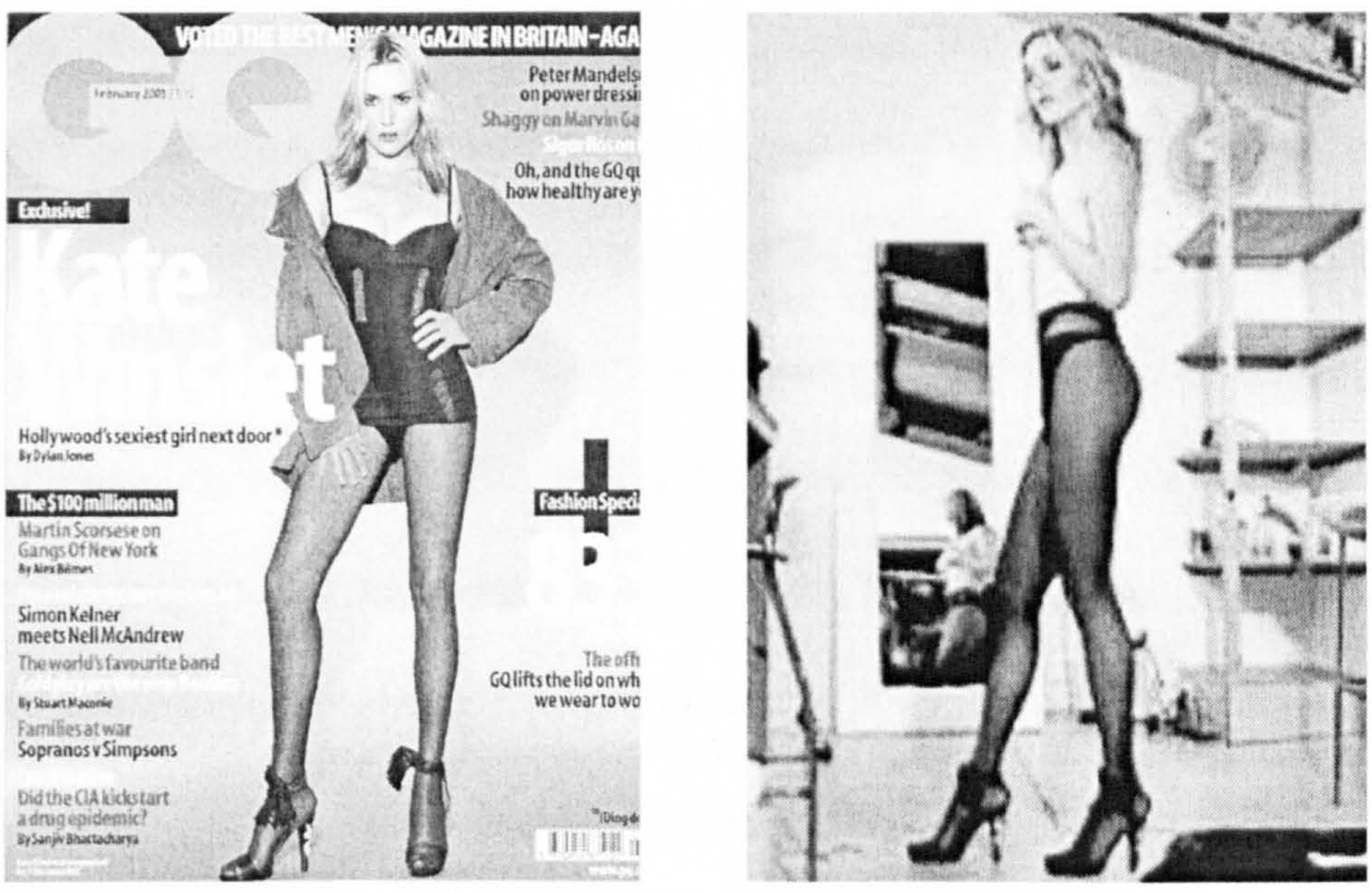
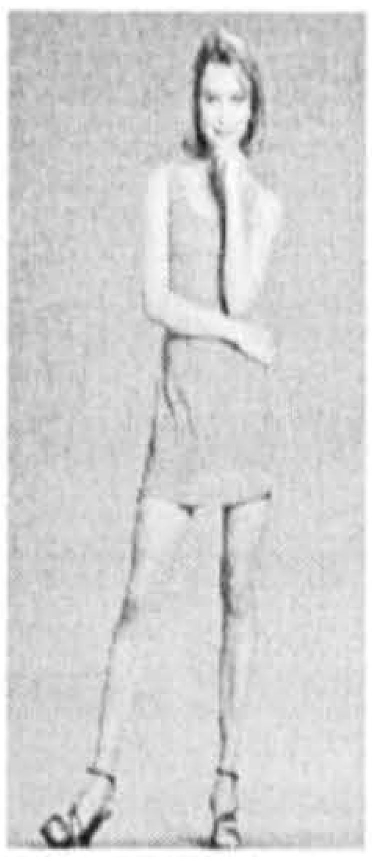


Figure 1.5 Retouched cover and feature photographs of Kate Winslet (GQ magazine).

advised to lose weight as it would help her to get more work. This led to an episode of anorexia when her BMI went below 18 and led her to becoming unwell. When she recovered she was determined to never again lose weight to get a part. The images reproduced from her GQ magazine interview (Figure 1.5) were later criticized because they were heavily retouched to show the actress as significantly slimmer (18) and the significance of this retouching is discussed in section 9.10. In a report called Eating Disorders, Body Image and the Media (12) published in the year 2000, the British Medical Association warned that actresses often diet to appear slim on screen and end up 20% thinner than average (compared to women in the general population of the same age) as they tend to overcompensate. The BMA went on to describe the media as a "negative reinforcer of overestimated body size in women with eating disorders".

So why is it that journalists and magazine editors feel free to make bodyweight assessments and negative comments about people they have only seen in photographs? Perhaps it is simply the old belief that "the camera never lies" is still at work: If people look overweight in photographs, *they must be overweight in reality.* Couple this belief to the other basic belief of journalists that "bad news sells" (11) and it is almost inevitable that defamatory photographs of the famous will be reproduced, accompanied by reporting of bodyweight that has no access to the true information about the subjects height, weight and ideal Body Mass Index and general health. Yet journalists themselves struggle with these issues and often write about the dangers of underweight models and actresses. Two stories from the Guardian newspaper "Thin, thinner, thinnest" and "Thin end of the wedge" (reproduced in full in Appendix D) are quoted from below:



Thin, thinner, thinnest

No, this isn't really how Calista Flockhart looks. But how long before it is? And how long, Anita Chaudhuri asks, before somebody says enough is enough?
The Guardian **Monday October 18, 1999**

"Cruel, yes, but celebrity weight-watching has become a spectator sport, fuelled by the ever shrinking waistlines of top female stars In America, this week's People magazine carries a cover story posing the earnest question "How thin is too thin?" alongside gruesome photographs of Jennifer Aniston and Courtney Cox, Victoria Beckham and Helen Hunt, as well as Flockhart. "I swear I eat more now than I ever did in my life," Aniston says, not entirely convincingly. Liz Hurley was also recently lambasted in the press for her malnourished appearance. In this month's Elle, she talks openly about the self-imposed pressure for women in the public eye to lose weight. "If it's any consolation, I threw away two-thirds of my wardrobe and lost 15 pounds after first seeing paparazzi pictures of myself - the celebrity version of a vicious Polaroid."

The article concludes in an interview with Deanne Jade, director of the National Centre for Eating Disorders.

"Half the problem with the girls I see is that not only do they have these skinny women in the public eye, but they have mothers who are on diets too. It's a powerful combination," Jade says. "For teenage girls, dieting becomes an endorsement of being female; talking about having to keep their weight down makes them feel like 'proper' women."

It is hard to see how any of these women, particularly Flockhart, can possibly be appealing role models, with their grey skin, lank hair and sad expressions. Not so long ago, Ally McBeal featured on a Time magazine cover alongside Gloria Steinem and Susan B Anthony, America's First Suffragette, under the headline "Is Feminism Dead?" Flockhart was not best pleased. "I was quite depressed. Ally McBeal is a woman who falls down, who throws her shoes. I'm afraid that's all I'll be known for." As her weight continues to shrink, we can count on her being remembered for something else entirely

The quotations below come from reporting on a "summit meeting" organised by the British Medical Association and the UK government that led to the publication of the discussion document "Eating Disorders, Body Image and the Media."

Thin end of the wedge

Hadley Freeman. Wednesday June 21, 2000

"Today, the government's 'super-waif summit' will debate the causes of eating disorders in girls. Hadley Freeman, who was hospitalised for four years with anorexia as a teenager, reveals that while images of skinny celebrities did not cause her illness, they did delay her recovery. In a way, Alexandra Shulman is absolutely right: you do not catch anorexia from Vogue magazine. At today's summit meeting, she and the other editors of fashion magazines can sit down, safe in the knowledge that the pages of their publications do not carry some kind of bacteria that will make their readers anorexic. Anorexia and bulimia are not trendy, zeitgeist illnesses. There are records of anorexics from as far back as the 16th century, so there is no point in trying to relate the cause of the illness to current pop-cultural issues."

"And yet. When models are becoming increasingly bony, when the ideal clothing size is diminishing every year, with actresses proudly showing off their hip bones and clavicles at this year's Academy Awards, when actresses and models who still need to wear a training bra are called "curvy", there is something very unhealthy going on. To watch the current series of Ally McBeal on television is a particularly fascinating experience, with actress Calista Flockhart's tendons becoming more prominent each passing week. Yet she is being celebrated as an icon of feminine beauty and independence in at least two recent magazines. People in the fashion and entertainment industries are deliberately and belligerently missing the point when they storm that they don't "cause" anorexia."

"Unquestionably, models and actresses are now expected to be thinner than ever before, and we are becoming inured to the situation. When these women become dangerously, bone-jutting skinny, they garner more publicity, more adulation, more success. This then becomes a vicious circle, with such images of skinniness being seen as the image of a Successful Woman. Calista Flockhart, Portia de Rossi, Lara Flynn Boyle, Sarah Michelle Gellar, Elizabeth Hurley, Courtney Cox and Jennifer Aniston were all, at most, moderately well-known actresses before their ribcages began to show. Now, particularly in the case of Flockhart and Aniston, they are to be seen marching down red carpets, clad in Versace dresses, beaming out from magazine covers, feted in high-fashion magazines which had previously ignored them. One US fashion magazine is said to have splashed out on a Marc Jacobs cashmere jumper to disguise what even it recognised as excessive skinniness during a cover shoot with Flockhart."

The following chapters and appendix H explain the problems of reproducing images of people in 2D media and why actresses and models pursue bodyweights that are much lower than the levels required for good health and fertility. Their aim is to add significant new knowledge to the basic theories that all photographers, researchers and scientists need to use in the execution of their work. The final chapter will offer solutions on revised photographic practices so images of people will soon be more accurate and less distorting than they have ever been before, and give journalists the information they need to interpret photographs correctly.

CHAPTER TWO

Reproducing Depth and Naturalness

"Not only **can** the camera lie, the camera **always** lies"

Malcolm Muggeridge, Christ and the Media, 1977.

2.1 Introduction

The difficulty of reproducing the depth and naturalness of real scenes on a 2-dimensional surface has long been appreciated. Leonardo da Vinci (1452-1519) in his book *Trattato della Pittura* wrote about binocular vision:

"A painting, though conducted with the greatest art and finished to the last perfection, both with regard to its contours, its lights, its shadows and its colours, can never show a relievo equal to that of natural objects, unless these be viewed at a distance and with a single eye."

(Howard and Rogers 2002, Volume I p.52)

It is obvious to us now that the only way to allow a more natural "relievo" of a real object is to reproduce it using stereoscopic imaging. Less obviously however in the same text, he also wrote: "The further that a spherical body is from the eye, the more of it you will see." This could be equally be interpreted as "the nearer you are to a convex object, the less of its surface you will see." It seems that Leonardo was well aware of the variability of an image caused by changes in *observer -to-subject* distance. Though in his case, he probably had no ability to magnify a distant image with lenses (8) and see in detail the effects he described.

Photographers use the term *camera-to-subject distance* in equations describing changes in imaging conditions. By using telephoto lenses similar in design to the earliest telescopes, photographers can record images that clearly show these changes with the same image size on an artificial retina (film, CCDs etc). However the Human Visual System is not able magnify an image in this way. Neither, it seems can we usefully determine if an image has been made by a telephoto or a wide angle lens (see chapter 6 ,section 6.1.1). In this thesis, the term *camera-to-subject distance* is used to describe the key variable in the perspective shown in any image. Photographers use telephoto lenses most often to make the subject appear to be closer to the camera. They seem to be so attracted to the advantages of the magnified image that they are unaware of possible deficiencies or unwanted side effects in the use of these lenses. However, photographers may chose to use telephoto lenses because they are believed by some to improve the rendering and size constancy in an image. The psychologist James E. Cutting of Cornell University contributed a chapter in a collection of articles called *Moving image theory: Ecological considerations* on conventions in cinematography, from which the following quotation is taken:

Perceiving Scenes in Film and in the World

"In photography relative size and density are manipulated through the use of lenses Swedlund, 1981). Perhaps the most familiar example of issues concerning relative size occurs in portrait photography. Here the photographer typically stands back from the subject and uses a long lens. For 35-mm film, the standard lens has a focal length of 50 mm, and a lens with a focal length greater than about 100 mm is considered a long lens, also called a called telephoto lens. With a short focal length lens on the camera, the camera must be placed close to the person being photographed, with the result that the difference between camera-to-nose distance and camera-to-ear distance is great, and the person's nose appears large. With a long lens on the camera, the camera can be placed farther away from the person

being photographed. With the camera farther away the difference between camera-to-nose and camera-to-ear distances becomes negligible, so the person's features appear close to their actual sizes. This is also one reason why most shot-reverse-shot sequences in cinematic dialogs are taken with relatively long lenses. They make the actors look better."

J. D. Anderson & B. F. Anderson (Eds.), Moving image theory: Ecological considerations. Carbondale, IL: Southern Illinois University Press.

Cutting appears to argue that the seemingly compressed perspective found with telephoto lenses is more accurate because it reduces the changes in apparent size caused by natural perspective. Conversely, "short focal length lenses" (normally called wide-angle lenses) are dismissed because they increase the differences in apparent size. Yet he does not consider the case of the "standard lens" in photography (see appendix H, section 5) which neither compresses or expands perspective and reproduces the appearance of natural size constancy when its images are viewed at life-size magnification. A more likely explanation for the use of telephoto lenses in film production is for the purpose of noise suppression in "talking movies" and is discussed further in section 9.12The significance of this convention is examined in the camera-to-subject experiments reported in chapter six.

2.2 The Fattening Effect of Photography.

It is commonly said in the fields of photography, film and television that the camera is so fattening "it can put 10lbs on you." Distortions are regularly mentioned anecdotally (Gunby, 2000; Kelly, 1998 and appendix D). Until recently however, one could find no academic or technical discussion of this effect. This is despite researching the phenomena with a number of institutions such as the British Journal of Photography, the Independent Television Commission, the Moving Image Society (BKSTS), the Royal Television Society, members of the American Society of Cinematographers and more conventional scientific resources. To date, only one reference has been found in any journal or by any technical author on the fattening effect of photography. It was discovered in a technical manual called Three Dimensional Photography: The Principles of Stereography by Herbert C. McKay published in 1953. In chapter 10, THE NUDE IN STEREO on page 163 appears the following paragraph:

"Then too there is the oft-repeated statement that women in photographs appear shorter and stouter than in real life. Frankly I do not see it that way, but perhaps those who are not trained to analytic vision are so deceived. At any rate I have had others speak repeatedly of the fact that stereo does not have this effect. I have had several women say about a stereogram, "Oh! I like that. It doesn't make me look as fat as I usually do in a photograph!"

Until the present study, it appears that the effect has been known by technicians but ignored by scientific research and no one has examined the fattening effects of photography in a systematic way.

2.2.1 Cyclopean Vision

Professional photographers are sometimes aware that the scenes viewed with their normal direct vision will differ from the 2D representations produced when they are imaged and transferred to photographic paper or a projection screen. Almost everything about the original scene is conveyed in a modified or degraded form. The descriptions of classical image aberrations (e.g. Langford, 1989, chap. 2) only cover the effects of simple uncorrected lenses on the shape or colour of the image. The most obvious loss in conventional imaging is presence derived from stereo depth information (Freeman, Avons, Meddis, & Pearson, 2000; Freeman, Avons, Pearson, & IJsselsteijn, 1999; Hendrix, & Barfield, 1996; IJsselsteijn, de Ridder, Hamberg, Bouwhuis, & Freeman, 1998). However, there are other more subtle effects of which we are often unaware: Peripheral vision objects and scaling cues are usually excluded from photographic images.

Photography almost always fails to reproduce scenes at their original (same-size) magnification. Even when this is achieved, it cannot reproduce the detail that can be seen with normal vision from the original viewpoint while maintaining the angle of view. Natural brightness ranges are immensely difficult to reproduce as each image generation can add contrast and lose shadow or highlight detail. Accurate colour reproduction too is almost impossible with conventional imaging and subject colour failure can be found in most types of image capture. These "fidelity failures" are often corrected by trial and error or custom and practice techniques derived from professional knowledge (Langford, 1989, chap.8).

The only thing that appears to be unchanged in a photograph is the point of view. However, the single-point perspective that makes a photographic image appear to be an accurate representation of the original scene can also convey inaccurate object information. Humans too perceive the world from a single-point perspective. By the process of cyclopean vision (Julesz, 1971 and figs. 2.1-2.4), we see the world through a "cyclopean eye" that generates a single artificial viewpoint from a location mid-way between each real eye. In normal human vision, the processes of foveal convergence, accommodation and stereo fusion allow the brain to construct a new

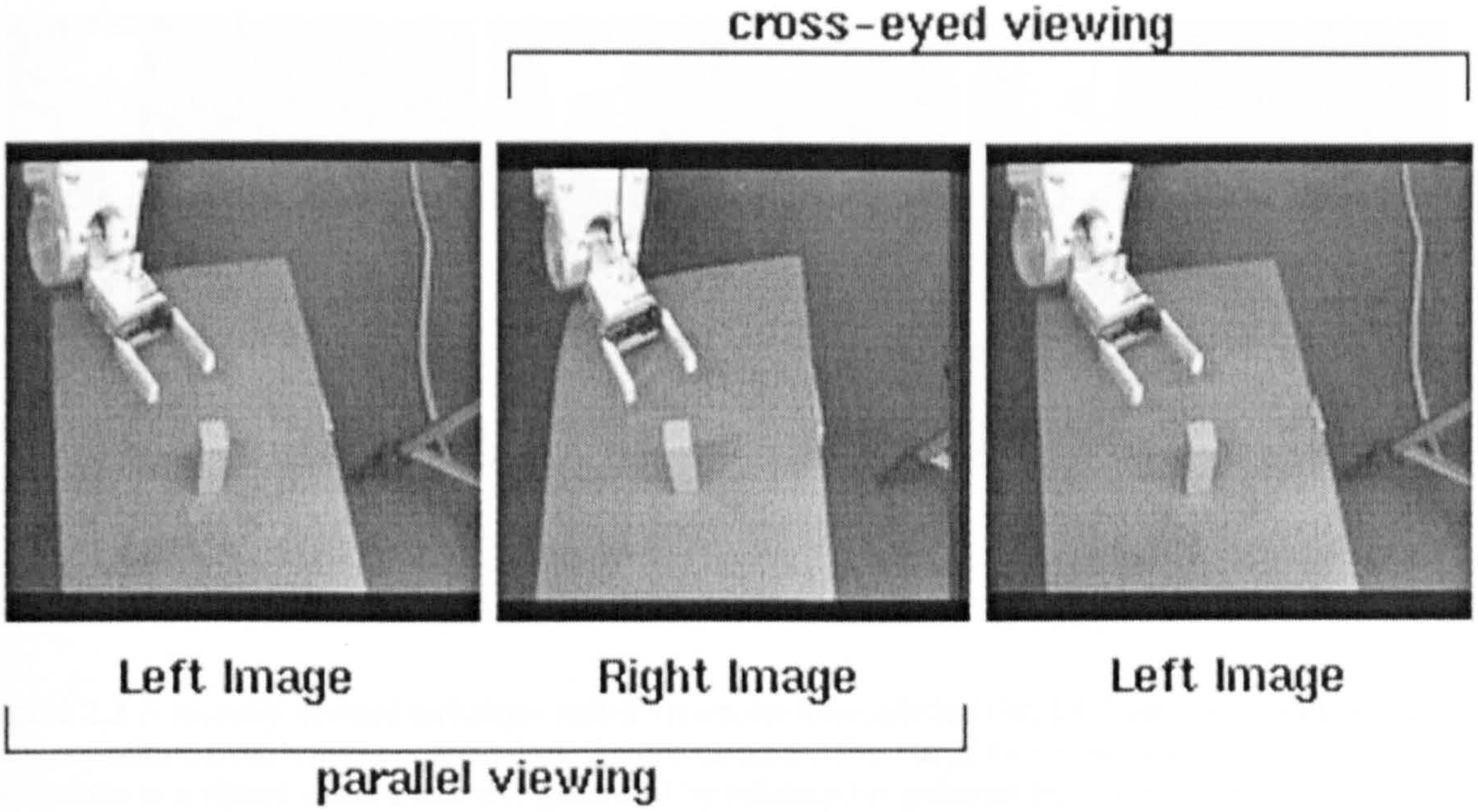


Figure 2.1 An example of a cyclopean view can be seen in these stereograms by Rolf Henkel (13). If one views the left and right image with a parallel stereo print viewer, or the central and left image by cross-eyed "free viewing," one can see a remarkable effect. The wooden peg appears to point directly towards a central viewpoint midway between each eye.

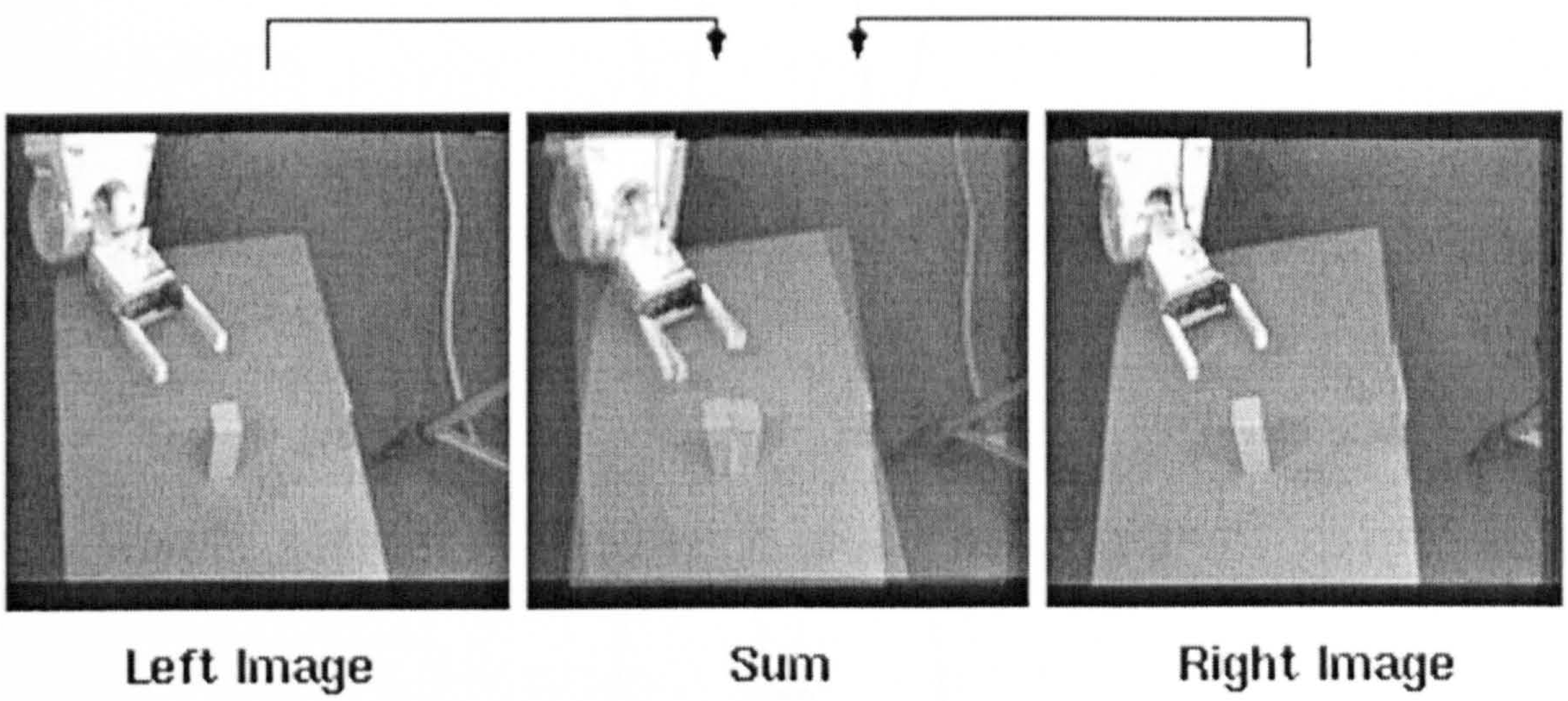


Figure 2.2 The central "sum" image shows there is no correspondence between the two images when they are placed in layers. To see the wooden peg without blurring, the HVS interpolates a new view-point from mid-way between each eye that combines information from both eyes into a single percept.

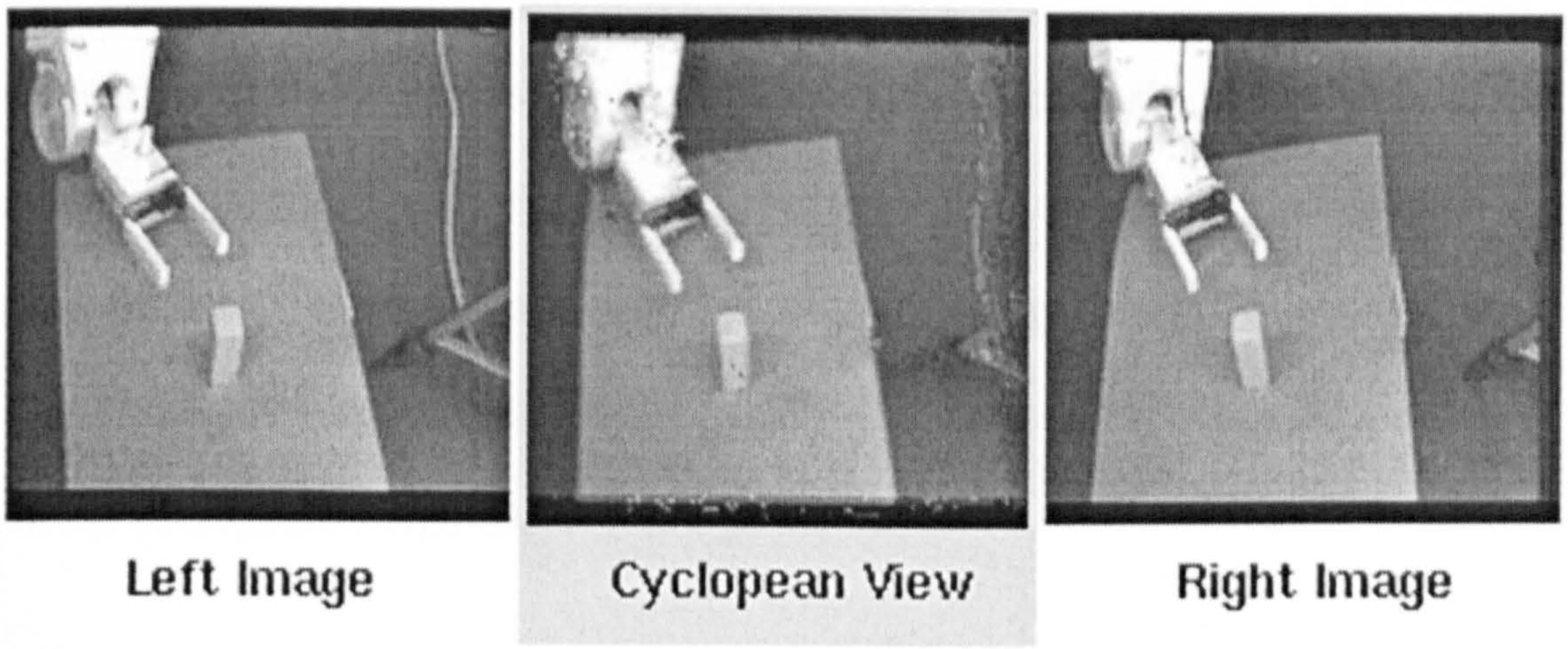


Figure 2.3 A recently devised technique called viewpoint interpolation (14) has been use to simulate a camera position from a mid-point between the two cameras. The visual direction of the wooden peg now points to a virtual camera that was generated by information gathered from both of the stereoscopic cameras and simulates the effect of a cyclopean eye.

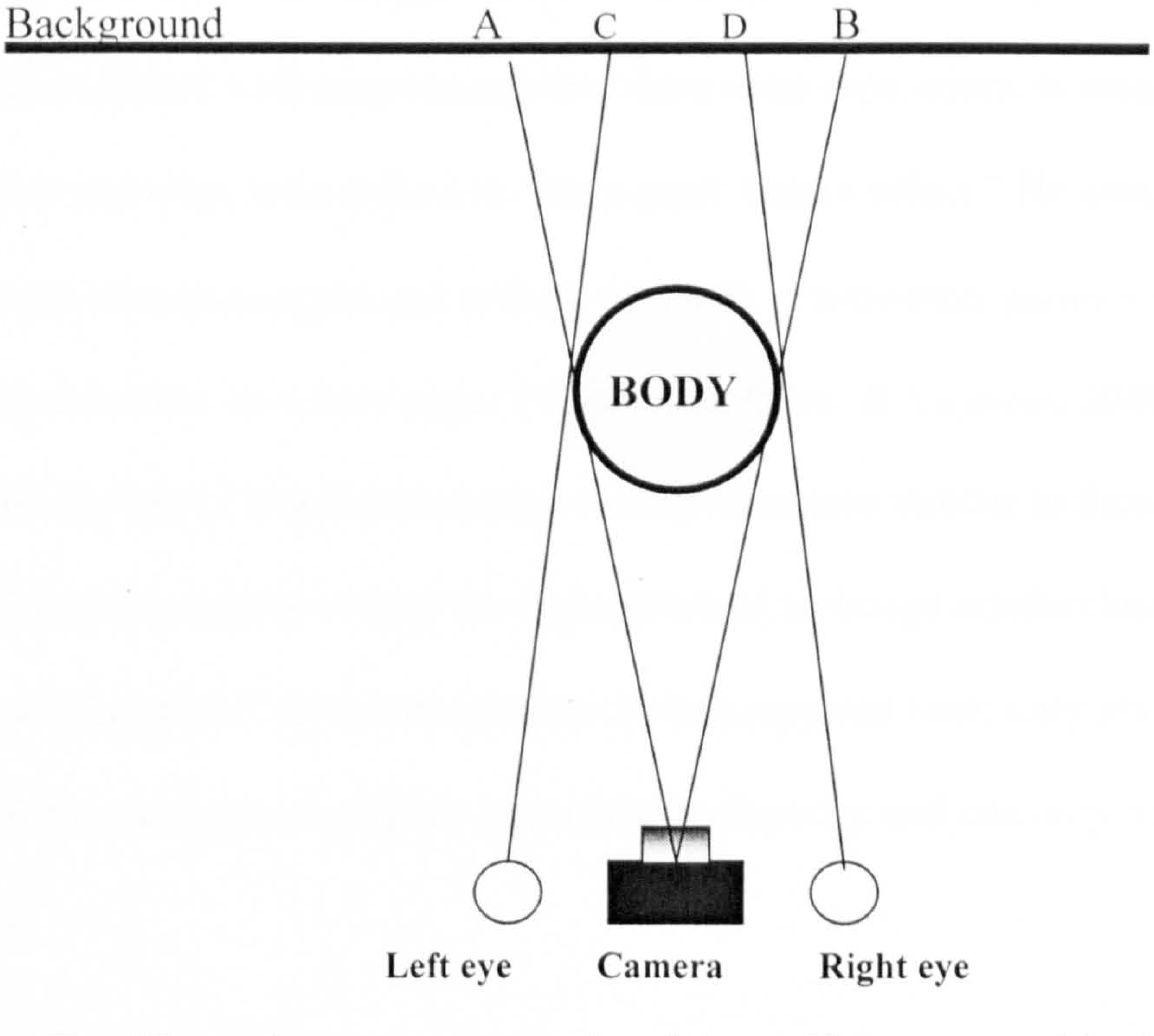


Figure 2. 4 The difference between a camera point of view and human stereo vision from the same position. A body occludes more of the background in a 2-D photographic perspective (A-B) than with stereo vision (c-d).

perspective that differs from those seen by either eye individually. This cyclopean point of view appears to be similar to a 2D photographic perspective. However, a single lens system cannot reproduce the way in which we can focus/fuse on an object with two eyes and see both diverging and converging optical paths from the same position. With close-up objects we have the ability to see the normal photographic perspective and also have "look around" vision from a single head position. The result is that close-up objects viewed stereoscopically occlude less of the background than their 2D photographic equivalents (see figure 2.4).

This thesis investigates the possibility that failure to reproduce this geometry in a display is a major cause of the fattening effects associated with conventional photographic images. A previous study using large HDTV cameras (Yamanoue, 1997) found evidence of changes in size estimations in stereoscopic conditions. His experiments linked widening camera lens inter-axial separations to smaller size perception and what is described as "the puppet theatre effect." He used direct observation of a mannequin and compared it with a same-size, parallel imaged stereo video reproduction. In a later paper (Yamanoue, Okui, & Yuyama, 2000), they supported the use of lens separations and magnifications similar to those of the human visual system in order to reduce the appearance of an image artefact known as the "cardboarding effect." In the stereo experiments reported here, only photographic images were viewed and only the stereoscopic disparity and convergences were changed.

2.2.2 The Orthostereoscopic condition

In psychophysical experiments, monocular vision has consistently been linked to lower performance when compared to binocular vision, with the exception of the horizontal-vertical illusion (Prinzmetal, & Gettleman, 1993). Tasks such as luminance increment detection, contrast sensitivity with sine wave gratings, colour discrimination, vernier acuity, letter identification and visual search (Banton & Levi, 1991; Blake, Sloane & Fox,1981, 1981; Jones & Lee, 1981) all show improved performance in the binocular condition. It is possible that whenever the visual system is presented with images that do not allow it to form a normal binocular cyclopean view, predictable perceptual disturbances will occur: the display medium may be flawed in its ability to convey objects and people in their original proportions, size and background occlusion characteristics.

The International Stereoscopic Union glossary (15) describes an orthostereoscopic image as "An image which appears to be correctly spaced as in the original view. See also, tautomorphic image. A tautomorphic image is a stereoscopic image which presents the original scene to the viewer exactly as it would have been perceived in life; i.e. with the same apparent scale, positions of scenic elements, and a stereo magnification of x1 for all subject matter in the view." By this, it means that it is an image that is free of artefacts caused by lens interaxial separations that are greater or smaller than the average adult human interpupilary distance of 65mm. It also describes Giantism as the jargon term for hypostereo: the impression of enlarged size of objects in a stereo image due to the use of a stereo base separation less than normal for the focal length of the taking lens(es). Dwarfism (or hyperstereo) is the opposite effect. Sometimes called "Lilliputism" it refers to the use of a longer than normal stereo base in order to achieve the effect of enhanced stereo depth and reduced scale of a scene. The name Lilliputism was coined because of the miniaturisation of the subject matter which appears as a result of using this technique and is often referred to as the "puppet theatre" effect.

The earliest reference to the technical advantages of orthostereoscopic imaging can be found in the October 1952 edition of the Journal of the Society of Motion Picture and Television Engineering. Written by Raymond Spottiswoode, N.L. Spottiswoode and Charles Smith (Spottiswoode et al 1952), they describe the technique as "the condition of perfect image reproduction - i.e. that in which the image as a whole is geometrically congruent with the scene it represents." Surprisingly, while recognising the theoretical advantages of orthostereoscopic imaging Spottiswoode et al argued that this would constrain the artistic freedom of directors and cinematographers into using fixed focal short length lenses and lens interaxials of only 65mm. Their pragmatic solution was to reject these constraints for more flexible and practical combinations of magnification, lens inter-axial separations and alignments. This often meant that images were captured using long telephoto lenses, wider than normal lens inter-axials, "narrower than natural" convergences and that the stereo window of reproduction was often placed behind the plane of focus/screen plane. Lenny Lipton in his book *"The Foundations of the Stereoscopic Cinema"* (1982) uses the term toti-orthoscopy to describe orthostereography in this way:

"Whatever its practical merit (which is limited), the toti-orthoscopic condition comes close to creating an isomorphic reproduction of the visual field within the stereophotographic field than other kinds of orthoscopy. It is as if we were looking through a rectangular aperture cut into a black material of endless extent. This aperture at the time of photography is the same size and shape as the screen on which the film will be projected. But only the spectator at the V =Mf, distance, which is equal to the distance from the camera to the object, will be able to enjoy the full, although dubious, benefits of toti-orthoscopy." Lipton (1982), p. 225

He seems to agree with Spottiswoode et al (1952) that the benefits of orthostereoscopic images are restricted to a single viewer seeing a life-size image from the same screen-to-viewer distance as the original camera-to-subject distance. However, for stereoscopic experiments where the goal is to produce as lifelike an image as is possible to a single observer only, an orthostereoscopic image capture and display system is the best method to reproduce natural viewing geometries and provide a more lifelike visual experience.

2.2.3 Waist-Hip Ratio

On viewing the first computer generated stereoscopic images made for these experiments (figure 5, appendix H), it was observed that the thickness of the models waist appeared to be slightly smaller in the stereoscopic images compared to the 2D monocular projections. The waist of this model was 50% of the size of the shoulder and hip areas and it bore little resemblance to the normal human form. Waist-to-Hip Ratio or Waist-Hip Ratio (WHR) is the term normally used to describe the ratio of the circumference of the waist and the circumference of the hips in humans. WHR is most commonly used to refer to proportions by which fat is distributed around the torso. The concept and significance of WHR was first theorized by Evolutionary Psychologist, Dr. Devendra Singh at the University of Texas at Austin in 1993, "Adaptive significance of female physical attractiveness: Role of waist-to-hip ratio." Journal of Personality and Social Psychology, 65. According to Singh, a WHR of 0.7 for women has been shown to strongly correlate with general health and fertility. Women within the 0.7 range are reported to have optimal levels of estrogen and be less susceptible to major diseases such as diabetes, cardiovascular disorders and risk for ovarian cancers.

WHR is considered to be a factor in judging female and male attractiveness. Women with a 0.7 WHR are rated as more attractive by men, regardless of their culture, ethnicity or any other subjective factors. Such diverse beauty "icons" as Marilyn Monroe, Twiggy, Sophia Loren, Kate Moss and the Venus de Milo all have ratios around 0.7 despite having significantly different weights. Consequently, there is strong evidence to suggest that from an evolutionary perspective humans use subtle biological cues, such as WHR, to indicate mate potential and fertility.

More recent studies using photographic stimuli, rather than the relatively crude line drawings used by Singh, have suggested that the primary cue of male attractiveness is the waist-chest ratio (Maisey, Vale, Corellison, & Tovée, 1999; Swami & Tovée, 2005b) which is the same as the WHR in the vertically symmetrical stimuli used here. Studies using photographs have also shown that relative body weight is arguably the dominant cue in females (Swami & Tovée, 2005; Tovée, Hancock, Mahmoodi, Singleton, & Cornellison, 2002). However, body weight and WHR are clearly closely linked for a particular height.

The 0.5 WHR of the peanut shape used in the first computer generated (CGI) images was therefore judged to have weak *ecological validity* if the results were to be generalised to judgments of the human form. So all subsequent CGI stereoscopic images were made with a WHR of 0.7. However, in order to disguise an overtly female interpretation of the shape, the rotational symmetry of the 0.5 shape was retained, as was its vertical proportions with the waist area centered between the hips and shoulders.

2.2.4 Ecological Validity

The Psychology Students Glossary (17) describes Ecological Validity as "the degree to which the behaviours observed and recorded in a study reflect the behaviours that actually occur in natural settings. In addition, ecological validity is associated with "generalizability". This is the extent to which findings (from a study) can be generalized (or extended) to the "real world". In virtually all studies there is a trade-off between experimental control and ecological validity. The more control psychologists exert in a study, typically the less ecological validity and thus, the less they may be able to generalize the data to real world phenomena. An important principle in the experimental designs used in this thesis was to ensure that the stereoscopic stimuli used in the experiments were as lifelike as possible. The virtual images of the people and CGI stimuli presented were to be so lifelike that they could act as substitutes for a real person or object: At a later date, it should it be possible to run versions of the experiments with real targets and use similar paradigms to discover if size estimations varied when real targets were viewed monocularly or binocularly.

2.2.5 Reproducing Stereoscopic Disparities.

The technical constraints of stereoscopic photography are such that few attempts are usually made to fully reproduce the normal geometry and vergences of the human visual system. The term "Parallel Stereography" is used here to refer to all stereo image capture geometries that do not converge the lens axes on the centre of focus & interest at the object plane and generate a double image at the plane of reproduction. Almost all stereography uses combinations of lens inter-axial separations, magnifications and convergences different from those the human visual system would use when viewing the original scene. For instance, the average human interocular distance is approximately 65 mm but stereo camera separations are often much wider than this. Also, they usually fail to reproduce the point of zero disparity from the original scene with zero disparity in the display. This means that they show a single point from the captured scene as two points on the screen and the viewers are required to "force fuse" these points to form a single stereo image.

The term "zero disparity" is used in perception research to describe the fixation point in a binocular field where the foveas of each eye align on the surface of an object or a point in space (Howard & Rogers 2002 Vol.1). Zero disparity can be considered as the point of monocularity within a stereoscopic field because there is no perceptual difference between inputs to the left and right channels The horizontal disparities in front of zero disparity (the fixation point) are described as "uncrossed" because unfixated points imaged by the left eye are found to the left of centre (cyclopean) view, while unfixated points to the right are imaged found to the right. However, the disparities behind the point of fixation are termed reverse disparities because the lines of fixation cross behind the point of zero disparity. So reversed disparities can also be described as crossed disparities.

Richards (1971) proposed that there are three types of binocular detecting neurons found in the visual cortex that are similar to those found in monkeys. In his model for stereopsis, the first type of detectors were termed divergent or uncrossed disparity detectors and were stimulated by binocular positions in front of the fixation point. The "on" distribution were stimulated by positions lying on (and to both sides of) the fixation point and were termed "no disparity" detectors. Unfixated points behind the no disparity detectors stimulated the third type, called "crossed disparity" detectors. Fisher and Poggio (1979) also found evidence of these three types of cells in striate and prestriate coretx of the monkey: the tuned neurons were sensitive to small non-zero disparities and subserved fine stereopsis, and the near and far neurons activated by larger disparities (crossed and uncrossed respectively) and operating at coarse depth discrimination. Maunsell and Van Essen (1983) reported that disparity detecting neurons were sensitive to vertical as well as horizontal disparities. Experiments by Duke and Howard (2004) examined the effects of vertical-disparity gradients on apparent depth curvature of textured surfaces. Their results suggested that vertical disparities alone can be used to determine distances to surfaces directly, rather than to simply estimate vergence as had previously been thought. So it seems that vergence mechanisms of the human visual system use disparity detecting neurons to allow stable fixation on an object in near space, while linked accommodation processes ensure that the fixated point is focused.

It is interesting to speculate of the role of disparity detecting neurons in the perception of stereoscopic displays. Conventional stereography tends to avoid the restrictions of orthostereoscopic imaging. Yet is only the orthostereoscopic condition that can reproduce the horizontal, vertical, zero and reversed disparities that the HVS uses in normal perception to fixate on the centre of attention and allow the depth information to be naturally decoded. It is likely that normal stereoscopic perception is highly dependent on these disparities being reproduced accurately and that any failure to reproduce these disparities could cause misperception in an artificial three-dimensional scene. Yet these disparities are rarely reproduced in conventional stereography, and reversed and zero disparity are actively avoided by most systems. The experiments reported in this thesis use orthostereoscopic imaging to minimise possible photographic distortions so that the stereo image capture geometry that is as close as possible to that of the human visual system.

It was considered that viewing comfort should also have a high priority in the presentations. There are limits (Panum's fusional area) to how far out of horizontal or vertical alignment binocular stimuli can be before there is loss of fusion and diplopia or suppression of one image (Howard & Rogers, 2002). To reproduce how the visual system achieves this, the point of focus for each camera should coincide with the convergence point for each lens axis, and this must be reproduced as a point of zero disparity in the display. The displayed image should be scaled to life-size (x1 magnification), so that the object being imaged has the same visual angle on the retina and is at the same distance when projected as it gave to the orthostereoscopic camera when the image was captured. This alignment was most likely to give comfortable viewing because when the points of each camera's focus are horizontally aligned in the display, the centre of interest (a face, for instance) appears as a single image. Zero separation in the display (no double image at the centre of interest) means that relatively flat objects can be viewed successfully without polarizing spectacles. Typically, this condition has a high degree of 2D compatibility as only the out of focus areas are not aligned at the screen.

Polarising spectacles allow the viewer to separate these areas into discrete channels by which they can then perceive the original scene depth. The principle that underlies all of the stereo experiments reported here is that orthostereoscopic images are presented to the participants for comparison with 2D images from the same viewpoint and camera to subject distance. In practice, this means that when participants are making size or shape judgements under experimental conditions, they are presented with images where the only differences are of disparity.

The pilot studies using the techniques described in chapter 2 are detailed in appendix H. They are included in this thesis as a guide to researchers wishing to reproduce the experiments reported in chapter 5. The following chapter, entitled "Body Image Distortion and the misperception of photographic images" reviews the previous research and papers that have been influential in the methodology reported in chapter 4 and the design of the experiments reported in chapters 5 and 6.

CHAPTER THREE

Body Image Distortion and the misperception of photographic images.

3.1 Introduction

There are three primary hypotheses that underlie this research: firstly, a photographic image can be distorting because it lacks the volumetric (z-axis) depth information that we normally perceive with direct stereoscopic vision. Secondly, the magnification of a photographic image rarely matches the viewing distance of the eye when viewing the original scene, so it is not reproduced with life-like size, distance or magnification information. Thus, the original distance to the object is not conveyed and can undermine size constancy derived from a perceptual process combining perceived distance and visual angle. A third effect is synoptic enlargement, whereby photographs that appear to have been imaged at a great distance also appear to be larger than a near object that subtends the same visual angle on the retina.

These factors seem to combine to form the characteristic flattening and fattening effect often observed in human portraiture and figure studies. This chapter reviews the previous research that informed the development of methods used in the initial pilot studies (reported in appendix H) and the more recent research in the fields of body image distortion and misperception of mediated imagery.

3.2. Body Image

The term Body Image refers to a person's appreciation or perception of their physical appearance. The concept was first formulated by the German writer Paul Schilder in his book, The Image and Appearance of the Human Body which was translated into English in 1935. He described Body Image as "the picture of our own body which we form in our mind, that is to say the way in which our body appears to ourselves." Later authors in the field have expanded the definition to refer to "the picture we have in our minds of the size, shape and form of our bodies: and to our feelings concerning these characteristics and our constituent body parts" (Slade 1988).

3.2.1. Body Image Distortion

The earliest observations of the phenomenon of body image distortion (B.I.D.) in people with eating disorders were made by Hilde Bruch in the early 1960's (e.g. Bruch, 1962). She had observed that many emaciated anorectic patients were adamant that their bodies were actually fat and came to the opinion that this was the primary diagnostic criteria for the illness. She argued that these distortions of body image were so great that they were "disturbances of delusional proportions," that were fixed and resistant to any attempt at external or clinical intervention. This resistance described by Bruch was also reported in many subsequent studies reviewed by Cash and Brown, (1987), Thompson

(1990) and Slade and Brodie (1994). The latter article was a major review of all of the studies which had measured body size estimation accuracy and it came to the following conclusions:

- Individuals suffering from anorexia or bulimia tend to overestimate their physical body size and especially, bodily widths.
- Randomly selected female controls and psychiatric groups also overestimate their body size, though to a lesser extent than eating disorder patients.
- A tendency to overestimate body size is therefore not unique to anorexia or bulimia patients, so cannot be used solely as diagnostic criteria.

Slade and Brodie went on to describe how these studies required a complete "reconceptualization" in a paper published in the Eating Disorders Review, also in 1994. They identified five areas where the recent literature indicated that eating disorder patients displayed a diagnostic difference in "body concept" but no difference between the controls in "body percept." These were summarized below as:

1.Perceptual Accuracy vs Attitudinal Bias

A study by Gardner and Moncrieff (1988) used a signal detection paradigm to separate "perceptual and attitudinal" bias. They found no difference between nine anorectic subjects and nine controls in terms of sensory sensitivity, but a significant difference in terms of bias. They concluded that body image judgments differed between groups primarily due to attitudinal variables.

2. Conitive vs Affective instructions.

A series of studies have examined verbal instructions on image accuracy judgments. Proctor and Morley (1986), Franzen et al (1988), Bowden et al (1989), Robinson et al (1996) all performed experiments to compare body image estimates by eating disorder patients and controls under two main conditions: Cognitive (make judgments on how you *think* you look) and Affective (make judgment on how you *feel* you look). All of these studies showed a consistent pattern of results: the eating disorder patients exhibited significantly greater overestimation of body size in the affective condition compared to the cognitive instructions. There was no difference between the controls in both conditions. So the "instruction effect" was confined to eating disorder subjects and remained consistent in patients re-tested at a later date.

3. Variability in Accuracy Judgements.

Collins et al (1987) analyzed their subjects' responses on the basis of accuracy of three matched groups of observers. The matching criteria were: Accurate, i.e groups of subjects whose average error was within 14% of the actual size; Underestimators, defined as people whose size estimations were 15% or more under actual size;. Overestimators, whose average size estimates were 15% over the actual size. The data was also grouped by normals (controls *n*=60), whose accuracy fell within 80% percent of the accurate range and characterized as accurate estimators.

The second group were bulimics (*n*=24), of whom 50% were accurate estimators and 42% were overestimators. Bulimics therefore were either accurate or overestimators of their own body size. However, the two other groups: obese (*n*=150) and anorectic (*n*=78) had high percentages in each of the three size estimation groups. Slade and Brodie (1994) suggest that obese and anorectic subjects may have histories of repeated weight change which may lead to a "loosening of their body image" and thereby rendering them more likely to be less accurate in their estimations of body size.

4. Effects of mirror confrontation on body size estimation.

A study by Norris (1984) reported the size estimation accuracy of four groups of subjects before and after a period of mirror gazing. Neither the normal control group nor the emotional control group were affected by the manipulation; while both the bulimic and anorectic groups showed a reduction in size estimation following the period of mirror confrontation. The manipulation produced a change in responses only in the eating disorder subjects.

5. Evidence that anorectic's exhibit fluctuating size judgments.

Brindred et al (1990) used a video based experimental paradigm with seven anorectic subjects over a four week period to examine reported fluctuating BID. They concluded "Although no consistent change in BID over time was evident for the group, all subjects showed significant changes, albeit in different directions. This change in individuals, obscured in the analysis of group data, suggests a lack of stability of BID over time."

3.2.2 Measurement of BID

Many methods have been developed for the measurement of body image, reviewed by Cash and Brown (1987), Thompson (1990) and Skrzypek et al (2001) and generally they fall into four groups as follows: Firstly, there are the Body Size/Body Site Estimation procedures. These require subjects to make judgments about single body sizes, normally in the horizontal plane, for example about the width of their face, waist and hips. Such methods include the following:

Visual size estimation

The Visual Size Estimation (VSE) or Movable Caliper Technique (MCT) was used in early studies comparing patients with anorexia nervosa with normal subjects (Slade and Russell, 1973). This was based on the technique developed by Reitman and Cleveland (1964) and requires the individual to estimate the width of each body dimension by reference to the distance between two movable lights.

Image marking procedure

The Image Marking Procedure (IMP) was developed by Askevold (1975) and requires the individual to mark their perceived sizes on a sheet of paper attached to a wall; these are then compared with actual sizes marked by the experimenter/therapist.

Adjustable light beam apparatus

The Adjustable Light Beam Apparatus (ALBA) developed by Thompson and Spana (1988) requires subjects to adjust the width of a light beam projected on a wall to match the perceived size of various body dimensions.

Body image detection device

The Body Image Detection Device (BIRD) developed by Ruff and Barrios (1986) requires subjects to adjust the distance between two lights to match their perceived size.

A second set of methods are the Distorting Image or Whole Body Adjustment procedures. These require subjects to correct a distorted image of their body as a whole. Such methods include the following:

Distorting photographic technique

The Distorting Photographic Technique (DPT) requires subjects to adjust a photographic image of themselves in line with their perceived body shape. This method was pioneered by Garner and Garfinkel in Toronto (Garner et al., 1976; Garfinkel et al., 1978; Garner et al., 1983) following on from the work of Glucksman and Hirsch (1969) with obese subjects.

Distorting mirror technique

The Distorting Mirror Technique (DMT) requires subjects to adjust an image of themselves in line with their perceived body size/shape. This method was pioneered by Traub and Orbach (1964) and has been used in several studies with eating disorder subjects since then (Huon and Brown, 1986; Brodie and Slade, 1988).

Distorting video technique

The final method in this category involves the Distorting Video Technique (DVT) which requires subjects to adjust a distorted video image in line with their perceived body size/shape. This method was first described by Alleback et al. (1976) and has since been used in a number of studies with eating disorder subjects (Freeman et al., 1984; Collins, 1986; Brodie et al., 1989).

A third set of methods involve the use of drawings or silhouettes which range from very thin to very fat outlines of the human form. They are normally used to determine the difference between perceived and ideal body size/shape. This method was first described by Furnham and Alibhai (1983) and has since been used in a host of studies.

Finally, a fourth set of methods involve questionnaires which are primarily used to measure Body Dissatisfaction/Disparagement. The original scale was that of Secord and Jourard (1953) which has become the prototype for most of the scales subsequently developed.

All of the BID assessment methods used until 1995 could be collectively described as miniaturized 2D representations or perspectives. Probst et al (1995) were the first researchers in the field to use life-sized projected images and utilized the video distortion technique. The revised technique was also used by Gardner and Bokenkamp (1996) and was reviewed by Skrzypek, et al

(2001). They concluded that the life-sized video distortion produced results in accordance with previous studies using smaller images.

3.2.3 BID and Perceptual deficits.

Skrzypek, et al (2001) in their review of body image assessment using body size estimation in recent studies on anorexia nervosa have mirrored the views of Brodie and Slade that BID is a symptom of body concept, rather than body percept. They wrote: "The concept of body image is thought to consist of two components: body size perception and attitudes towards the body. Correspondingly, two distinct modalities of body image dysfunction seem to be relevant for anorexia nervosa: perceptual body size distortion and cognitive-evaluative dissatisfaction.

Consequently, various techniques to assess body image have been developed to assess body image disturbance in eating disorders, particularly anorexia nervosa." They compared the numerous methods of assessing body image used in recent studies on anorexia nervosa and discussed the various findings. The findings suggest that body image disturbance is not due to any perceptual deficit, but is based on "cognitive-evaluative dissatisfaction." Although overestimation of body size is not a universal symptom in eating disorders, this issue was considered important in terms of prognostic and therapeutic implications. Thus, body size estimation remains a worthwhile approach to assessing body image disturbance in eating disorders. The identification of sensory and non-sensory factors in body image was considered to be a promising issue for further research."

Their conclusion however that "sensory factors" are considered as a promising area of research is surprising, as their review found little evidence that revised BID measurement techniques gave new insights into its perceptual aspects. They suggest that as "Cognitions and affective components are widely recognized as part of the perceptual process, it would therefore be helpful to take into account both "sensory and non-sensory" factors into account when defining body image." However, in a major review of eating disorder studies by Williamson et al (2005), the view that sensory factors and body perception were important elements of eating disorders was rejected. Their conclusions concurred with previous studies that BID was primarily a disturbance of body concept, rather than body percept: "These findings provide support for cognitive behavioral models of eating disorders that environmental and emotional events, especially those that activate weight concerns, are associated with increasing body size overestimation." However, none of these studies considered the possibility that viewing two-dimensional photographs is an essentially unnatural process and that this may in itself lead to perceptual distortions which might contribute to the BID observed in anorectics. The experiments described in this thesis investigate for the first time the various possible sources of perceptual distortions when making size judgments with photographically presented images. As well as developing our knowledge of the perceptual processes involved, they may therefore also throw some light on the origins of BID in anorectics.

3.2.4 BID and the Media

In its report, Eating Disorders, Body Image and the Media (2000) **(12)** The British Medical Association considered the role that the media can play in shaping people's attitudes to eating and body shape, and developing self-esteem in the young who are at the greatest risk of developing an eating disorder. The report considers whether the media play a significant role in the causation of eating disorders, where they can 'trigger' the illness in vulnerable individuals, by suggesting that being 'thin' means being successful, and how they affect young people who may have low self-esteem or unhealthy attitudes towards food. More positively, the media may be able to significantly contribute towards developing high self-esteem in young people, and actively participate in health promotion to combat the mistaken belief that 'thin equals healthy' and that 'dieting', rather than healthy eating and regular exercise, is the way to achieve a healthy weight.

The BMA report accused the media of being a "negative reinforcer of overestimated body size in women with eating disorders". It went on to conclude that "Female models are becoming thinner at a time when women are becoming heavier, and the gap between the ideal body shape and the reality is wider than ever. There is a need for more realistic body shapes to be shown on television and in fashion magazines." The BMA also reported that actresses were "on average, 20% underweight when compared to women of their own age in the general population." This is likely to be unhealthy, because staying below a Body Mass Index (BMI) of 18 for extended periods has been implicated in permanent infertility in women.

Extensive studies by by Professor Rose Frisch of Harvard University has confirmed that body fat and fertility is intimately linked in pre-menopausal women. Frisch (1987) reports "We can think of body fat as Sex Fat, because it provides the energy required for reproduction." What she found as most surprising was that there is a razor-thin borderline where losing just 3lb (1.4 kg) can tip a normal sized woman into infertility. "Many women who maintain the body shape made popular on the catwalks are completely infertile. These women can, however, continue having periods and not guess that anything is wrong until they try, and fail, to get pregnant. If low body-fat switches off a woman's reproductive system for too long, it may be too late to regain fertility simply by putting on weight. Women need to gain weigh at a younger age. A 5ft 3in (average height) girl must weigh at least 90lb (41 Kg) if she is to become fertile, while a pre-menopausal woman of the same height needs to over 101lb (45Kg) if she is to continue to ovulate." The implication of this research is that it may not be too long before a women sues their employer because they insisted she maintained a bodyweight that improved her appearance, but seriously undermined her fertility.

3.2.5 Perceptual restrictions of 2D Media

Size constancy refers to the tendency to perceive objects as retaining their size, despite the increase or decrease in the size of the image projected on the retina caused by moving closer to or farther from the objects. (fig.3.1)

Photography, and indeed all images with perspective use this basic perceptual effect to represent "real world" depth in a medium constrained to only 2 dimensions of spatial reproduction. And as can be seen by the illustration above, the HVS can be easily deceived as to the true size of an object if the perspective is not life-sized and accurately reproduced. An illusion that can dramatically demonstrate the power of this effect was first demonstrated by the American ophthalmologist Adelbert Ames, Jr., who first constructed such a room in 1946 (Figure 3.2)

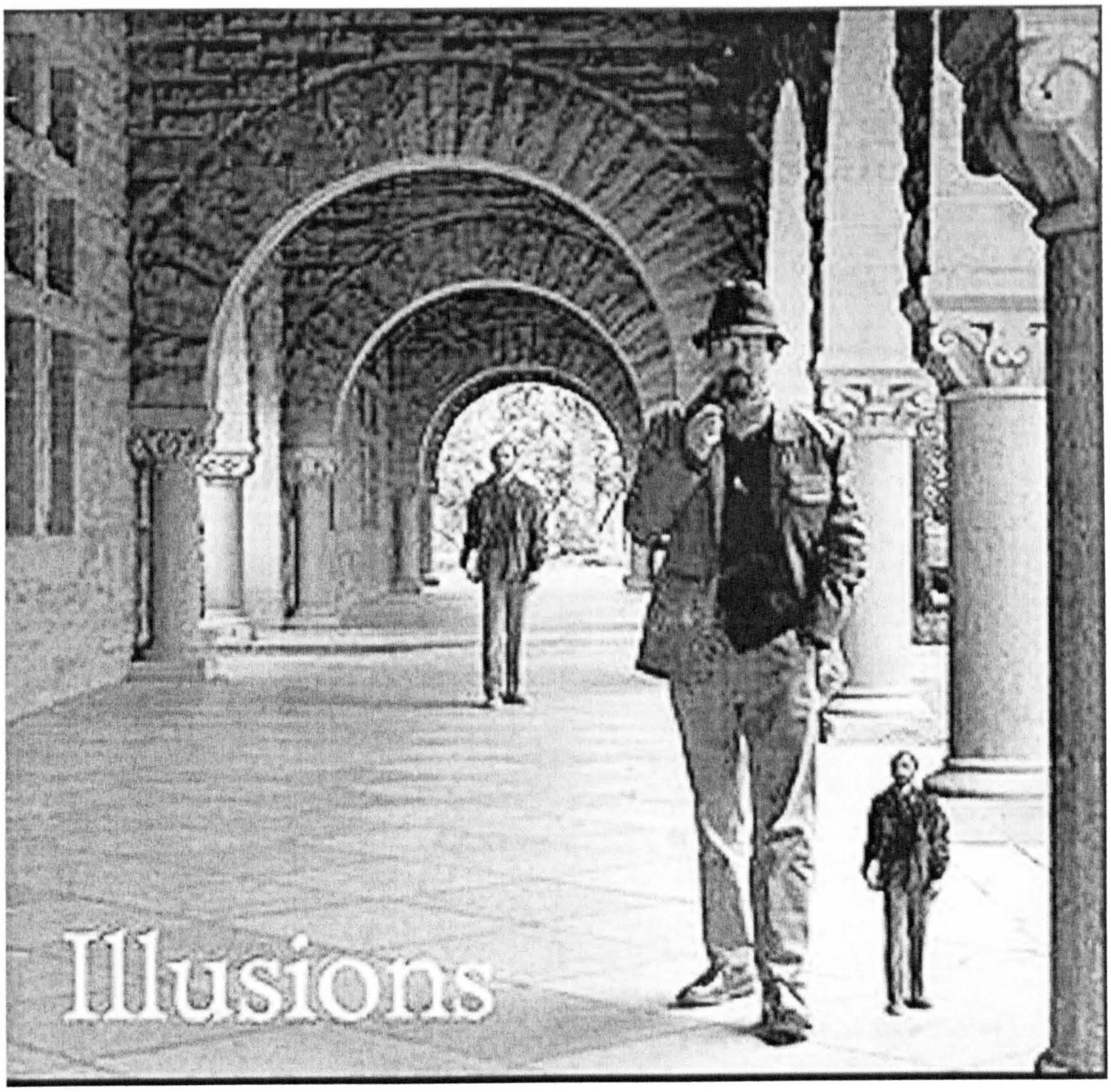


Figure 3.1 An illustration of Size Constancy (www.illusions.com) shows how the background figure appears to be "life-sized" in relation to a foreground figure, even though the "miniature " person on the right actually subtends the same visual angle on the retina as the person in the background.

Figure 3.2 An Ames room demonstration (photo supplied by www.psychologie.tu-dresden.de)

When you look through a peephole "to remove any cues from stereopsis" into an Ames Room (Ames, 1952), the room looks normal and cubic, but its true shape is distorted. The floor, ceiling, some walls, and the far windows are actually trapezoidal surfaces. Although the floor appears level, it is actually at an incline (the far left corner is much lower than the near right corner). The walls appear perpendicular to the floor, although they are actually slanted outwards. The generally accepted explanations states that given this ambiguity of perspective, your visual system relies partly on past experience with normal cubic rooms to judge the shape of the room. This explanation is an often repeated argument in favour of top down processing, i.e., that your visual system resolves ambiguity based upon knowledge of the external world.

Gregory, (1963) has written extensively about the Ames Room and has noted however that the Ames illusion is not a perfect illusion. He observed that there was a certain "oddness" to its illusion and felt that this oddness was perhaps due to irregularities in its construction. It is more likely however, that this oddness comes from trying to accommodate to the two far corners. The HVS may assume that the two corners are equally far away, but they do have two different focusing distances.

Accommodation (focusing) however, is not a strong enough cue to break the stronger constraints that give rise to a cubic perception. Seckel and Klarke (1990) fitted the peephole with a pinhole, which removed any cues from accommodation (normally thought to work over only small distances). Observing through a pinhole, any signs of oddness disappeared. An important limitation to the power of the illusion was noted by Ames in his original paper, which is that the illusion is entirely monocular in nature. When viewed binocularly, the illusion breaks down and the true shape of the room is revealed.

3.2.6 Perceptual advantages of Binocular Vision

In psychophysical experiments monocular vision has consistently been linked to lower performance when compared to binocular vision, with the exception of the horizontal-vertical illusion (Prinzmetal & Gettleman 1993). The horizontal-vertical illusion is the tendency for observers to overestimate the length of a vertical line relative to a horizontal line that has the same length. One explanation of this illusion is that the visual field is elongated in the horizontal direction, and that the vertical-horizontal illusion is a kind of framing effect (Künnapas, 1957a, 1957b, 1957c). Since the monocular visual field is less asymmetric than the combined visual field, this theory predicts that the illusion should be reduced with monocular presentation. This prediction was tested in five experiments in which the vertical-horizontal illusion was examined in a variety of situations. However, in the binocular condition, the horizontal component of the image was in fact seen as smaller than when viewed monocularly. This effect was not explained by the authors, but it is in accordance with the slimming effect along the horizontal axis in the stereoscopic experiments reported in chapter 5 of this thesis.

3.3 Apparent Size with Distance for Monocular and Binocular Views

Holway and Boring (1941) studied the relationship between the perceived size of targets and size constancy predictions in three monocular conditions and one binocular condition. They showed that the apparent size of an object did not vary greatly when the distance from the observer varies, provided the observation is made with uninterrupted binocular vision. They further showed that the function by which apparent size is dependent approximates the law of visual angle (Hering 1861), though their data showed that monocular observers (wearing an eye patch) were less accurate at size estimation. The monocular data was later confirmed by Taylor and Boring (1942) when the experiment was repeated with one-eyed observers only. Their results suggested that the misperception of size was a function of monocular perception alone and not caused by temporary impaired vision of normally binocular observers.

The relevance of both studies to this thesis is that they are the only ones found to date that use monocular and binocular views in an experiment measuring apparent size. These two factors are the predominant issues in the studies reported in chapter 5 of this thesis. Using Holway and Borings' results as guide, one would predict that binocular viewing would give more accurate size estimations than the monocular conditions. This prediction was tested in the experiments 1-4 in chapter 5.

3.3.1 Varying of Apparent Size in Stereoscopy with Varying Binocular Views

Yamanoue (1997) investigated the relationship between size distortions and shooting conditions in stereoscopic video productions. Of particular concern was the need to avoid the "puppet theater effect" whereby objects are reproduced as smaller than life size. This effect is also known as dwarfism and according to Yamanoue is made worse by the presence of background planes placed well-behind the target to be reproduced. The target, a life sized mannequin, was imaged using differing camera lens separations of 65mm, 95mm and 125 mm in front of a black background and standing in a hallway. Both backgrounds at 65mm gave normal to slightly large size estimations in the responses from 13 observers.

However, the image with a real background images with 65mm of lens separation produced the most consistently life-sized responses. At 95mm and 125mm of lens separations, the observers report smaller size estimations that correlated to interaxial lens separations: the wider the interaxial separation, the smaller the size percepts were reported, but in the realistic background condition only. The black background conditions gave the same slightly large size estimations in all three conditions. Yamanoue concludes the use of orthostereoscopic imaging with a lens interaxial separation of 65mm gave accurate size estimations of the target, but only with a realistic background image. The stereoscopic experiments reported in chapter 5 of this thesis use orthostereoscopic imaging with 60-65mm interaxial lens separations and realistic background planes behind each target.

3.3.2.Display Size and its Effects on Size Estimations

Bradshaw et al (1996) investigated the effect of display size on disparity scaling from differential perspective and vergence cues. They compared the relative effectiveness of differential perspective and vergence angle manipulations in scaling depth from horizontal disparities. When differential perspective and vergence angles were manipulated together (to simulate a range of different viewing distances from 28 cm to infinity), approximately 35% of the scaling required for complete depth constancy was obtained. When manipulated separately the relative influence of each cue depended crucially on the size of the visual display. Differential perspective was only effective when the display size was sufficiently large (i.e., greater than 20 deg) whereas the influence of vergence angle, although evident at each display size, was greatest in the smaller displays.

For each display size the independent effects of the two cues were approximately additive. Perceived size (and two-dimensional spacing of elements) were also affected by manipulations of differential perspective and vergence. These results confirm that both differential perspective and vergence are effective in scaling the perceived two-dimensional size of elements and the perceived depth from horizontal disparities. The display size in the experiments reported in this thesis were always of images that were life-size and subtended a visual angle of greater than 20 degrees to the observers.

3.4. Image Constancy in Pictures Viewed from Oblique Angles

Hagen, (1976) investigated the influence of picture surface and "station point" (now referred to as CoP, for centre of projection) on the ability to compensate for oblique views in pictorial perception. She also investigated the importance of awareness of the pictorial surface and point of observation, in 40 4-yr-olds, 40 7-yr-olds, and 40 university students. Information for the pictorial surface was varied in 2 ways: by use of 2 points of observation for pictures, 1 congruent with the correct center of projection of the pictures and 1 not; and by direction manipulation of the pictorial medium, photographic prints contrasted with transparencies. Observers were instructed to pick the larger member of a pair of photographed objects viewed monocularly through a peephole. She reported that there was a clear improvement with age in perception of pictorial size-in-depth relations.

There was also a strong effect of the visual angle relations of the pictured pairs. The effect of CoP interacted with pictorial surface and age, suggesting the development of a mechanism of compensation for the perspective distortion of oblique view. This existence of this mechanism was supported in subsequent studies by Rosinski and Faber (1980), Rogers and Bradshaw (1995) and Argrawala et al (2000), but refuted by Vishwanath et al (2005). They manipulated images by viewpoint and by the apparent orientation of the picture surface. They found that picture invariance was achieved by a local estimation of surface orientation, not from geometric information in the picture. The Hagen and Vishwanath studies both conclude that 2D pictures are perceived with shape constancy from a central viewpoint. This conclusion was supported by pilot study experiments conducted in advance of the stereoscopic experiments reported in chapter 5. By projecting the first swimsuit images (see appendix H) to a group of five observers (arranged in a diamond formation, 1 metre distance from an observer at the correct CoP), it was found that reduced size estimations in binocular conditions were independent of viewing position. All five viewers recorded size estimations that were consistent, regardless of their viewing position. This showed that even large changes in viewing position had no effect and confirmed that a stereoscopic image can be successfully viewed from areas adjacent to the ideal viewing position at the centre of projection. All the experiments reported in this thesis were, however, run with the observer at the correct CoP.

3.4.1. Photographs vs Reality

Evans (1976) made the important observation that "all photographs of any kind are always distorted relative to reality." In a seminar presentation by Marc Green (21), an accredited expert witness in court cases he observed: "I'm writing this section because of several calls that I've recent received from attorneys. In each case, the attorney had taken photographs at an accident scene and wanted me, as an expert, to confirm that the photographs could be presented to the court as a valid representation of what someone would have seen at the time of the accident. I've had to decline because they cannot." He went on explain further why photographs and animations cannot be used for this purpose. "The reasons fall into three categories, *physical, sensory and cognitive*."

Physical Reasons

Green makes many observations that are similar in nature to the photographic limitations of film and photo-mechanical reproduction described in section 1.2. He then reports "Camera optics produce other distortions. Lens choice can have a large effect on spatial relations. Short focal length (wide angle) lenses tend to cause 3D spatial relations to distort because the lens is typically closer to the subject. Conversely, long (telephoto) tend to compress objects together. The angle of the photograph also affects relative size of objects at different distances. Lastly, there are common perspective errors. To be viewed accurately, jurors would have to be placed at the correct eye position to view the images.

Sensory Reasons

"In theory, some film nonlinearities, and possibly range compression, could be corrected to accurately show the scene brightnesses and color accurately. This is an involved process and, in any event, could not compensate for several sensory factors (reduced dynamic range, for example), as well as the cognitive problems discussed below. Sensory problems occur because photographs are easily distorted vs. reality by visual factors at the time they are viewed: Photos also lose the visual cues if binocular disparity and texture."

Cognitive Reasons

"There are many psychological biases in interpreting images. Blurred and low contrast objects appear further away, objects on the left of an image often appear closer than objects on the right, etc. Perception involves knowledge and expectation. The court cannot put itself in the mind of the accident participant. Perception is a function of time. The court can casually study the photos. In an accident, the participant is usually caught off-guard and must respond in a split-second glance."

Summary

"Photographic images should not be used as graphic representations of reality for many reasons. There is, unfortunately, no real substitute for going to the accident scene, measuring luminances of object and background, comparing these values to visual contrast threshold norms and then taking into account visual and cognitive factors. Lastly, these comments apply only to the use of photographs as representations of what a person would have seen. Photographs may serve other valid purposes to aid the court in finding of fact. However, any visual image can create subtle cognitive effects."

Green's assertion that photography should not be used as indisputable representation of reality is in accordance with the main theoretical and methodological approaches used throughout this thesis. However, in the photographic experiments reported in chapter 5, orthostereoscopy is used to add life-like depth to the 3D photographs that are otherwise identical to the 2D comparison stimuli. This means that the imaging issues described above are common to both types of 2D and 3D stimuli, and are therefore balanced in each presentation: Only the disparities presented are different in each condition and as such, they simulate the viewing conditions most commonly found either in 2D photography or the direct binocular viewing of real scenes.

3.4.2. Photographs and Boundary Extension

The first researchers to identify the phenomenon of boundary extension were Intraub and Richardson (1989). They had observed that when subjects see a scene in a photograph or even a three-dimensional representation with a clear border, their memory of that scene tends to extend beyond the original boundary: you remember the scene as larger than it actually was, sometimes even just a few seconds after seeing it. They described the effect as a "picture-memory phenomenon in which subjects' recall and recognition of photographed scenes reveal a pronounced extension of the pictures' boundaries." After viewing 20 pictures for 15 s each, 37 observers exhibited this striking distortion; 95% of their drawings included information that had not been physically present but that would have been likely to have existed just outside the camera's field of view. To determine if boundary extension was limited to recall and drawing ability, a second experiment tested recognition memory for boundaries. Eighty-five undergraduates rated targets and distractors on a boundary-placement scale. Subjects rated target pictures as being closer up than before and frequently mistook extended-boundary distractors as targets. They speculated that the phenomenon may be due to the active creation of a mental representation of a scene. As most perceptual constructs are derived from scant available information (the area visible by the fovea as the eye scans parts of the scene), it makes sense that the HVS might also construct a representation of parts beyond the boundary of what is actually viewable. They concluded "If observers fill-in the blanks between places they have actively looked at, why not extend their representations beyond the boundaries of what they have actually seen?"

Intraub et al (1996) tested the possibility that "boundary extension" occurs soon after picture perception. In their first experiment, 151 undergraduates viewed 7 close-up or wide-angle photographs for 250 msec or 4 sec. Recall and recognition tests followed. Brief presentations yielded as much boundary extension as long presentations. In Experiment 2, picture triads were presented at a rate of 333 msec per picture with no interstimulus interval. After 1 sec, one picture repeated and remained in view while 60 Ss indicated whether it was the same or showed more or less of the scene. Even when conditions mimicked a series of rapid eye fixations, boundary extension occurred. They concluded that the presentation of a picture appears to activate a perceptual schema that allows observers to understand it in a larger context, and this process distorts memory for its actual boundaries.

Intraub (2004) extends the phenomenon of boundary extension to a new modality: touch. She showed participants six different "scenes" composed of real, physical objects (groups of ordinary things like toys, books, and toiletries), each demarcated by a "boundary" of black cloth. She then had an assistant remove the boundaries and asked the participants to mark where the boundaries had been. As expected, they placed the borders well beyond where they had been in the original scenes. She then blindfolded another set of participants and showed them the same scenes, with an easily detectable three-inch-tall wooden "boundary" replacing the original cloth border. When they returned later to the same scenes with the border removed (and still blindfolded), they were asked to place wooden blocks where the borders had been.

Finally, she repeated the "blindfolded" condition with a volunteer who had been deaf and blind from early childhood. Intraub calls this participant (KC) a "haptic expert" because she has spent her entire life negotiating the world by the sense of touch. At the age of 25, KC was a successful college student who could easily identify the objects in the experiment by touch (the only difficulty was reminding her to use the entire 30 seconds allotted for each scene, so as to match the blindfolded group). The results for KC were the same as those from the blindfolded group: She extends boundaries in the same way sighted people do. There was one difference between the blindfolded and the vision group: the vision group showed more boundary extension than the blindfolded group, suggesting that the modalities of sight and touch are not precisely analogous. The experiments reported in chapter 6 of this thesis appear to show a phenomenon that it similar to boundary extension. Viewers were asked to judge the bodyweights of models from face and neck portraits only: The models bodies were never shown in any condition, and yet the observers were able to complete the task without questioning the difficulty of size estimations of targets that cannot be directly perceived. It may be that in normal perception, people are highly adept at visualising scene content that is not directly seen and that this phenomena can be demonstrated in photographs of people in addition to landscapes, inanimate objects and other non-human scenes.

3.5. Visual Direction and Cyclopean Vision

Van Ee et al (1999) examined the effects of perceived visual direction near an occluder. When an opaque object occludes a more distant object, the two eyes often see different parts of the distant object. Hering's laws of visual direction (1879) make a prediction for this situation: the part seen by both eyes should be seen in a different direction than the part seen by one eye. They examined whether this prediction holds by asking observers to align a vertical monocular line segment with a nearby vertical binocular segment. They found it necessary to correct the alignment data for vergence errors, which were measured in a control experiment, and for monocular spatial distortions, which were also measured in a control experiment. Settings were reasonably consistent with Hering's laws when the monocular and binocular targets were separated by 30 arcmin or more. Observers aligned the targets as if they were viewing them from one eye only when they were separated by 2 arcmin; this behavior is consistent with an observation reported by Erkelens et al (1994). The same behavior was observed when the segments were horizontal and when no visible occluder was present. They observed that perceived visual direction when the two eyes see different parts of a distant target is assigned in a fashion that minimizes, but does not eliminate, distortions of the shape of the occluded object.

3.5.1 Ecological Plausibility of alternative spatial configurations.

Nakayama and Shimojo (1990) investigated "da Vinci stereopsis": depth and subjective occluding contours from unpaired image points. They reported; "Distant surfaces are occluded by nearer surfaces to different extents in the two eyes, leading to the existence of unpaired image points visible in one eye and not the other. An ecological analysis of the real world situation that could have given rise to such unpaired points indicates the presence of a depth constraint zone, defined by visibility lines between which possible real world points must lie." The leading edge of this zone starts at the edge of a fused binocular occluding surface and recedes linearly with increases in horizontal distance to the unpaired point. Their psychophysical evidence indicated that the human visual system makes use of this unpaired information in a remarkably adaptive manner, showing an increase in perceived depth for increasing horizontal separations between the unpaired target and fused edge, at least over a significant angular range (approx. 25-40 min arc). They also showed that unpaired points in binocular images can lead to the formation of subjective occluding contours and surface having the qualitatively appropriate sign of depth. Furthermore, they showed that the visual system could not recover depth of unpaired points camouflaged from the other eye against silhouettes. Their findings indicate that the visual system makes use of occlusive relations in the real world to recover depth, contour, and surface from unpaired points.

They finally reported that the novel emergence of subjective occluding contours from unpaired monocular stimuli "raises the possibility that this process is mediated by visual experience, built up by the association of unpaired points and occluding contours." The visual stimuli used in all of the stereoscopic experiments reported in this thesis feature unpaired image points visible in one eye and not the other. These points are of special relevance in they are found near the centre of the virtual stimuli from which subjects were asked to make critical size estimation judgements. These zones are reproduced in a spatial configuration that exactly matches the same "monocularly viewed" areas found on directly viewed real scenes and objects. So the stereoscopic targets used in this thesis match the real world image characteristics of large double-curvature objects and allow these virtual images to act as accurate "stand-in" representations so that real people and objects did not need to be used in the methodologies used in experiments 1-4 reported in chapter 5.

3.5.2 The Role of Cyclopean Vision

Erkelens and van Ee (2002) studied the role of the cyclopean eye in vision and came to the conclusion that it was "sometimes inappropriate, always irrelevant." They describe how during binocular fixation, the eyes usually point in different directions, and yet, each object is judged to lie in a single direction. They report that "it is commonly believed that a particular location in the head serves as the origin for such directional judgments. This location is known as the cyclopean eye. We argue here that observers can judge visually perceived directions from angular information alone, and do not require positional information supplied by a cyclopean eye. We show that experimental findings reported as evidence for the cyclopean concept can also be explained solely by angular information without the need for a cyclopean eye. Recent findings concerning binocular shape perception and the cyclopean illusion demonstrate that binocular perception is incompatible with vision from a single vantage point. The concept of the cyclopean eye is sometimes inappropriate and always irrelevant as far as vision is concerned."

The Erkelens and van Ee claims were rebutted by two papers. The first, by Ono, Mapp and Howard (2002) argued against their paper that the concept of the cyclopean eye is "always irrelevant as far as vision is concerned" and that perceived direction during monocular viewing is based on the signals of the viewing eye only. In their first experiment, they presented a pair of small lights on a visual axis and measured the absolute visual direction of the near light with reference to different parts of the face. The near light appeared in front of the bridge of the nose or very near it, contrary to what was expected from Erkelens and van Ee's s claim that monocular stimuli are seen in their correct locations.

In their second experiment, they replicated Erkelens' experiments with measurements of "phoria" and analyses of eye movements. The results confirmed his finding that the cyclopean illusion occurred rarely in the monocular condition, but their phoria and eye movement data provided the basis for a very different interpretation. The data showed that the oculomotor signal in his particular monocular condition was considerably weaker than in his binocular condition; therefore, the rarity of the monocular cyclopean illusion is not surprising. Moreover, since both claims above are based on an over-generalization of the results of Erkelens study, neither claim was "persuasive."

Kokhotva et al (2005) also challenged the Erkelens and van Ee claims. They said: "Wells-Hering's laws summarize how we process direction and predict that monocular stimuli appear displaced with respect to the viewer, but not with respect to other seen objects. Erkelens, C. J., & van Ee, R. (2002). The role of the cyclopean eye in vision: sometimes inappropriate, always irrelevant, criticized this view and claimed that there is no perceptual displacement of these stimuli. We challenge their claim and improve on shortcomings of past studies. LEDs were monocularly presented to the observers, without their knowledge of which eye was being stimulated. Viewing distance was 9-10 cm; fixation distance was 30 cm. and the observers reported the perceived relative and absolute directions of monocular stimuli. Our results are consistent with Wells-Hering's laws."

The paper by Erkelens and Van Ee (2002) argues that the cyclopean view the brain forms from a binocular input is essentially the same as 2D image perceived from the same viewpoint. However, previously published research by Harper and Latto (2001), reproduced as appendix A in this thesis, showed that this is not the case and normal human perception is measurably different between monocular and cyclopean percepts. It is likely that the Erkelens and

Van Ee paper was accepted for publication before the Harper and Latto paper was published, so it is unsurprising that they make no mention of our findings in their review of previous studies. However, had they been able to so, it may have been difficult for them to make such claims. It is interesting to note that they have not followed-up on this study with new experiments that deal with the counter claims made by Ono, Mapp and Howard and Kokhotva et al. It may be that the role of the cyclopean eye is in fact a well understood and robust phenomenon that is the key explanation as to why objects can appear to change shape when viewed in monocular and binocular conditions, one of the topics of investigation of this thesis. The following chapter reviews the methodology used for the stereoscopic imaging techniques used that were informed the psychophysical research detailed in chapter 3. These experiments are reported in chapter 5.

CHAPTER FOUR

Orthostereoscopic methods and equipment

4.1 Introduction

The pilot studies described in Appendix H indicated that the methods and equipment used were inadequate in a number of ways. This chapter describes the development of the othostereoscopic imaging techniques revised methodology used in all of the subsequent stereoscopic experiments reported in this thesis. The details of how these images were used in a series of three experiments are reported in Chapter Five.

4.2 Experiments for Projected Orthostereoscopic Images.

4.2.1 The camera and projection system. The Pentax LX cameras used in the stereoscopic photography reported in appendix H were of a professional specification and build quality, and as such tend to be larger and heavier than lower-specified contemporary designs. In side-by-side comparisons, their 55mm standard lenses also gave a very slightly larger and therefore more telephoto image than a standard 50 mm lens (see Section 3.4). When attached to the flash-bar base plate so that they could be mounted on a tripod, the interaxial lens separation of the cameras was 72mm. In the stereography literature reviewed in Chapter Two, the technique used to describe photography that matches the interaxial separation and magnification of the human visual system (HVS) is orthostereography. However, the Pentax lenses used had greater magnification than the HVS and that the cameras bodies were over 10% too far apart. This would induce the effect known as dwarfism (15), which is an artefact of stereoscopic photography whereby lens interaxial distances wider than 65mm give objects the appearance of being smaller than life-size. Giantism is the opposite effect whereby interaxials smaller than 65mm give scenes the appearance of being larger than life-size. Interestingly, the standard technique uses parallel axis lenses (see figure 4.1 below) and does not converge each lens axis on the centre of focus (region of interest) and in this respect does not match the normal conditions of the HVS, as shown in figure 4.2.

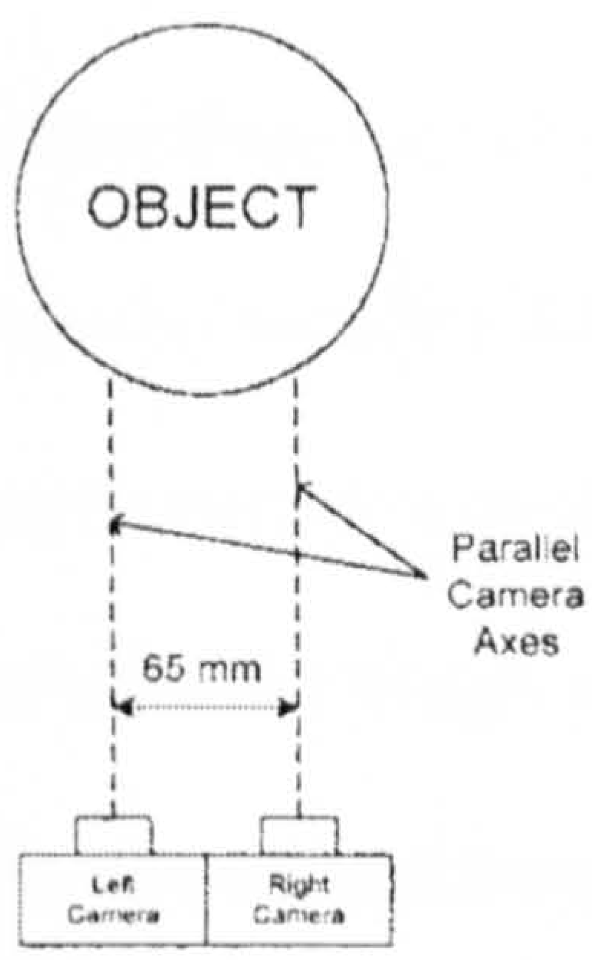


Figure 4.1 | The standard parallel axis camera configuration used in most stereoscopic imaging

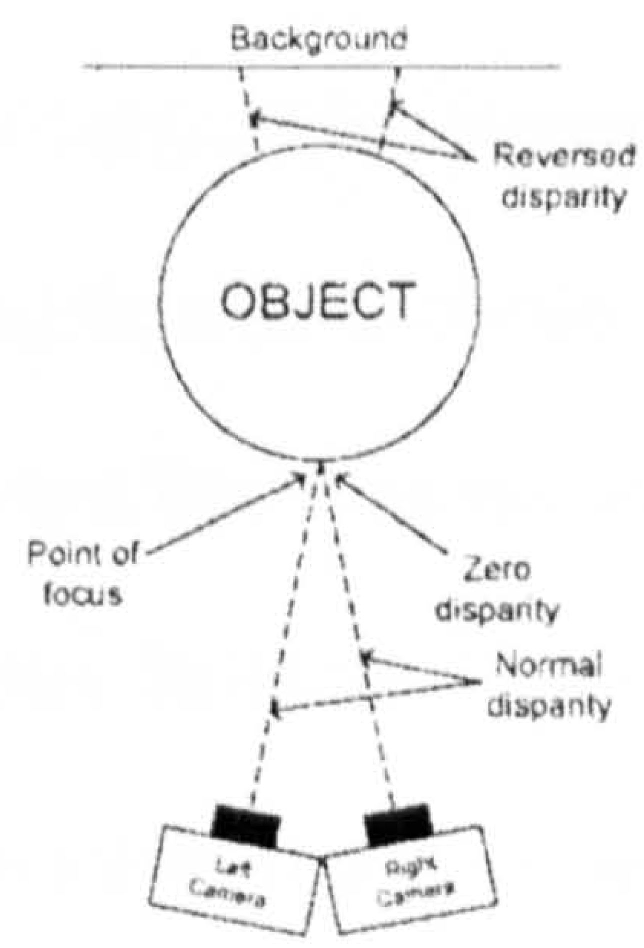


Figure 4.2 A simulation of the naturally convergent orthostereoscopic configurations using cameras whereby both vergence angles are linked to the distance or object that the lens is focused on.

This seems to be done for reasons of technical simplicity or because it has not been considered to be of importance. However, as the plane of sharp focus for optically corrected lenses is a plane at 90^0 to the lens axis, a parallel lens system cannot reproduce the effects of convergence normally experienced by the HVS. Thus, a parallel lens system cannot reproduce the focused and defocused regions of a converged axis system and was the initial reason why it was not favoured as a basic design feature of the second stereo camera.

Two 35mm SLR cameras were required to make a stereo camera capable of producing true orthostereoscopic images. They must have 50mm standard lenses to match the focal length of the human eye. They must also be small enough so that when the two bodies were attached to a tripod mount they would give an interaxial lens separation that approximated the HVS interocular distance of 65mm. The Olympus OM1 was found to be ideally sized and specified. Though considered to be a semi-professional camera, its lens resolution was of the highest quality and could deliver high resolution 35mm transparencies. It was also a much smaller 35mm single lens reflex camera than the LX (the smallest professional 35mm camera available at the time) and could therefore deliver a smaller lens interaxial distance. The combined cameras gave a lens interaxial distance of 66mm when first mounted onto the tripod attachment bar with the lenses in the parallel position. Balsa wood wedges were used to converge the camera bodies (as on the LX stereo camera) and it was found that the interaxial measurement from the centre of each lens was now approximately 62 mm when the lenses were converged at the required 1.68 metres. The stereo camera was fixed into this configuration using high strength glue which allowed the camera to dispense with the previously used straps used in the earlier camera that made it impossible to change films without dismantling the whole arrangement. The dual shutter-release cable used previously was fitted and the revised stereo camera was suspended from a tripod mount. The lens height (measured from the lens centre to the ground) was again set to 1.33 metres.

The projection system too was revised. The previous Elmo projectors were of low brightness (150W.) and had 100mm projection lenses that meant they had to be over four metres away from the screen in order too give life-size magnification. The Kodak EKTAPRO. 35mm projectors used for the later experiments had 250-watt lamps and 85mm F2.8 lenses that gave an image almost twice as bright as it predecessor. This advantage was reinforced by the use of a new high-gain (Widescreen Centre lenticular element 1.4metre) metalized screen that gave an image that was twice as bright as the older design. A calculation was made to establish what would be the correct lens interaxial separation of the optical axes of each projection lens required to project the images at the same angle of convergence as the original camera configuration. As the 85mm lenses on the Ektar projectors are 1.7 times longer than the 50mm lenses on the Olympus cameras, the correct interaxial separation is approximately 110mm (65mm x 1.7). This distance was established using the reflex mirror method (figure 4.b). The 1.68 metres camera to subject distance multiplied by 1.7 gives a projector to screen distance of 2.87 metres. Thus when the projectors were placed at this distance from the screen, a life size image was projected on the screen: The projected model's interpupilary distance was measured, it was found to be almost identical to the 63mm interpupilary distance measured from the real person. The projectors were placed on a flat sheet of plate glass (0.5 x 0.3 metres) that acted as a stable platform that could be set to be perfectly horizontal using shims and a professional quality spirit level. The height of the lens centres were set to 1.33 metres, to match the elevation of the Olympus camera lenses used for all of the photography.

A reflex mirror arrangement was required to achieve the interaxial distance between the lenses of the two projectors. The normal parallel lens axis arrangement is shown below in figure 4.3. This is because they are very large devices and the nearest the lenses can normally be to each other when the projectors are placed side-by-side is over 350 mm. And even this distance is too close to allow the fans enough air to cool the powerful halogen lamps. The reflex arrangement allows the optical axes to be as close as is required for the correct interaxial separation. A surface-silvered mirror was used to ensure there was no secondary image from a glass surface. And in order to match the geometry of the images captured by the cameras, the projectors were converged to align of the centre of the projection screen (see figure 4.4.). The transparencies in the associated projector were laterally reversed (in addition to the normal inverted loading) in order for the image to be projected with the correct orientation. Ideally, the polarizing filter for this optical axis should be much larger than the screw thread filter used in the previous studies and placed in front of the

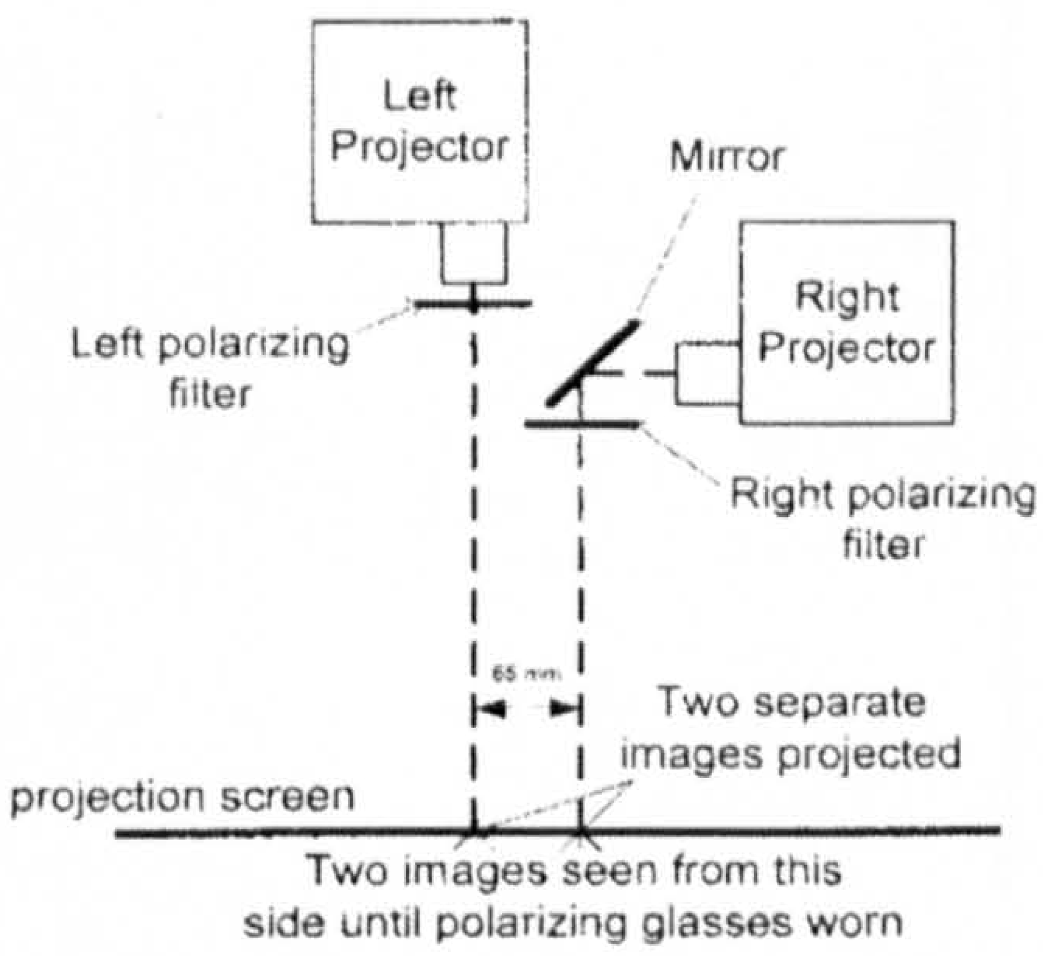


Figure 4.3 The matched rear-projection configuration for the parallel axis camera configuration of Figure 1

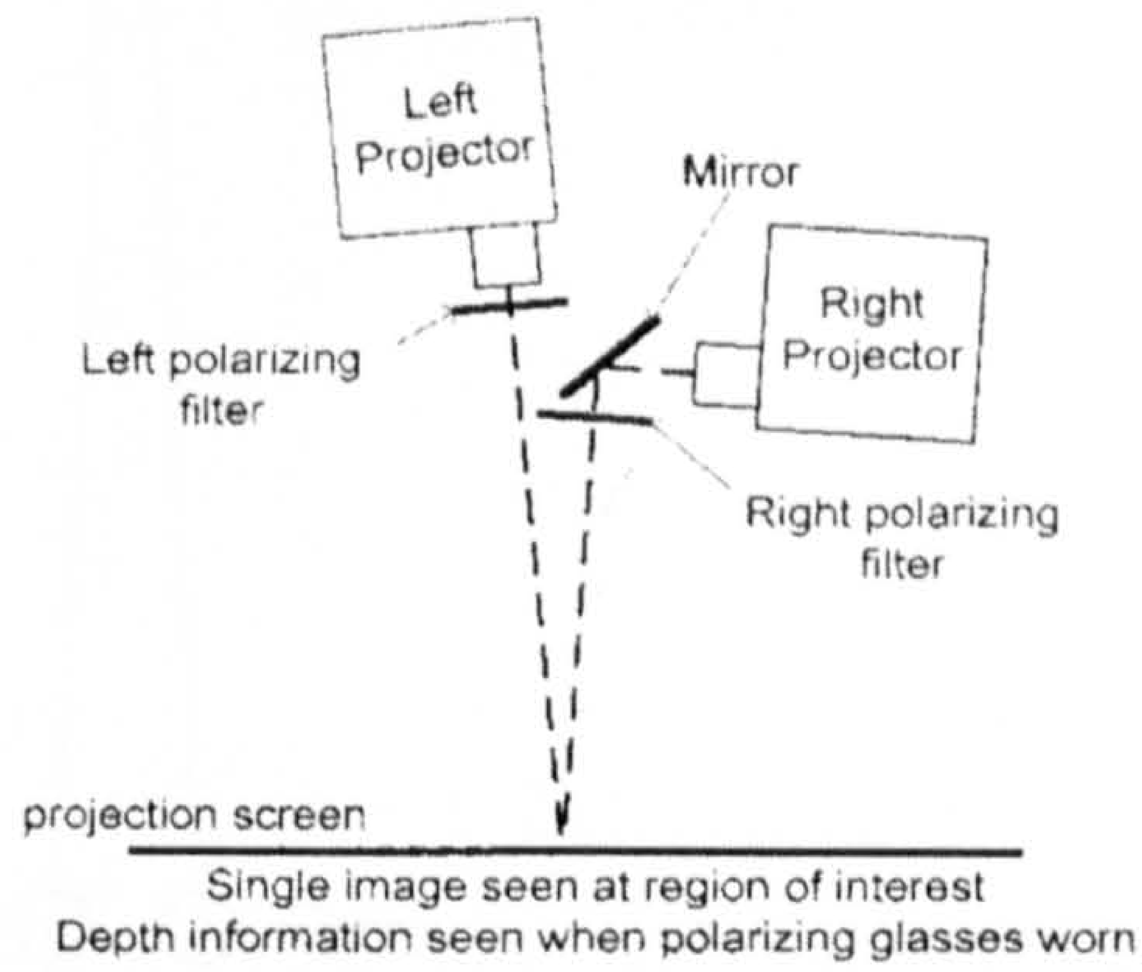


Figure 4.4 The matched convergent projection configuration for the orthostereoscopic image camera configuration of Figure 2.

mirror. This is because a surface-silvered mirror can slightly depolarize any light reflected from it due to dust on the surfaces of the lens, filter and mirror surface. So it is better if image is polarized after it is reflected. This was not done in these experiments, but careful cleaning of these surfaces ensured that depolarization could not be observed and that cross-talk between the optical channels was kept below 5% as measured using a photographic light meter.

The participants viewing position is shown below in figure 4.4b. The observers viewing position was carefully controlled by locating the eye distance from the screen (suing a plumb line) at 1.68 metres. To avoid the participant occluding the image from the projectors, an adjustable stool was used to lower each participant to the required viewing height.

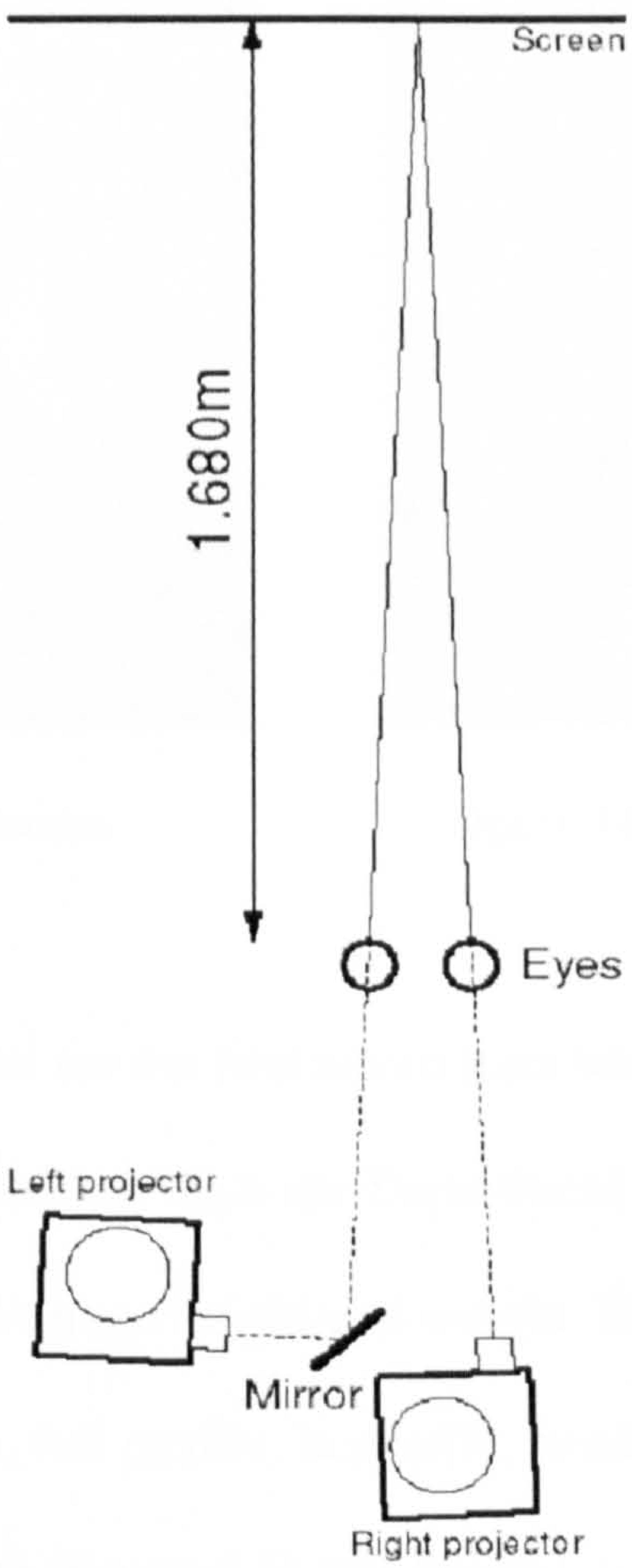


Figure 4.4 b. Shows a plan view of the participants position for the first experiment

Figure 4.5 is right camera image.

Figure 4.6 is the left camera image

4.2.2 Test Image. The model for the first stereo tests with the new camera was a female colleague (Gabrielle Salfati) from the Department of Psychology. She was in her late twenties and was of average height and weight. She posed in a number of positions (full face to camera, full profile, ¾ profile, hands on hips etc). The pose chosen for the test stereograph (figure 4.5) was helpful in that it extended the stereoscopic depth of the area reproduced in comparison with the previous images. This was because the Coke can is closer to the camera and therefore the viewer when it is reproduced with stereoscopic projection. The Coke can also gave a useful standard object to see if the display suffered from size or shape distortions (this image was subsequently always referred to as the "Coke can" photograph). The dress was made from a highly reflective yarn that directed slightly different brightness to each camera. The flat photographic background was the same one used in all of the photography to date, as was the lighting plan layout of the scene. The transparency film used was a hi-resolution Fuji 50ASA film that also featured a finer grain structure and slightly more accurate colour reproduction.

To establish the levels of naturalness possible with this camera, film and projection system, a unique side-by-side demonstration was staged between a real and virtual image. The measurements of the photographic layout of the test photograph were recorded and the set-up was accurately reproduced some weeks later. The model was asked to return and to adopt the same pose (holding the Coke can). Adjacent to this set up was the stereoscopic projection system described in section 3.2. The screen was placed at the same distance from the camera position (next to the model) and her image was projected at life-size magnification on the screen. Thus, it was possible for a viewer in the camera position (the camera having been removed) to observe the real model and her virtual reproduction by a simple switch in gaze from left to right. The result was remarkable, in that one got an impression of the model having a virtual twin: *The stereoscopic depth and character of the reproduction seemed to be almost identical to the real scene. The model had the same size and proximity to the viewer in both scenes, the Coke can was the same size and distance away, as was the blue background.* The relative brightness between the two scenes was also surprisingly well matched. The demonstration was viewed by five people, and it was agreed that the only significant differences between the real and virtual model were the slightly lower resolution in the projected image, higher contrast and slight shifts in the consistency of the colour reproduction. As these factors were not considered to be relevant to the proposed experiments, it was agreed that the reproduction was both life-like and free of subtle geometric distortions. [i] This combination of equipment and film was used in the stereoscopic experiments reported in this thesis. When computer generated images (CGI) were used, the virtual cameras were set to the same specifications as the real cameras. This ensured that the projection conditions could remain constant throughout the stereoscopic experiments, regardless of the method of image capture used to generate the stimuli. The test image was shown to every participant in the stereoscopic studies as part of a familiarisation procedure before the main experimental images were shown.

4.3 Convergence, Vertical Disparities and Orthostereoscopic Imaging

The author was asked by his sponsors, The Independent Television Commission (ITC), to organise an "Image Science Workshop" to be held at their offices in London in November 2000. The two-day event was intended to bring together leading neuroscientists, psychologists and imaging technologists to discuss the brain and perception of mediated images. The event was also an opportunity for the author to repeat his presentation of stereoscopic images and data (first shown at the Presence 2000 Conference in Delft) to a more specialised and potentially more critical audience. It was expected that if there were any flaws in the stereoscopic techniques used or that the slimming effect observed was caused by an artefact of the display, this audience would be able to identify the cause. Amongst the many distinguished academics present were Brian Rogers, the co-author of Binocular Vision and Stereopsis (1994) and Seeing in Depth (2002) with Professor Ian Howard. These reference works are the regarded as the most comprehensive text in stereoscopic perception and are cited by the most authors and researchers in the field. After seeing the presentations of the slimming effect of othostereoscopic images, most of those present reported it to be a clear effect that they had not seen before.

In Delft and at the ITC workshop, all of the 2D stimuli were projected from identical slides in both projectors and converged on the screen. In the pilot studies, the

2D swimsuit images were projected synoptically this way, but the 2D peanut images were presented on cards held up at arms length, subtending the same visual angle as the 3D image. There was some discussion between Michael Morgan, John Frisby and Brian Rogers about the stereoscopic techniques used and if they could cause the slimming effect. Brian Rogers made an important observation about incorrect vertical disparities in the 2D/synoptic images used in the experiments. He argued that when identical 2D images are projected from two converging projectors, the vertical disparities caused by convergence are unnatural and could cause misperception of the 2D stimuli. The author replied that the synoptic projection technique used in the current demonstration was only used in the swimsuit experiment. The peanut experiment used a 2D target so the vertical disparities seen there were correct and yet the apparent slimming effect appeared to be the same. He replied that this was true, but there may be some unknown effect of accommodation because the stimuli being compared were at different distances. So viewing hand-held 2D targets at the same visual angle as an image presented stereoscopically might cause a difference between apparent sizes in any experiment using that technique. He further noted that in all of these experiments, the only image with veridical vertical disparities was the 65mm stereoscopic condition used in Experiment 3 (Section 5.4.1). It was agreed that in all subsequent experiments, the distance, visual angles and vertical disparities in all stimuli should be as closely matched to the "natural" condition in all of the 2D and 3D stimuli presented. He was invited to Liverpool to inspect the revised methods and subsequent experiment when they were ready in the next year.

Brian Rogers came to Liverpool in October 2001 to see if images made with corrected vertical disparity images and accommodation had addressed the problems he had identified the previous year. It was extremely important that he accepted that the orthostereoscopic projection technique used in these experiments was both ecologically valid and that any claims made for it are supportable. Surprisingly, he had no knowledge of any previous experiments that used lens/projector convergences to same-size magnification and the use of lens interaxials that approximate the human inter-pupilary distance. He was also not familiar with the important Spottiswoode and Smith papers on 3D film that was used by the ITC and its associates (Dumbreck 1995) in the design of their orthostereoscopic television systems. Their papers were not referenced in "Binocular Vision and Stereopsis," however, this was corrected in the updated edition "Seeing in Depth" published in 2004.

Brian Rogers viewed all of the stimuli shown in the four stereoscopic experiments reported in Chapter Five. His attention was especially drawn to the final CGI peanut images rendered with fully corrected vertical disparities. And once again, the slimming effect was apparent and similar in size to the images seen previously in London. He agreed with the Spottiswoode and Smith assertion that orthostereoscopic imaging was "the condition of perfect reproduction" so it may be that the 2D condition was presenting the inaccurate size information. Most importantly, he confirmed that linking convergence angles on the camera and projector (and therefore reproducing the point of zero disparity in the original scene without disparity in the display) was a technique that would assist viewers to fuse the stereo images and be free of avoidable distortions. As to the general claim that objects can "change shape" in a stereo image, he accepted that this was possible. But he reserved judgement on the assertion that 2D images cannot provide a condition of perfect reproduction of close-up 3D objects until he had seen experimental evidence to support it. These experiments and revised methods used are reported in Chapter Five.

CHAPTER FIVE

EXPERIMENTS 1-4

The Slimming Effect of Orthostereoscopic Images

5.1 Introduction

The following four experiments use orthostereoscopic imaging to investigate the slimming effect of binocular disparity and the corresponding fattening effect of viewing images synoptically, without disparity. The first three were reported in the MIT Press journal, Presence: Teleoperators and Virtual Environments. Entitled *Cyclopean Vision, Size Estimation and Presence in Orthostereoscopic Images by Bernard Harper and Richard Latto,* it appeared in Presence Vol. 10, No. 3, June 2001. (A copy is included in appendix A.) The technique used in all four experiments mimics the geometry of the human visual system and uses convergent axis stereography with the cameras separated by the human interocular distance. It simulates viewing angles, magnification and convergences so that the point of zero disparity is reproduced without disparity in the display. The image sizes of the stereoscopic stimuli are identical to the synoptic 2D stimuli so that in all conditions, the displayed target subtends the same visual angle on the retina. The only difference between the images in each experiment is the stereoscopic disparities conveyed. The experiments compared size, weight or shape estimations (perceived Waist-Hip ratio) in 2D and 3D images of the human form and real or virtual abstract shapes:

- Experiment 1: Swimsuit Experiment.
- Experiment 2: Slimming Effect with a non-human object.
- Experiment 3: The First Peanut Experiment.
- Experiment 4: The Second Peanut Experiment

5.2 Experiment 1: Swimsuit Experiment

5.2.1 Introduction

The pilot swimsuit study described in section 3.3.1 was flawed in a number of ways that this study was designed to address. The Pentax cameras were replaced by the revised orthostereoscopic (Olympus OM1 based) stereo camera. The pilot study also had possible confounds in the image capture procedure used because some of the volunteers had been swimming and were wet. Others were canvassed before swimming and were dry. All wore their normal swimsuits (and were hard to identify afterwards because they were un-numbered) and it was realised that many of these garments were specifically designed to subtly disguise the underlying shape of the wearer. As this could possibly affect the outcome of the study, three black single-piece bathing suits of a "one size fits all" design were purchased for the female participants. These would be worn by any model who did not bring a black, single piece swimsuit of their own. Each model was canvassed by an assistant who recorded their name, height, weight, contact details and asked them to sign a release form allowing the images to be used for scientific purposes only. Each model was then asked to change into a black swimsuit and to wear a number indicating their position in the sequence of photographs. The models were then photographed wearing the model number and in a standardised pose. The pose was a semi-profile (3/4 view) with hands on hips as it was considered as the most representative from the various posed used in the first study and the test shots from the Olympus stereo camera. In all other respects, the images were captured in the same way as in the pilot study.

5.2.2. Method

5.2.2.1 Equipment and experimental design. The participant's stereo-acuity was measured using the TNO test for stereoscopic vision (TNO 1972). It uses red-green anaglyph glasses to view printed stereographic plates in a book. The subject wears the anaglyph spectacles provided to view a series of stereographs that decrease in depth with each subsequent presentation. A score in arc-seconds is recorded for the last two correct depth judgements made by the user. The range of the test is 480 to 15 arc-seconds which covers over 95% of the normal human range. Each individual stereo acuity score could then be used to compare the relative stereoscopic vision of each of the participants. People with very low scores would indicate poor stereoscopic vision and their data could be excluded from the analysis.

The projection system used was as described in section 3.2. The hi-gain metalized screen was used and the viewing position was set at 1.68 metres from the screen. This was the same distance as the original camera to subject distance and ensured that when the image was projected orthostereoscopically to life-size magnification, it was a close as possible to the ideal geometry described in section 4.1.2 of the Coke can photography. As with the previously reported method of presentation, the participant's task was to assess which one of the original seven bodyweight descriptors was most appropriate to describe the apparent bodyweight of the models in each of the projected images. However, in this experiment the participant would see a person's projected image once only, so that there could be no "memory effect" of seeing the same person projected twice. Also, the 2D images would now be seen while wearing the polarizing glasses. This meant that the 2D condition would need to be presented synoptically[1]. This was achieved by copying all of the left camera images from each of the stereo-pairs to a same-size reproduction. The left camera film was chosen as the models images were larger on this camera because it had a more frontal view. The right camera image which had a viewpoint that was more of a profile and therefore could be interpreted as being slimmer if it were copied synoptically. Copying the "fatter" image would mean that the relative slimming effect of binocular disparity would be smaller, but avoided a possible artefact using this method. Fujichrome copying film was used in a dedicated 35mm transparency copying system to accurately reproduce the left camera transparencies to the same magnification, brightness, colour and contrast as the originals. These copies were physically measured using callipers and were later confirmed as being identical in size to the originals and two sets of copies were made. The pairs of left slide copies were then mounted into registration mounts and projected using the EKTAPRO projectors. The images were also measured at the projection screen and found to be almost exact copies of the originals.

The same-size synoptic pairs allowed two sets of photographs to be made. Each set had images of ten different people, five to be stereoscopic and five synoptic. A random draw from the original set of body images was used to establish which ten images should be in the study and then which should be stereoscopic or synoptic. So randomised mixtures of ten different stereo and synoptic images were presented to each of the participants, and each viewer did not see the same person projected twice.

[1] Synoptic presentations are sometimes called psudoscopic stereo, as it was noticed with the early stereoscopic viewing devices that if two identical photographs were inadvertently loaded as a stereo pair, many viewers reported a compelling stereo effect. This effect is discussed in Chapter seven of this thesis.

The presentation method was the same for all images and the polarizing spectacles were worn throughout.

A series of familiarisation stereo images were gathered from the original stereo test images and familiar local landmarks photographed by the author. These were shown before the experimental stimuli were presented to ensure that the participants were comfortable with the procedure before they were required to make a size judgement. The participants saw the images from behind a small desk that had a low powered lamp directed at the response sheet. The lamp was baffled to ensure that light was well-directed so that no spillage could fall on the screen and reduce the quality of the projected image. In order to ensure that the descriptors were easy to understand, a large card with the descriptors was placed next to screen and the descriptors were written on the response sheet adjacent to the tick boxes. After the study was completed, a short de-briefing session was used to explain the experiment and get feedback from the participant.

It was considered that as the slimming/fattening effects appeared to be large, a small sample size might be appropriate. A calculation would confirm that a small sample size was adequate, especially as the pilot studies showed a large effect with a small number of people and the effect was demonstrable using the most simple of descriptive statistics. The pilot study indicated that as each test could take 30 minutes to run with each participant, a low sample size was desirable. A sample size of 28 participants was agreed because it was approximately double the number previously tested and would allow for a large number of participants data to be excluded if they did not have adequate stereo vision or had spoiled their response sheets. Their data therefore would be unusable either because they had monocular perception (in that they could not fuse the stereo images) or because they had written a response in the wrong box for the appropriate stimuli. It transpired however that all of the responses gathered were used, even though some participants scored very poorly on the TNO stereo acuity test. Two people did not record any score on the TNO test as they were almost blind in one eye and they decided not proceed any further with the study. To analyse the data, the responses were converted into a Likert scale. Thus a score of 4 equalled a response that the body image presented was the correct weight, 7 was very overweight, 1 very underweight.

Descriptor	Abreviation	Likert value
VERY OVERWEIGHT.	V.O.	7
OVERWEIGHT.	O.	6
SLIGHTLY OVERWEIGHT.	S.O.	5
CORRECT	C.	4
SLIGHTLY UNDERWEIGHT.	SU.	3
UNDERWEIGHT	U.	2
VERY UNDERWEIGHT	V.U.	1

This data was transcribed into a statistical analysis program (SPSS) and an analysis of variance (ANOVA) was completed so that the mean perceived weights of the models in synoptic and stereoscopic conditions could be determined.

5.2.2.2 Participants. Twenty-eight Liverpool University undergraduates were tested individually.

5.2.2.3 Procedure. Participants began by taking a stereo acuity test (TNO test forstereoscopic vision) and viewing a series of 3D slides to accustom them to stereo viewing. They were then shown images of the ten models in a random sequence of stereo and synoptic 2D images, so that each model was never shown to the same subject in both 2D and 3D. Half the subjects saw Models 1-5 in stereo and Models 6-10 in synoptic 2D, while half saw 1-5 in synoptic 2D and 6-10 in stereo. Trials were self-paced and during each presentation subjects rated the bodyweights of each model on the 7-point Likert scale labelled from VERY OVERWEIGHT to VERY UNDERWEIGHT.

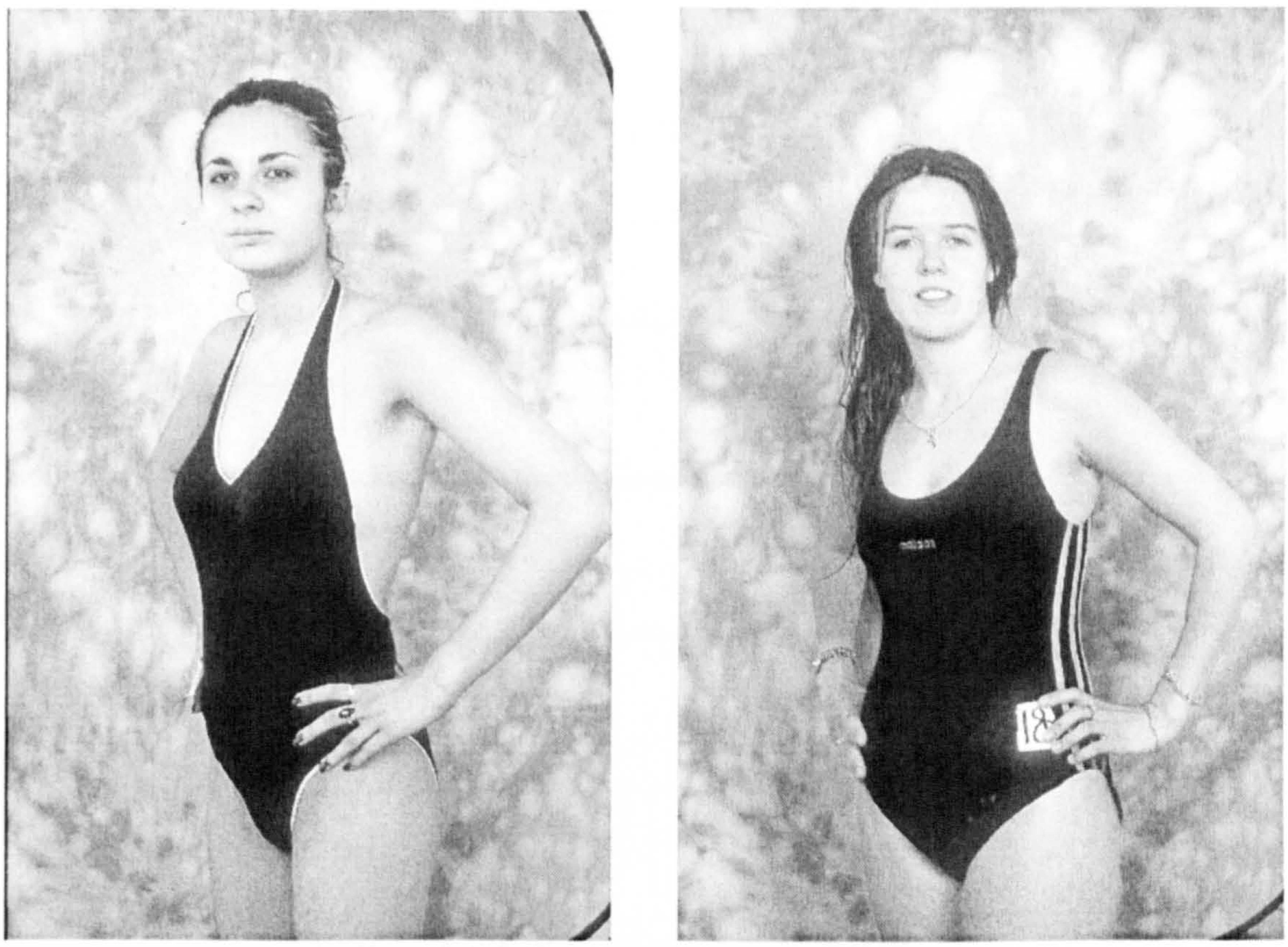


Figure 5.1 These images are from the swimsuit experiment sequence. The model on the left is wearing the "one size fits all" swimsuit used by most of the models. The model on the right is wearing her own black swimsuit, which was allowed under the previously agreed methodology. The presentation order was derived from the order in which they were photographed so that the consecutive number order remained intact.

5.2.3 Results.

The mean perceived weight estimates of the ten models viewed either stereoscopically or synoptically are shown in Figure 5.2. As the means and the very small standard errors indicate, there was a strong centralising tendency in the participants' judgements, partly because the range of bodyweight in the models was not high. The debriefing sessions also indicated that there was some reluctance by the participants to make negative judgements on the models. Nevertheless, the models were rated slimmer when viewed stereoscopically ($t(9) = -3.55$, $p < 0.01$) and the effect size ($r = 0.76$), calculated using the procedure in Field (2005, p.294), was substantial.

a one factor (viewing condition) ANOVA showed that the models were rated as significantly slimmer when viewed stereoscopically ($F=15.072$, $df=1$, $p=0.001$).

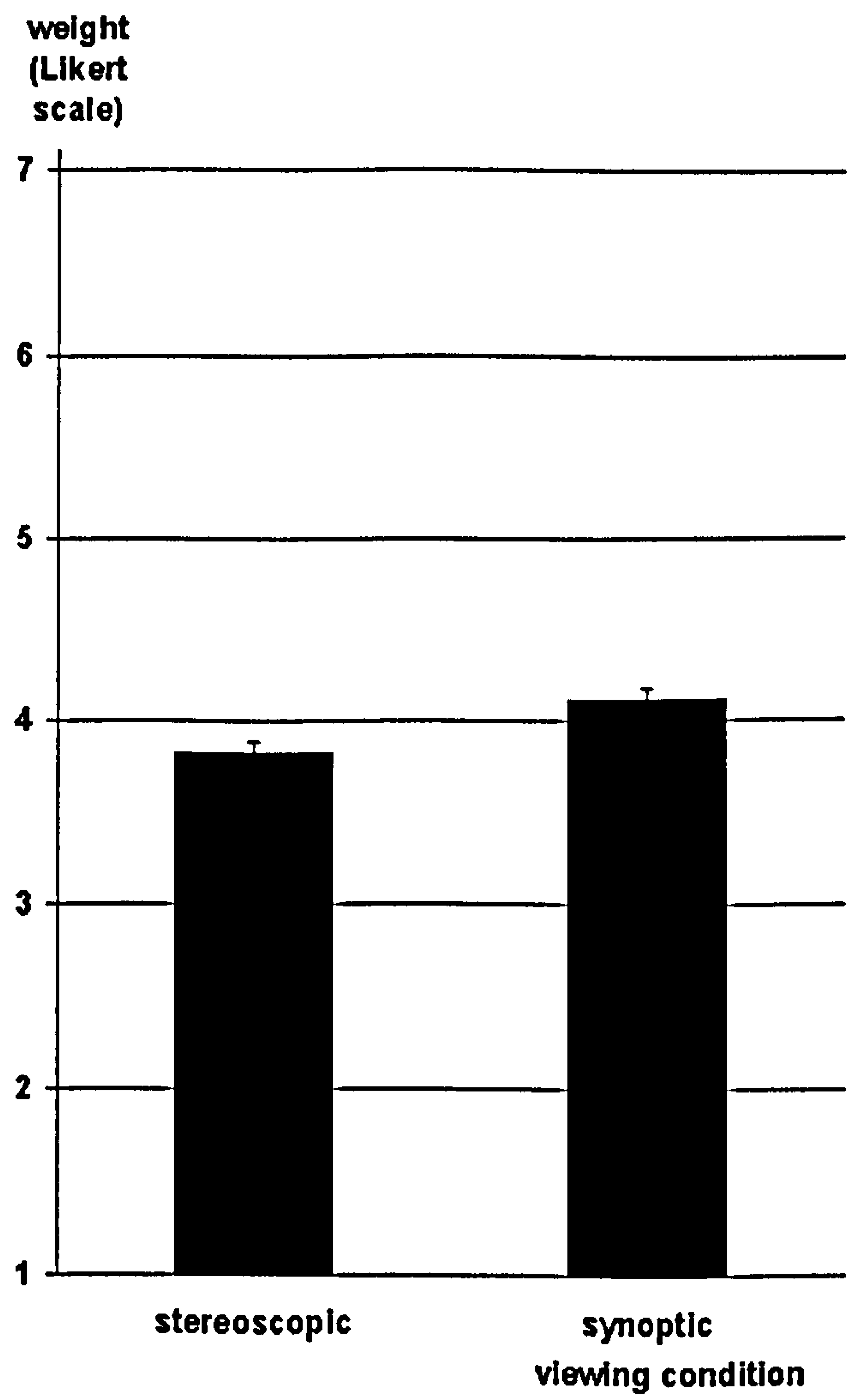


Figure 5.2 Experiment 1. Effect of viewing condition on mean perceived weight. The error bars on the op of each column indicate the standard error of the mean.

5.3 Experiment 2: Slimming Effect with a non-human object

5.3.1 Introduction

Each participant in the swimsuit experiment was given a short debriefing after the presentation. This was to explain what the basic principle was behind the experiment, discuss their individual data (especially those who recorded low stereo acuity on the TNO test)[2] and to get some feedback from them about the experience. Of particular interest was to find-out if any of the participants suffered from discomfort or eye-strain during the presentation. Clearly, any reports of difficulty in seeing any of the stimuli might be diagnostic of problems with the stereo image quality or of suppression that would cause a monocular percept of the stereoscopic stimuli.

Most participants were unaware that the presentations were randomly changing between monocular and stereoscopic conditions. This is not too surprising as the monocular condition was synoptic. Throughout the history of stereography such psuedoscopic images (Seeing in Depth VII, p.491) have often been mistaken for true stereographs. Secondly, informal reports from several participants suggested that they sometimes felt they were in the presence of "real" people. Perhaps increased presence in the stereoscopic condition may have led the participants to give judgements that were less harsh to models that they felt were more present in the laboratory? Although this seems unlikely, particularly as most viewers were unaware that the presentation

---

[2] For ethical reasons, it was decided that all of the participants should be informed of the significance of their individual TNO scores, especially if they had lower than average measured stereo acuity. The advice taken from the Department of Orthoptics at Liverpool University was that people are always concerned whenever they are told they have a low TNO score, but that in almost all cases simple reasons can be found to explain the effect. Typically, these are caused by the users not wearing their normal prescription glasses or contact lenses. This meant they were not able to resolve fine enough stereo detail to fuse the images, or were colour deficient in one or more of their RGB channels. Occasionally people can register no score with the TNO test. Yet if they persist with it, or return to it later to try again they can sometimes see enough depth information to record an average score. The two people who did not register a score on the TNO test were both aware that they were effectively blind in one eye and decided not to proceed any further with the study.

mixed 2D and 3D images, it was decided in Experiment 2 to test this finding using an inanimate object. This would ensure that human presence could not be a factor in the task of size estimation. It was also decided that validity of the convergence orthostereography used in all of the previous studies should be tested. Of particular concern was the reversed disparities found behind the centre of interest. In all of the previous stereo photography, there was a flat photographic background approximately 1 metre behind the centre of convergence where the models stood. It was comparatively easy therefore to switch ones gaze between the foreground and background to fuse the disparities, even though the disparities were reversed on the background. This is because the eyes can diverge, focus and fuse a limited amount of reversed disparities such as those found in an auto stereogram. However, if the background had been at optical infinity, the reversed disparities behind the object might appear to be unnaturally large and impossible to fuse. This is because increasing distances from the centre of convergence give increasingly wide disparities. With foreground disparities, these are limited to a maximum of 65mm as they cannot be wider than the stereo base. However, the disparities of the image planes behind the object plane (centre of convergence) can be much wider and will vary based on the camera to subject distance and the distance of the background from the subjects. If the background disparities are so wide they are impossible to fuse, this would reduce the validity of the photographic technique used to a very restricted range of subjects that had a background placed immediately behind the subject.

A further problem was realised when it came to taking photographs of an object 1.68 metres from the lenses that was placed in front of a distant landscape. A depth of field calculator indicated that the f1.8 50mm Olympus lens used for the swimsuit study would not be able to give a background at optical infinity that appeared to be as sharp as the foreground. This was because the minimum aperture of the Olympus lens is f16 and yet an aperture of f32 would be required to give the required depth of field. It was felt however that the slightly unsharp background given by this lens would be acceptable as it correlated well to the actual condition of direct perception. The eye, like a camera lens cannot focus on a foreground object while maintaining a sharp background. However the depth of the field of the eye is good, so we rarely notice that there is a difference in focusing between the foreground and objects at optical infinity, particularly as attention is usually concentrated on the foveal image where accommodation is optimised.

5.3.2. Experiment 2 Method

5.3.2.1 Stimuli and equipment. The criteria for the target object for this photography were easy to establish but hard to find. The object should have dimensions and shape similar to a human female torso. Singh (1993) has indicated that a healthy adult female has a preferred waist-hip ratio (WHR) of 0.7. By this, he meant that the waist is 70% of the hip size in women with the highest ratings of attractiveness and fertility. It was felt that the object should have a rotational symmetry and look circular in plan view. This was to ensure that the object did not remind viewers, subliminally or obviously that it looked like a human female. The waist measurements of the object should approximate the average adult female (approximately 60cms) and the waist area should at its slimmest point be about 250mm above the widest point. It should also be light coloured and have some surface texture to allow the visual system to easily focus on it and fuse its stereo details.

An arrangement of two large plant-pots (figure 5.2) was found to match the criteria in all respects for making the photographs of the experiment stimuli. A location for the photography was found on a vacant residential site near Liverpool University. An

Figure 5.3. The stimulus used in Experiment 2.

overcast day was chosen to take the photographs to ensure that the depth information in the photographs came from an evenly illuminated scene that had little modelling in it from directional sunlight. The same 100ASA Kodak film, height settings and convergences were used. The stereo camera was set-up in front of a low wall and the plant-pots placed on it and arranged into the waisted shape seen in figure 5.3. The cameras were set at the required 1.68m distance and the same lens convergence from the previous photography was used. Internal camera meters established that the correct exposure was 1/15th sec at f16. The aperture was kept constant as a number of exposures were made using different shutter speeds. "Bracketing" exposures in this way ensures that there would be enough transparencies made with varying image density (apparent image brightness) for a matched stereo pair.

The resulting stereoscopic transparencies were mounted into 35mm mounts and loaded into the stereoscopic projectors used in experiment 1. Despite the very large reversed disparities in the background, the stereoscopic image looked surprisingly natural. The author had no problem fusing the background image and all of those people who saw the projected image were impressed by its natural appearance. This would almost certainly not have been the case if the much wider than life-size (hypersterescopic) lens interaxials of conventional stereoscopy had been used. This is because it is possible for hyperstereoscopic and telephoto stereography to have almost no corresponding points in the two scenes imaged behind the foreground object, unless the lenses are set to a parallel configuration. Any such low correspondence is likely to give rivalrous stimuli in each of the binocular channels and an unnatural flickering appearance to the stereograph.

The projected stereo photograph of the plant pots was measured from the screen and its proportions matched the original object exactly. And as with the other stereographs in this study, when one light path was interrupted to switch the display from 3D to 2D, the previously observed flattening and fattening effect was observed. The relative slimming of the "waist" area in the stereo presentation seemed to be particularly strong. The effect was strong enough to consider using a simple forced-choice paradigm in a size estimations task, rather than use the 2D/3D projected images as the stimuli for an experiment.

In experiment 2, the stereoscopic stimuli would again be copied to make a same-size synoptic pair using the specialized transparency copying technique previously described in section 5.2.2. The stereoscopic and synoptic pairs could then be loaded into each of two Pinsharp™ hand-held slide viewers for simple side-by-side size comparisons.

The original improvised stereo viewer made from two GEPE slide viewers taped together was replaced by a conceptually similar Pinsharp™ 3D viewer (figure 5.4).

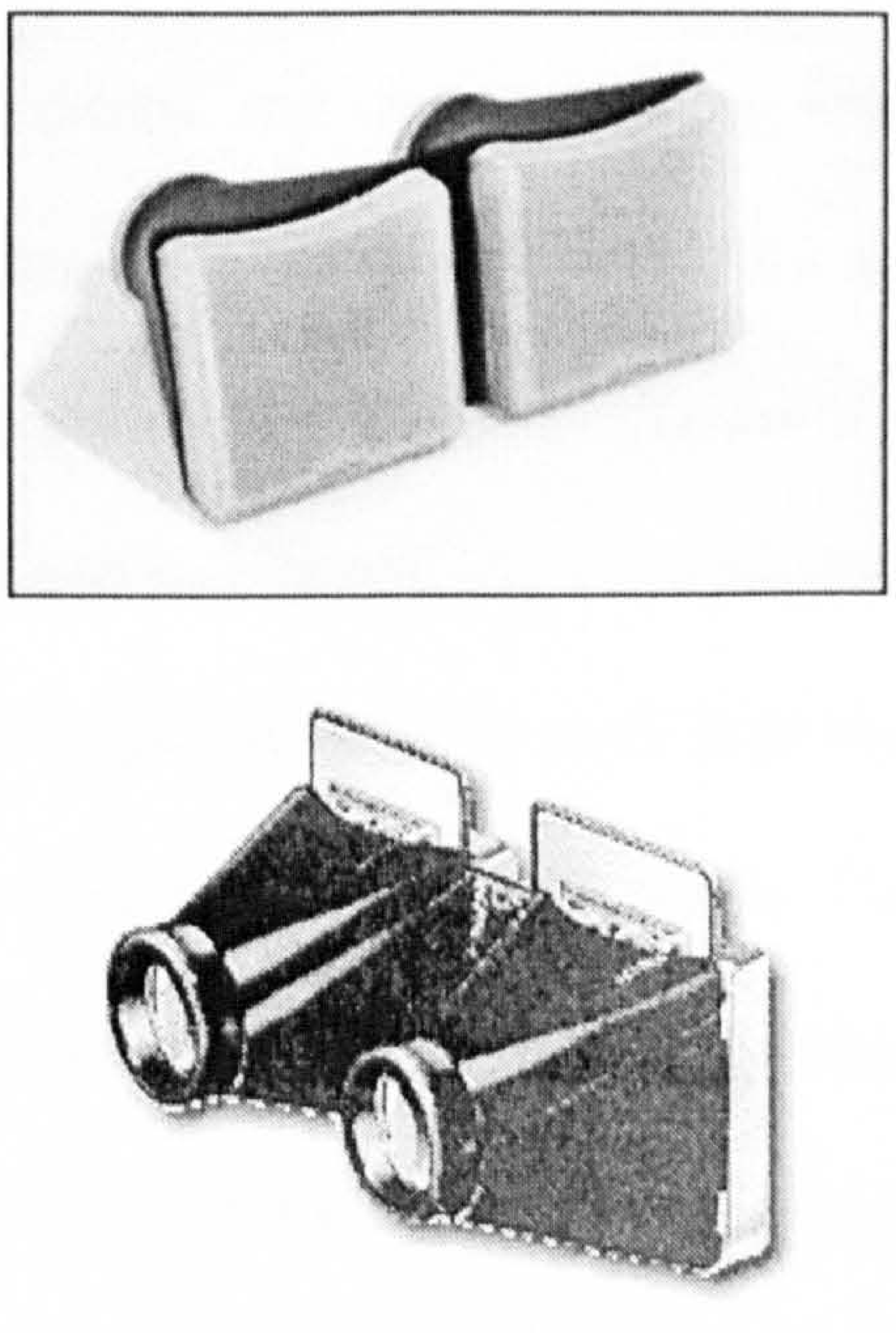


Figure 5.4 Front and rear views of the Pinsharp™ 35mm stereoscopic transparency viewer.

This device however was fitted with a pair of highly-corrected glass meniscus lenses rather than the plastic lenses in the GEPE viewers. The image size in the Pinsharp was larger than the original GEPE slide viewer (estimated at 80% of life-size) when the images were compared with the real stimuli at the same viewing distance. The Pinsharp™ viewer also has a flexible joint between the two slide holders to allow the users to adjust the inter-pupilary distance (IPD) or the convergence for images made with convergence stereography. It was found that daylight was not a reliable light source to rear-illuminate the transparencies in each of the viewers. So a 25-watt tungsten-halogen table lamp was used as a light source and the viewer to lamp distance was standardised at 50cm.

5.3.2.2 Participants. Twenty Liverpool University undergraduate participants were tested individually

5.3.2.3 Procedure. Each of the twenty participants were shown two sample stereographs that had been pre-loaded into two Pinsharp™ viewers. This was to accustom them to using the device and viewing stereographs. It also ensured that the first time they saw a stimulus using the viewer they were not asked to give responses that would be used in the experiment. They were shown how to adjust the IPD using a simple single-hand method and how to quickly view two consecutive stereo views that were held in the left and right hand. They were then shown two other viewers. One (marked A) carried the stereo pair of the plant pots. The other (marked B) carried the synoptic pair. The participants were asked to look at the whole objects in each viewer and especially to look at the waist areas. They were encouraged to switch between the two images a number of times and to judge their sizes at their own pace. They were then asked to indicate if any of the images were clearly smaller or larger than the other in any dimension. The participants were encouraged to report a size difference, even if they considered the task too difficult. The specific question asked was "Compare the size of the object in the first viewer with the size of the object in

the second viewer and indicate which viewer carries the image with the larger object."
For those who could not see a size difference between the stimuli, a no difference response was recorded. All of the responses were recorded by the experimenter on a printed response.

5.3.3 Results

The results shown in figure 5.5 confirm the prediction that the waisted object was viewed as slimmer or smaller in the stereo presentation ($\chi^2(2, N = 20) = 13.3, p < 0.001$). Almost three times as many viewers saw the object as slimmer or smaller when viewed binocularly compared to the synoptic image. And 95% of all responses gave different size estimates of targets that subtended the same visual angle and were at the same distance from the observer. Only one participant in the study saw the stimuli as identical in size in both conditions.

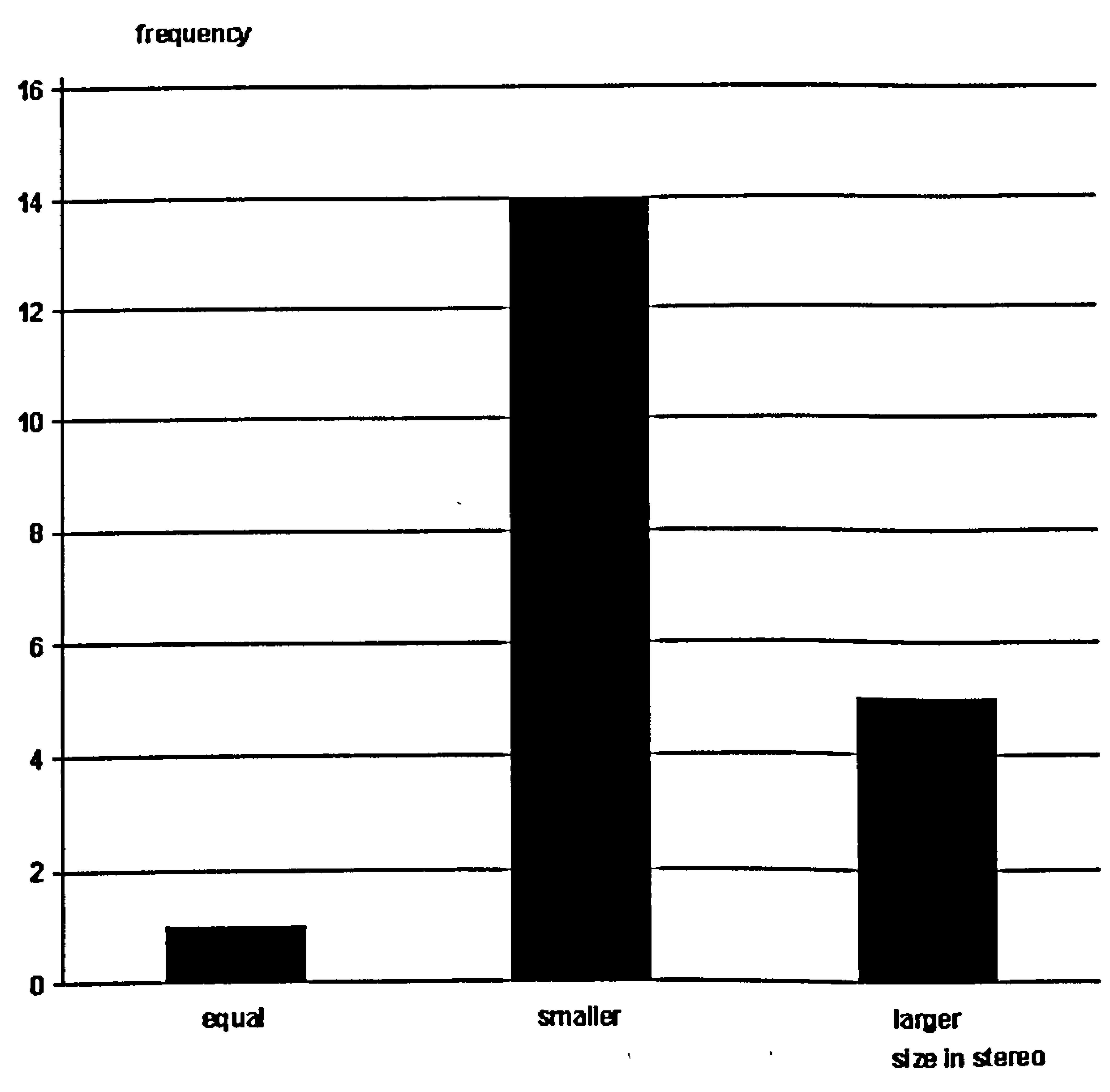


Figure 5.5. Size comparisons of the stimuli in Experiment 2 in synoptic and stereoscopic conditions.

5.3.4. Discussion

The experiment 1 data supported the subjective impression that people looked slimmer in stereoscopic images. The data from experiment 2 confirmed that the slimming effect seen in experiment 1 was not caused simply by participants acting more favourably to a more realistic human presence and was robust enough to be seen with a non-human target. It also suggested that the computer generated peanut shape used in the first CGI pilot study was seen as slimmer because it was a stereoscopic rendering, and not because the background design drove participants to make slimmer size estimations with stereoscopic projection. *A research question that had yet to be addressed in any experiment so far was to establish which type of stimulus (2D or stereoscopic) gave the most accurate representation of the stimulus' actual measured size.* Experiments 3 and 4 address this question. They also look at the effect of varying the disparity in the stereo pairs to test the hypothesis that the slimming/fattening effect is related to the amount of disparity.

5.4 Experiment 3: The First Peanut Experiment

5.4.1 Introduction

The first virtual object pilot study (peanut shape study, section 3.3.2) indicated that a waisted cylinder peanut-shaped object gave a strong impression of slimming in both of its stereoscopic conditions when compared to the otherwise identical synoptic 2D object presentations. However the technique for generating the images was inadequate for the task in that one could not establish any "real-world" equivalent measures or values for the stimuli. If it had been possible to simulate photography of a "life-size" peanut with advanced CGI software, then it might be possible to establish which method of imaging (2D or stereoscopic) gave the most life-like reproduction of the real size of the object. ... The only reliable measure derivable from this technique was the convergence angle of the virtual cameras in degrees. A core failure with the technique used was that the virtual camera lenses were set to converge on the central axis of the peanut. This is only possible in reality if the object is transparent. As the virtual stimuli was intended to be opaque, this was another confound in the original study. Also, it was noted that a lighting artefact that made the object appear to be slightly smaller on the stereo presentations due to incorrect rendering of the highlight and shadow areas meant that the methodology of the first peanut study was clearly flawed. These factors, coupled to the poor overall control with the software meant that this method would require major revisions before it could be used in another experiment. The method of presentation too would require modification as the viewer positions and projector alignments were far from ideal. The card-matching aspect of the first study too required some revision, as it was realised that the participants were given only one opportunity to make a size estimation in each condition. In subsequent experiments, each stimuli was presented twice and in reverse to counter-balance any order effects in the presentation..

The concept in this experiment is to present various types of stereoscopic and monoscopic stimuli and asked the participants to compare the shape of the object on the screen to the object on the various cards in front of them. The shape of the object on each of the cards would vary in slimness, with only one card actually matching the shape of the projected object. A set of varying waist-hip ratio peanuts would therefore need to be rendered and printed on large cards: their size would be controlled so when held up by the participants, their image should subtend the same visual angle as the large projected image which is further away. The projected object would be rendered only with the 0.7 waist-hip ratio shape. Therefore any change in perceived shape, as matched to the varied peanut shapes on each of the cards would be due to the stereoscopic content affecting the perceived shape of the peanut.

5.4.2 Method

5.4.2.1 The Revised Virtual Stimulus. For the third experiment, a computer-generated virtual peanut object was designed with the same imaging geometry as Experiments 1 & 2 (see Figures 5.6 to 5.8). The overall design aims with these stimuli was to produce a monochrome image that looked like a real object photographed using the existing camera system. The virtual object should have a shape that has human qualities, such as slim and broad areas that could correlate to the waist and hips or jaw and neck of a person. The virtual object would also need to be rendered in a larger number of stereoscopic convergences than was used previously to investigate the relationship between size judgements and degree of disparity. In addition, multiple 2D renderings of the revised object would also be required rather than the single matching card used previously.

The "non-human" peanut shape was again used for the revised stimuli, so that the viewers would see it simply as a large double-curvature 3 dimensional object. It therefore required very accurate design and rendering so that a real object could be built from scratch in reality to match the virtual stimuli if it were ever required. As

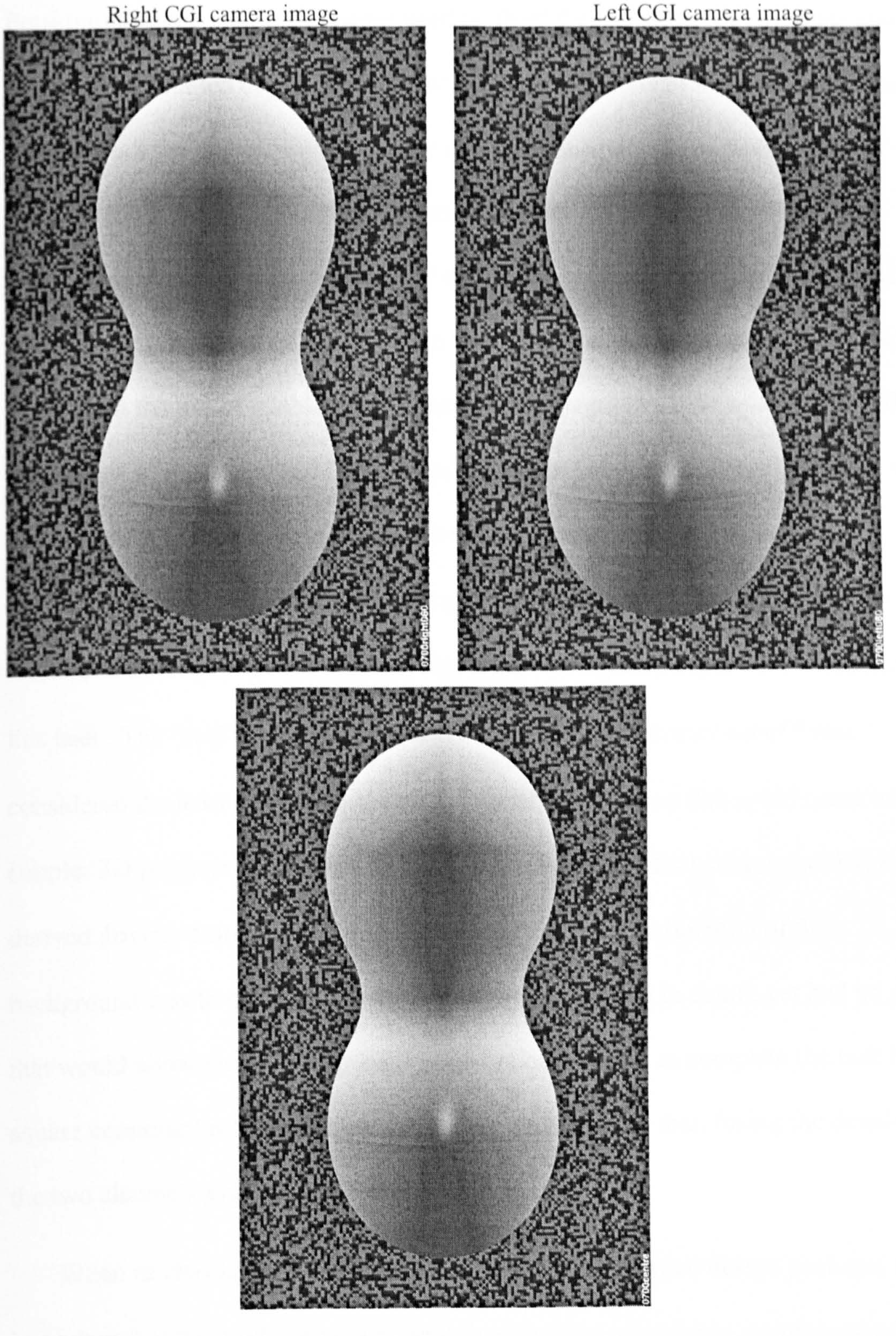


Figure 5.6. The 65 mm separation stereo pair and the centre camera image used in experiment 3.

this task was beyond the expertise of the previous programmer. The computer aided design lecturer from Liverpool University Department of Architecture (Mr Philip

Berridge) designed the CGI images used in all of the following experiments and pilot studies. The detailed design for the new stimulus was derived from what had been learned from the previous peanut study and the plant-pot stereographs. The 0.7 waist-hip ratio of the previous plant-pot shape for the overall shape of the object was retained, as the small waist hip ratio of the first virtual peanut had proportions (0.5 waist-to-hip ratio) that did not correlate to any natural human proportions. It was also agreed that the use of visible wire-frames in the images could be misleading and that the surface of the object and background should be rendered only with lighting and surface texture. The construction of the stimulus was completed using the architectural computer aided design program StrataVision™ 3D 4.0 from Strata Inc as it featured the required sophisticated rendering and lighting capabilities necessary for this task . The "real world" image quality available with StrataVision™ was considered the least likely to generate the variable pixilation that could occur with simpler 3D programs. It could also incorporate a random dot background that was derived directly from stock Adobe PhotoShop™ files. The function of this background was to define a flat virtual plane that was rich in detail, yet had no pattern that would allow viewers suppressing one optical channel to complete the task by square counting (as was possible in the first study), rather than fusing the details from the two channels available optical channels.

When rendering stereo disparities using a computer aided design package, it is important that the model is very accurately described using large uncompressed data files. Small changes in topography or brightness due to aliasing can alter the stereoscopic detail within the image. This is because the program generates pixels that correlate in size to the final size of the data file. The less data used to describe a surface detail, the more likely it becomes that points in virtual space will be become larger and move relative to the original data points when the file is rendered and transferred to film for projection. A typical stereoscopic artifact caused by data and rendering effects could turn a surface detail in half-shadow to black pixels in one rendering (a left channel image for instance) and white in another. This would lead to retinal rivalry as one tries to view this part of the image's surface, only to find that it has brightness differences that do not allow stereoscopic fusion to take place. Another artifact observed during the design phase was that some outlines had a saw tooth or stepped appearance rather than the continuous/smooth appearance of real objects that have been conventionally photographed. This problem can never be completely resolved, as every digital image is made from square pixels that will have a saw tooth effect on oblique lines if the image is enlarged enough, or has low enough resolution to see the effect at normal magnifications. This was avoided in all of the experimental images by raising the resolution at which each object was rendered and saving all the renderings as PICT ™ files that retain detail without compression artefacts.

The widest part of the peanut stimulus is described in these experiments as the "hips." The narrowest is the "waist." The upper section is sometimes referred to as the "shoulders," though the object shape has inverse and rotational symmetry, so this area is identical to the hip region in all of it dimensions. <u>All of the stereoscopic and synoptic peanut images in the experiments are of this 0.7 ratio. Only the 2D matching images are ever shown with variable physical measurements of their waist-hip ratio.</u> Originally, it had been planned to cover the surface of the object with the same random-dot rendering that was used for the flat background. However, this would have taken a greater amount of computing power and rendering time than was available. So, its surface was rendered without texture and the only stereo information available to the viewer was from lighting derived contour and shading and the trapezoidal distortion (perspective keystoning) of the background.

The lighting design of the CGI stimuli required some experimentation. StrataVision™ has a full range of virtual lighting systems that allows the programmer to simulate sunlight, artificial lighting and mixtures of both types of illumination. The style of illumination used previously in the stereoscopic photography of people and objects is often described as "flat" lighting by photographers. This means that the lighting is diffuse with little directional content. This gives very soft shadows (or sometimes no shadows at all), a reduction in texture depth and gradual reduction in brightness from the near scene to the background. Point source lighting (such as spotlights) give very strong shadows and can convey strong shape, depth and texture with cues that are monocular. These lights also fall under the constraints of the inverse square law, which says that with a point source of light, the intensity at the surface is inversely proportional to square of its distance to the source; e.g. doubling the distance reduces the intensity of the illumination by a factor of four. In the original stereoscopic photography this type of lighting would have given undesirably strong cues to object shape and texture in the foreground, independently of the stereoscopic depth information. It would also have excessive fall-off, making the background very dark (requiring extra lighting that was not available) and greatly reducing the amount of detail that could be seen or rendered on the film. With flat lighting, most of the depth information would therefore be conveyed by the differences, or disparities between the stereo channels and the naturally receding perspective of the 50mm lenses rather than the conventional 2D cues.

The swimsuit photography and the plant pot photography used flat lighting techniques that allowed the viewers visual systems to fuse the fine details in the image to recover the scenes depth information. The flat lighting in the swimsuit photography was directed at the subject from beside the camera at a level just above the lens height. The plant pot photography had flat illumination from "top lighting" that was inevitable from daylight with an overcast sky. However, the peanut design and rendering processes were very expensive in terms of computer power and time. The plan to render the peanut with a random dot textured surface, similar in character to random dot background was rejected as it would take too long to render all of the possible versions of the stimuli with realistic surface details. This would mean that the object had little fine detail on its surface (other than the joins between the sections of the peanut shape[3]) from which the visual system could derive a similar amount of stereoscopic information. Yet the previous stimuli had large amounts of surface detail stereoscopic information that would be missing from these renders.

It would have been straightforward to use virtual spotlights to give the peanut a distinctive 2D shape with strong shadows. These shadows would have given strong stereoscopic cues to the objects shape due to the large differences this would give between the left and right camera angles. However, this lighting would not correlate in any way with what had been used in the real images. Various types of flat lighting combinations were tried in order to give the peanut shape the required stereoscopic details without diverging away from the principles used in the previous stereoscopic images. A large diffuse light source was placed the equivalent of 10 metres above the top of the peanut shape. This approximated the even lighting given by an overcast sky and gave the shape the subtle cues required to reveal its overall shape. However, this

[3] Wire frames with simple rotational symmetry were used to construct the upper and lower domes of the peanut shape. The waisted area used a different method of generating wire frames that was derived from a series of algorithms used to construct the variable waist shapes for the 2D matching cards. The two domes and waist area were joined together and a surface rendering program was used to make the surface appear smooth. The joins between the three sections were visible as very fine lines that were barely visible, yet followed the precise stereoscopic shape of the object.

would put the waist area in shadow and put the centre of attention (and expected area of change) in the shadow from the shoulder region and make it the darkest part of the object. Two diffused floodlights were chosen from the tool box and placed on either side of the virtual lens to fill-in the shadow area at the waist to about 70% of the brightness of the highlight areas. A final spotlight was directed at the centre of the object. This was to give some additional stereoscopic definition to the front of the object and it brightness was set at about 10% brighter than any other area in the rendered scene.

The geometry and lighting for the centre camera image were established before rendering the 65mm stereo pair that represents the standard human interocular distance (see figure 5.6). However, if two camera angles 32.5 mm on either side of the centre camera position are used, then the images from those positions will be further away from the object. This would mean that the object would get smaller in the stereo renders and this would undermine the basic premise of the experiment, that images differing only in disparity can induce different size estimations. To keep the size of the object constant in all the renders, a curved path was programmed to ensure that the object magnification remained constant in all of the stereo pair renders (see fig 5.7).

In total, eleven sets of stereo pairs and one synoptic pair were rendered, as it was predicted that increasingly wider disparities would induce a dose-response curve of diminishing size estimations. The pairs were of varying camera lens separations measured in millimetres, from zero disparity (two images from the central camera position to make a synoptic pair with no disparity between them) called 00, out to a stereo pair with 300mm of disparity. The full list of the first pairs of renders are: 00mm, 30mm, 60mm, 65mm, 90mm, 120mm, 15mm, 180mm, 210mm, 240mm,

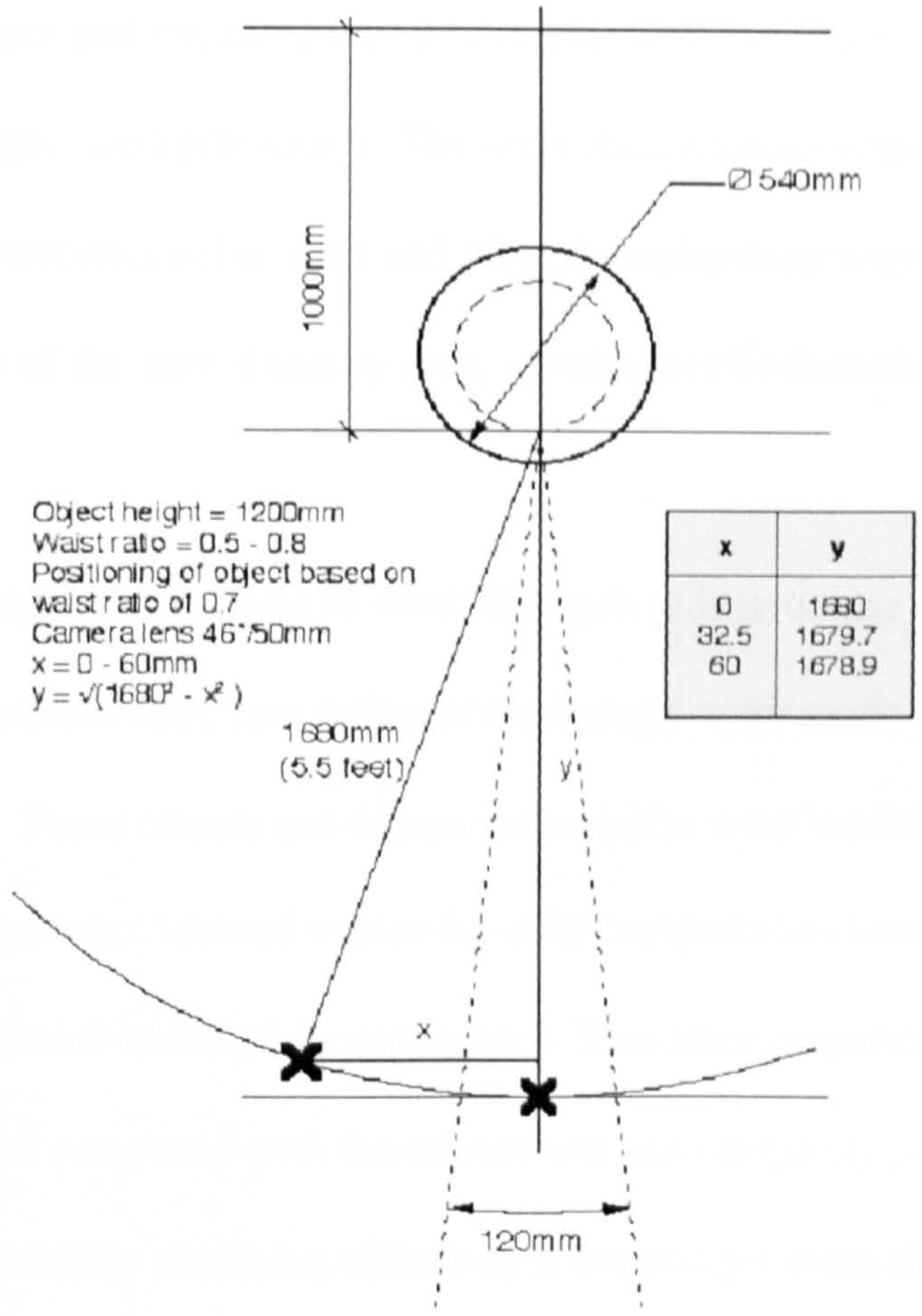


Figure 5.7. Plan view of the dimensions of the peanut shape used in Experiment 3, its relationship to the virtual camera positions and the plane of the background. The virtual cameras generated views at each disparity in an arc to ensure that the magnification was constant in every image. The right hand bold X shows the position of the camera when it was in the straight-ahead position (zero disparity). The left hand bold X shows the position of the left-hand camera at a distance x from the straight-ahead position. The disparity this generates is defined as 2x mm.

270mm, and 300m. An additional condition named 65P was added to these renders. This condition was used to reproduce the parallel axis (non-convergent) lens axis images used in conventional stereographs. There were two reasons for using this type of image. The first was to investigate if the conventional parallel technique was perceptually different from the more ecologically valid technique of convergence stereography. Additionally, this technique does not produce reversed disparities behind the centre of attention and may have provided an alternative stimulus for use in the experiments if the convergence technique was difficult to fuse by some of the participants. The pairs were transferred to Kodak 64 ASA slide copying film using a

Lasergraphics printer and mounted into 35mm glassless mounts with the same registration technique used previously. The were then marked with a code to denote if they were right or left channel images and what disparity they were (L65, R65 etc), with the exception of the zero disparity pair, which were both marked as 00 as they were identical.

Since people might be sensitive to small changes in interocular separation around the natural IPD measurement, two different renderings were made to test the validity of this hypothesis. These 60mm and 65mm stereo pairs were loaded into two Pinsharp viewers. The images were viewed in side-by-side comparisons and found to be very easy to view and almost identical in appearance. The other disparities were loaded into the viewers and compared with the 65mm and zero disparity pairs to see if there were any obvious relative slimming effects or if the images were difficult to fuse. The most obvious variation between the stereo pairs was that different disparities give varying object-to-background distances. The 60/65mm pairs showed the background at an apparent distance of about 0.75 metres behind the object. This was disappointing, as it should have appeared to be at 1.0 metres. It was hoped that this was an artefact of the smaller magnification found with the Pinsharp viewer and that the distance would be correct when the images were projected orthostereoscopically.

The 30mm pair conveyed the background distance as just behind the object, while the wider disparities increased the perceived background distance in the predicted way until the images were too difficult to fuse at around the 270mm distance. The object shape in these renders appeared to be surprisingly constant and not obviously slimmer with the wider disparities. They all however appeared to be slimmer than the synoptic zero disparity pair.

5.4.2.2 Design of comparison images

Additionally, 13 varying waist-hip ratio 2D centre camera renders were made. Only one of these renders was an identical shape to the 0.7 waist-hip ratio of the synoptic/stereoscopic targets, and this was image number nine in the sequence in the sequence of waist-hip ratios from 0.5 to 0.8. These renders used a dedicated algorithm to calculate the curvature of the object in such a way that each waist diameter differed by exactly 0.025 of WHR from the next image in the sequence (figure 5.7).

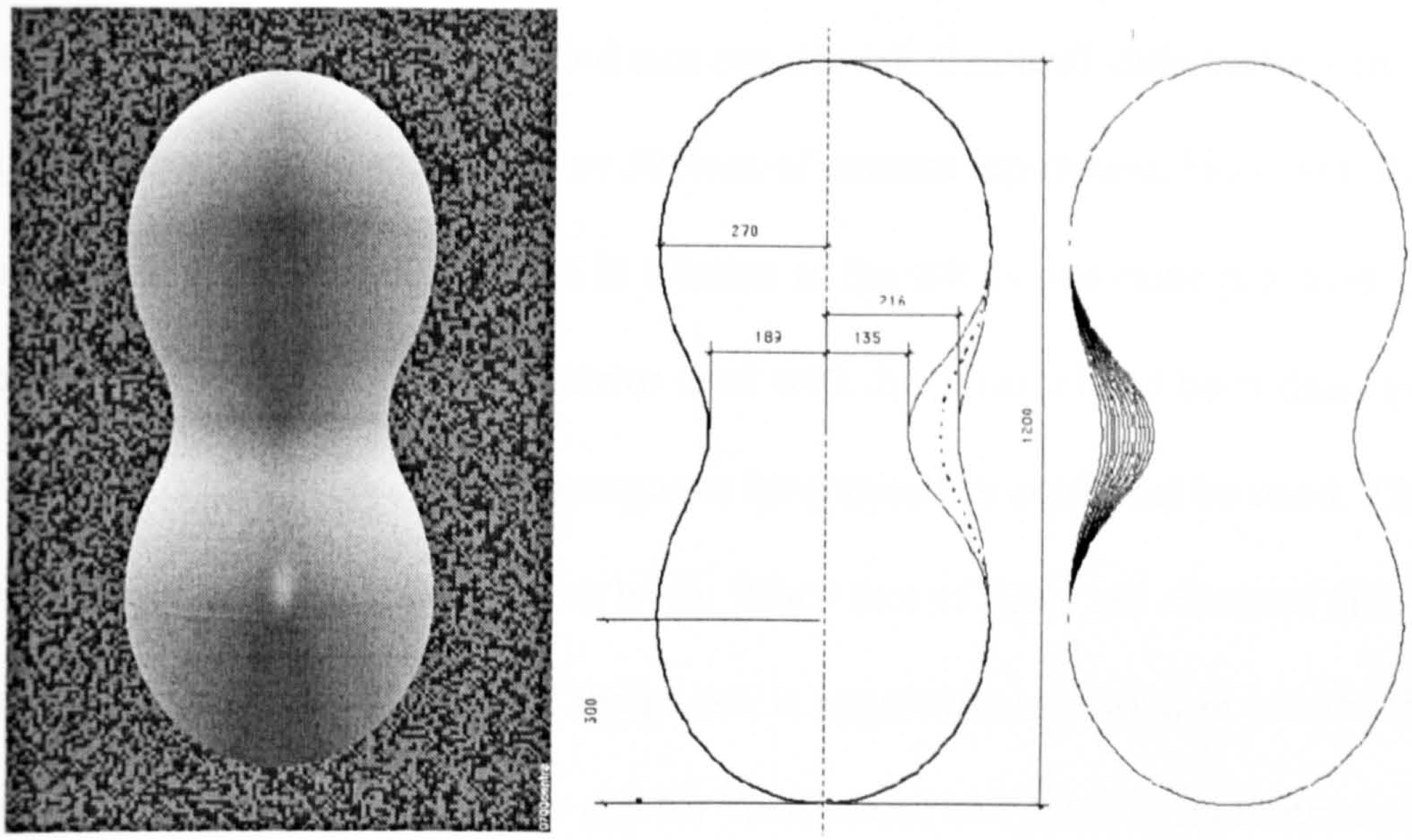


Figure 5.8. The left image is the centre camera render of the 0.7 WHR peanut. This was modified to have varying waist waist-hip ratios for the thirteen comparison stimuli used in Experiment 3. Each stimulus was printed onto A4 card . Card 1 had the slimmest waist-hip ratio of 0.5. Each of the subsequent cards had a ratio that increased in 0.025 graduations. Card 9 was the same 0.7 ratio as the stereo and synoptic images. Card 13 was at a ratio of 0.8. The central diagram shows the largest and smallest physical dimensions of the varying waist sizes. The diagram on the right shows all of the intermediate ratios. The card images were scaled so that they were approximately the same-size as the projected image when held at arms length.

This algorithm was specific to these "vari-waist" images and could not be used to make renders for the stereoscopic disparities. This meant that care had to be taken to ensure that the vari-waist images were identical to the stereo renders in all respects except disparity. The images were printed out onto A4 cards that were laminated to project the surface of the images during the expected degree of handling. It had been estimated that an image correctly sized on standard photo cards would subtend a visual angle similar to the "real" object if held comfortably at arms length. This was also the same visual angle for the peanut when it was projected to life-size on to a projection screen from the revised othostereoscopic projection system described in chapter four, figure 4 b.

5.4.2.3 Projecting images with varying disparities.

The 13 pairs of images were loaded into two Kodak Carousel slide carriers in sequence from zero disparity out to 300mm of camera separation. However, the right channel images were not reversed in relation to the left as was done previously. The reason was the surface-silvered mirror used with this channel had been damaged and the correct orthostereoscopic convergence for projection could not be used. The Kodak Ektar projectors are 330mm wide. When two of them are mounted side-by-side the minimum lens interaxial separation is approximately 400mm (due to the cooling fan vents requiring an air gap for ventilation) and leads to an interaxial distance that is almost four times too wide for the task. When the renders were projected using this arrangement, curvilinear distortion was visible on almost all of the images. The synoptic pair displayed a slight barrel distortion in that the edges of the background appeared to curve away like the front of a beer barrel. The synoptic condition gave some appearance of being stereoscopic, but in this case the depth seemed to be created by the vertical misregistration of the edges of the images that looked like the vertical disparities always found in convergent stereoscopic images. The 30mm renders gave a background that was clearly behind the object and quite flat in appearance. The 60 and 65mm stereographs showed classic pincushion distortion on the background where the curvilinear distortion was in the opposite direction to the synoptic pair) and this effect increased with the other disparities making them unfusable beyond 200mm of disparity.

Despite the problems with curvilinear distortions, the initial examination of the images had showed bright and detailed synoptic and stereographic images that were centrally aligned on the screen. This suggested that the programming matched the original image capture in the most important respects. The peanut height was found to be very close to the predicted height of 1200mm and the screen had to moved backwards a short distance to ensure the magnification of the image was correct. The peanut images were then measured and found to be the equivalent of life-size, with the background subtending 31.6° wide by 47.0° high and the peanut subtending 18.6° wide by 39.3° high from the viewing position 1.68metres away from the screen. Its waist subtended a visual angle of 13.0 ° in all of the projected images.

When a replacement surface-silvered mirror was sourced, the projectors were returned to their orthostereoscopic positions. As expected, the backgrounds of the synoptic, 30, 60,65, 65P and 90mm projected pairs were now perceived as flat when viewed using the correct projection geometry: Curvature of field was clearly evident however in all of the renders above 120mm disparity. It was also noted that the background plane behind the object was now approximately one metre behind the convergence point at the waist of the object in the 60 and 65mm conditions. This was exactly as it was designed to be. Subsequent distance perception tests between real targets 1 metre behind (and to the side) of the screen confirmed that the background was perceived at approximately this distance during the regular preview demonstrations.

It was decided that all of the renders above 120mm were unusable, as they had field-curvature artefacts from the projection geometry that made them appear to be unnaturally distorted. The 120mm condition was interesting in that it appeared to have the beginnings of slight curvature of field to the author. However, other viewers thought it was as flat as the 65mm condition that the projection geometry was designed to be compatible with. The 30mm condition appeared to be almost identical to the 65mm condition, with the exception of the background distance being perceived as closer to the peanut. It was concluded that if the number of disparities presented during the experiment needed to be reduced, this condition could be removed without adversely affecting the data or outcome.

Though the author could fuse on the waist area of the object with all of the disparities, the background area was almost impossible to fuse with the 270mm and 300mm stereo pairs. An interesting aspect of this effect is that it was expected that most of the images would cause difficulty in fusing the background details. The reason is that convergence stereography always produces reversed disparities in the background. This means that right eye objects are reproduced to the left of the image centre line, and vice versa. The eyes have to diverge a little in order to fuse reversed disparities like this, and it was expected that it would be very difficult for most viewers. Unnatural eye vergences such are thought to be a cause discomfort (Lipton 1982 page 12), however most viewers had no problems fusing reversed disparities provided they were from images generates at or below 65mm of stereoscopic disparity. So participants were tested using the 00, 30, 60,65 and 120 mm disparities.

5.4.2.4 Participants. Twenty Liverpool University undergraduates were tested individually.

5.4.2.5 Procedure. The thirteen comparison cards were randomized and the participants asked to place them in order from slimmest waist to fattest waist in order to familiarise themselves with the stimuli. They were then shown the first projected image of the sequence of varying disparity images and asked to pick a card that matched the shape of the peanut as it appeared on the screen. This was repeated with the remaining three images. However in the parallel condition (65P presentation), the projectors were de-converged to give the images the correct separation required for parallel stereography.

5.4.3 Results

Figure 5.9 shows the overall group means for these choices. Figures 5.10-5.13 show the frequency distributions of participants' matches for the four different disparities. A one-factor (disparity) ANOVA found an overall effect of disparity on size judgement ($F(2,38) = 7.628, p = 0.002$). Post-hoc paired comparisons showed that the only significant differences were between the 0^0 (synoptic) and the 65^0 (stereo) ($t(19) = 3.367, p = 0.003$, two-tailed) and between the 0^0 (synoptic) and the 120^0 (stereo) ($t(19) = 3.286, p = 0.004$, two-tailed). The 65P condition was not significantly slimmer than the 0 condition.

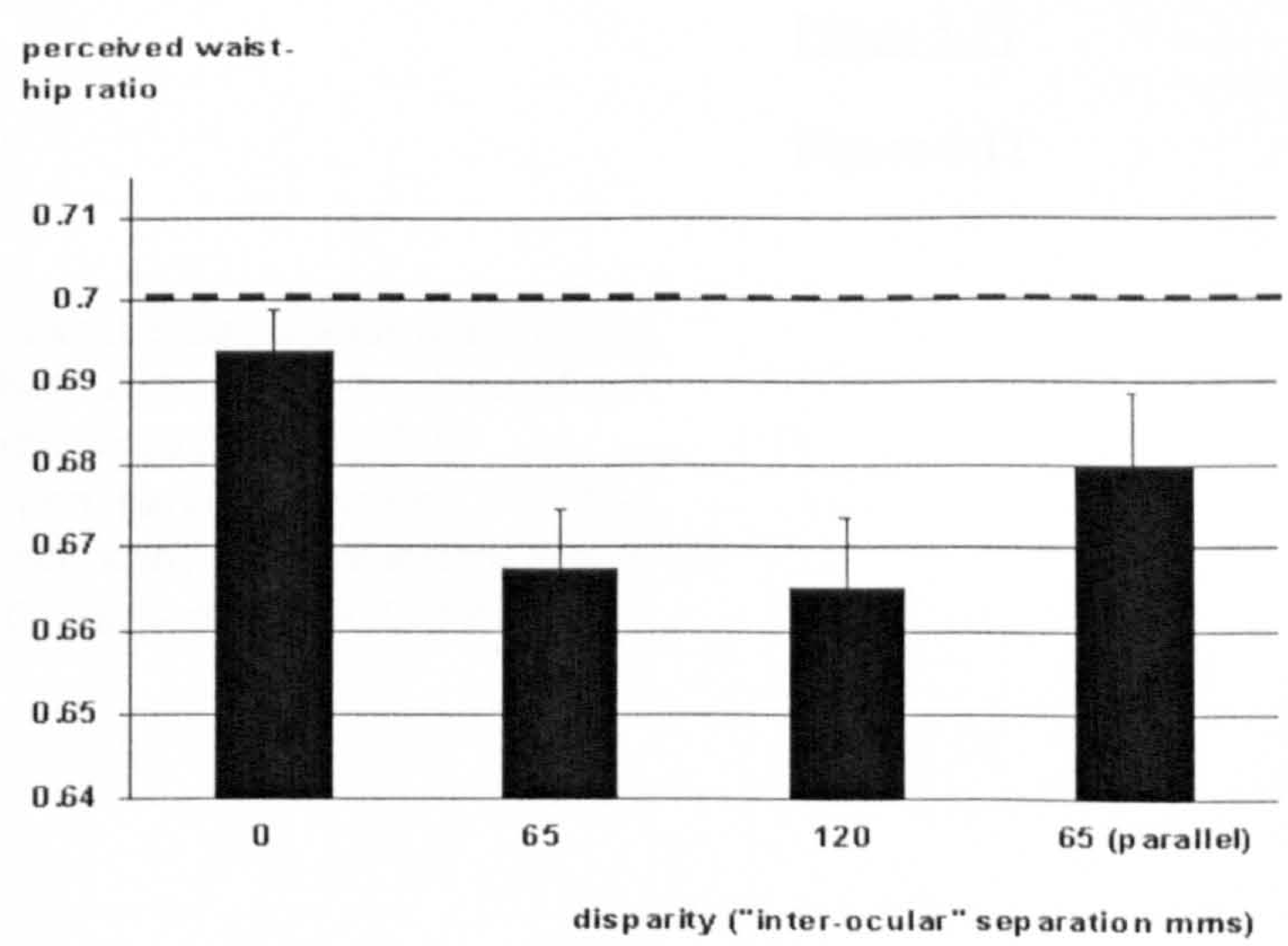
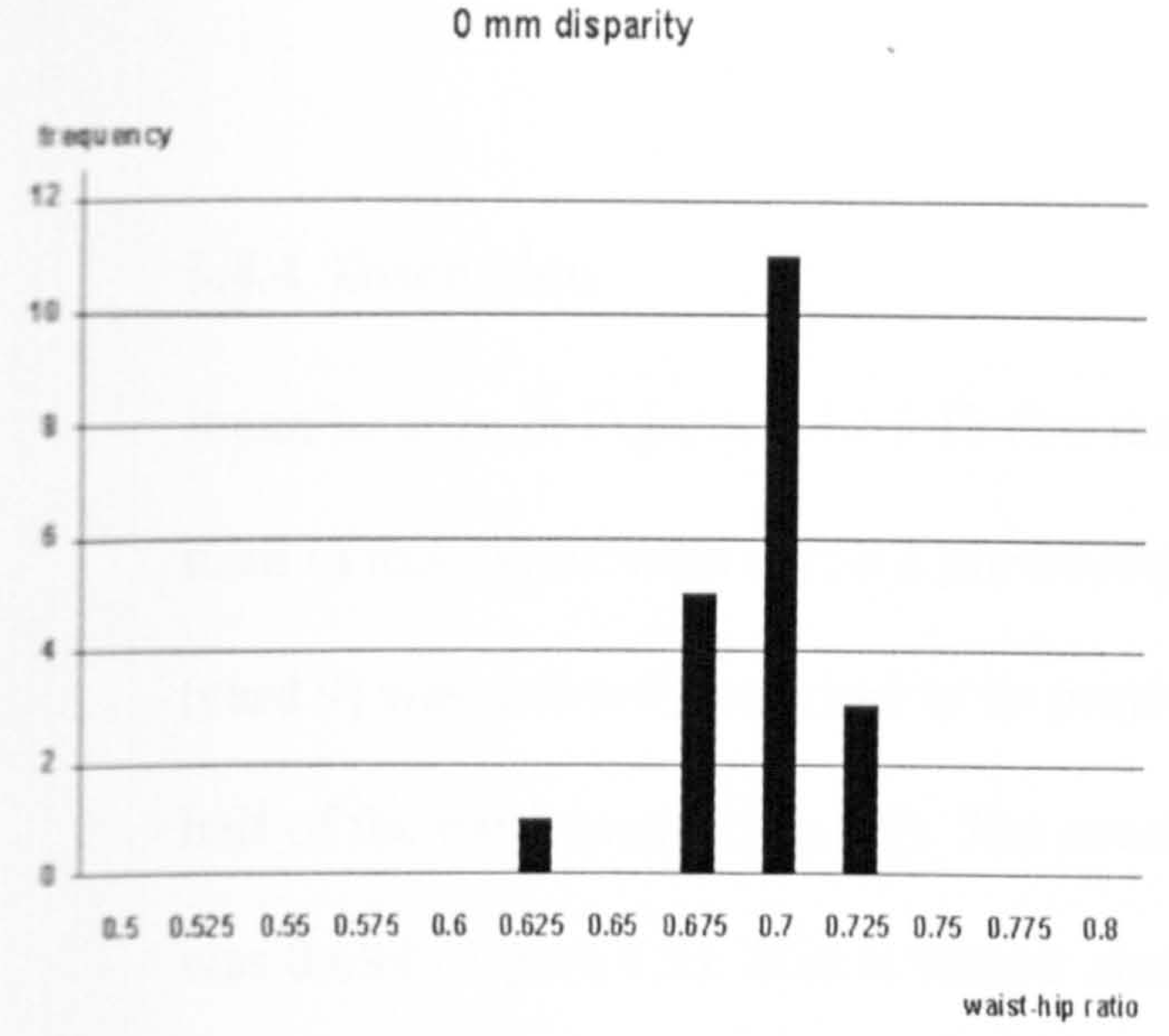


Figure 5.9. Perceived waist size when the object was projected in the four different disparities used in Experiment 3. The dashed line shows the actual waist-hip ratio of the stimulus.



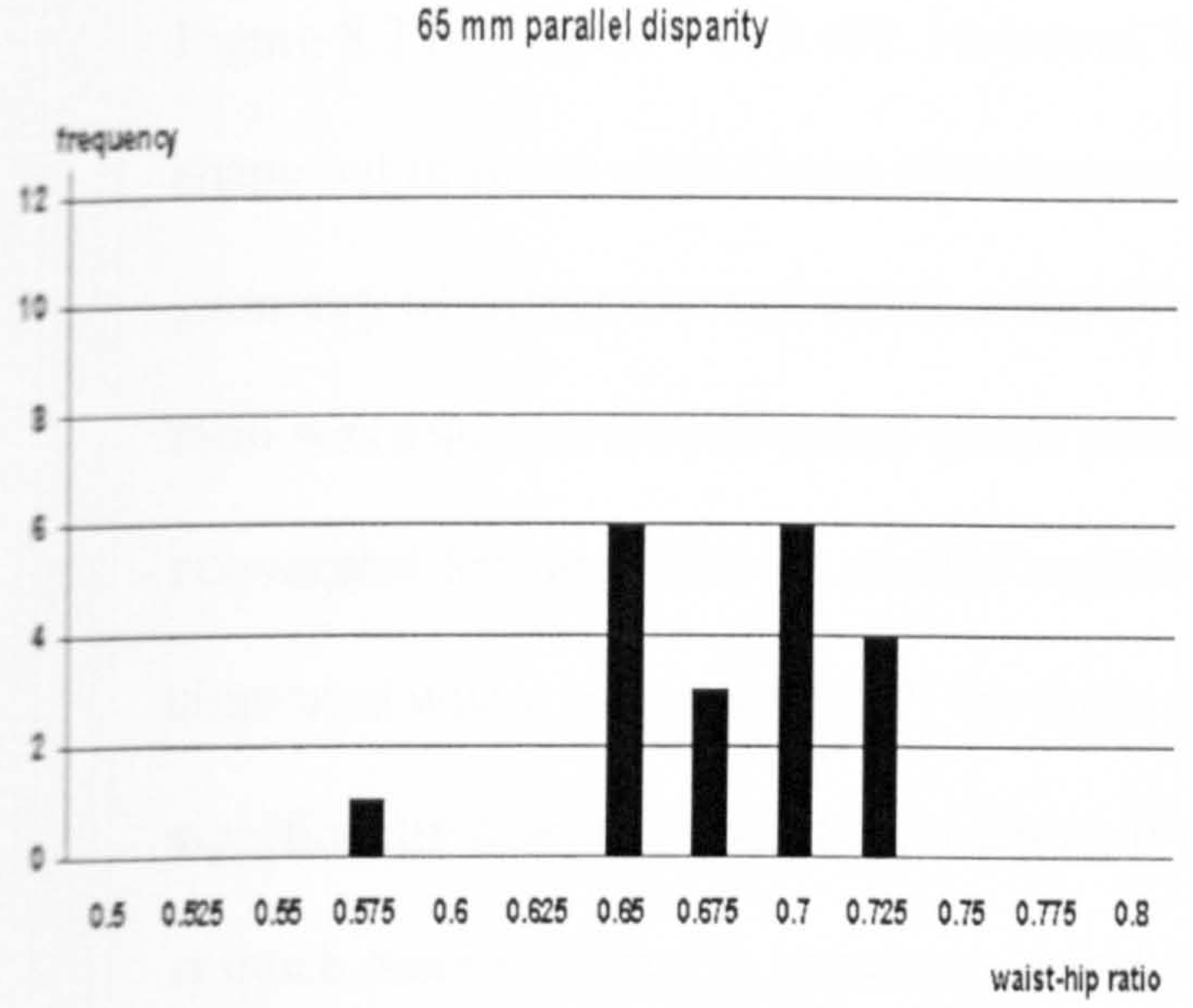
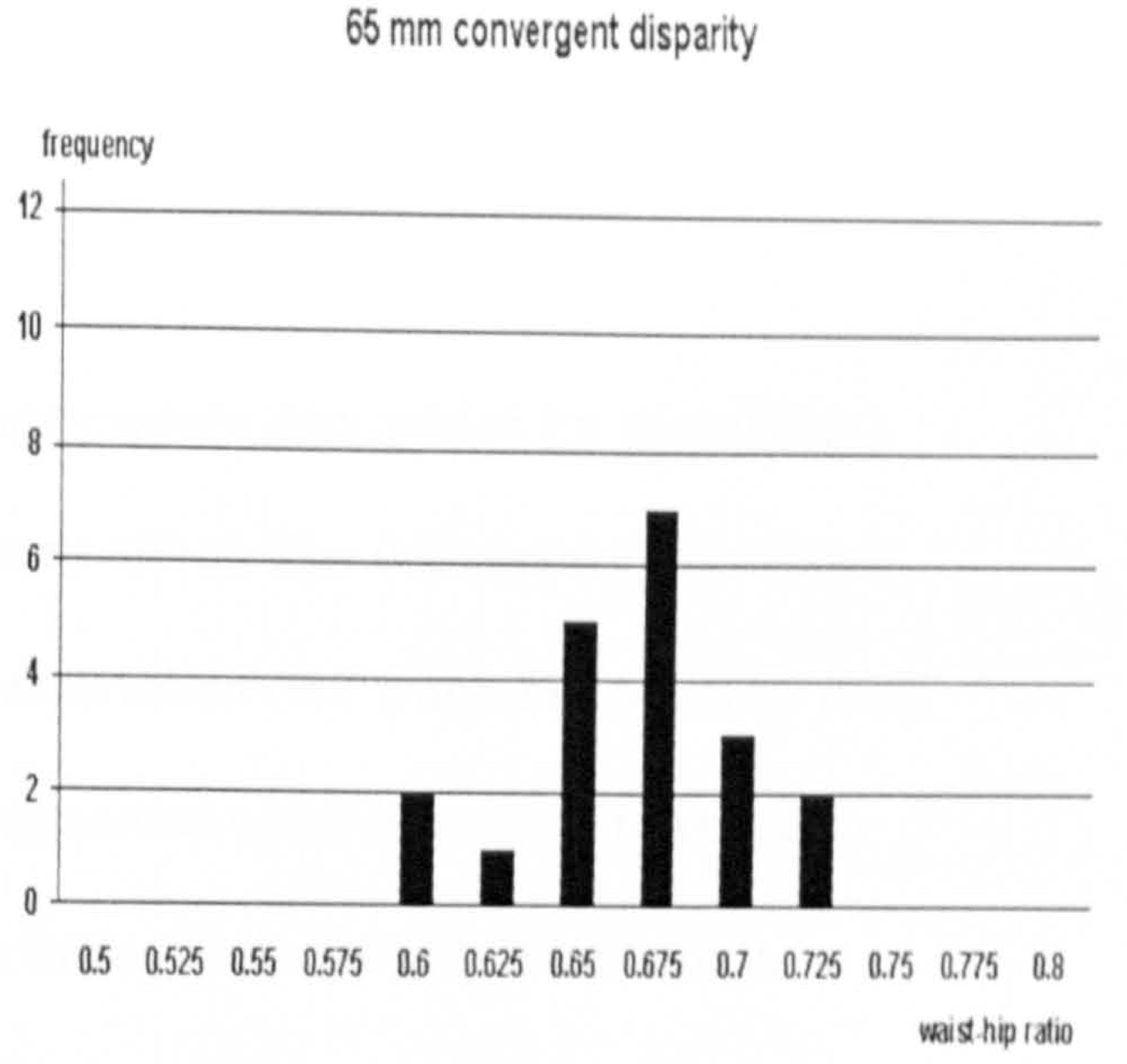


Figure 5.10

Figure 5.11



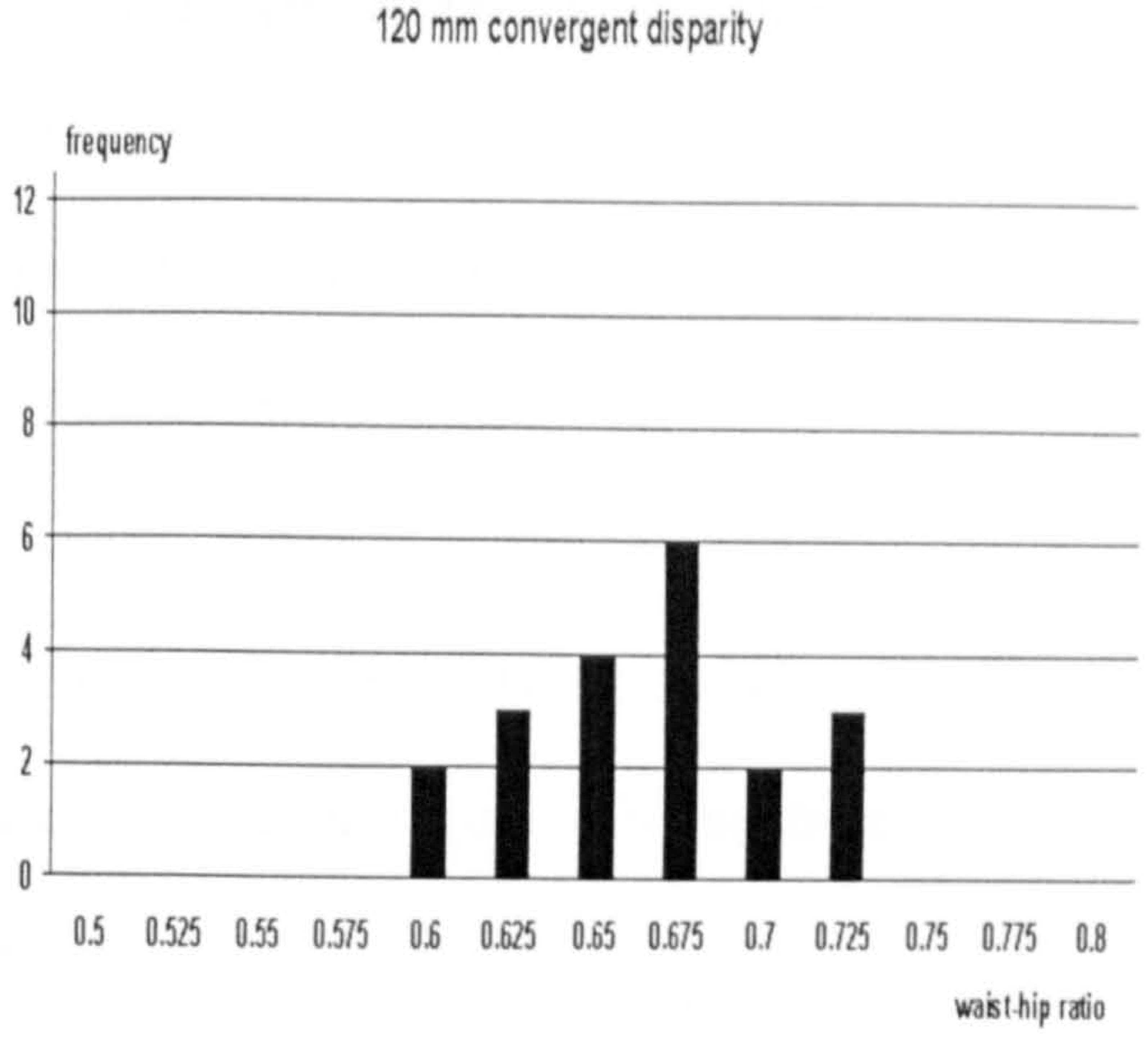


Figure 5.13

Figure 5.12

Figures 5.10 to 5.12. Shows the matches the participants made when shown the shape with a waist-hip ratio of 0.7: Fig.5.9 synoptically (0 mm. disparity); Fig.5.10 stereoscopically with 65 mm, parallel disparity; Fig. 5.11 stereoscopically with 65mm, convergent disparity; Fig. 5.12 stereoscopically with 120 mm, convergent disparity.

5.4.4 Discussion

It can be seen in Figures 5.10-5.13 that the image capture geometries (or disparities) used in this experiment reveal a previously unseen effect. The 0 Disparity 2D stimuli (card 9) was correctly matched to its projected equivalent (0.7 waist-hip ratio) by over half of the participants (Fig 5.9). The average perceived waist-hip ratio of the group was 0.694 (Figure 5.8). This is almost identical to the occluded area as shown in Figure 8.2 of chapter 8 of 0.692. However, when the viewers were shown the same shape but in stereo with 65mm of convergence disparity (corresponding to the normal geometry of human stereo vision), a match with a significantly slimmer waist-hip ratio was selected. Conventional stereo cameras do not capture images with convergent lens axes but use parallel capture geometry. When this condition was simulated with a test image (65P) the mean perceived waist-hip ratio did not differ significantly from the synoptic condition. It can also be seen in Figure 5.10 that there is much more variation in responses in this condition.

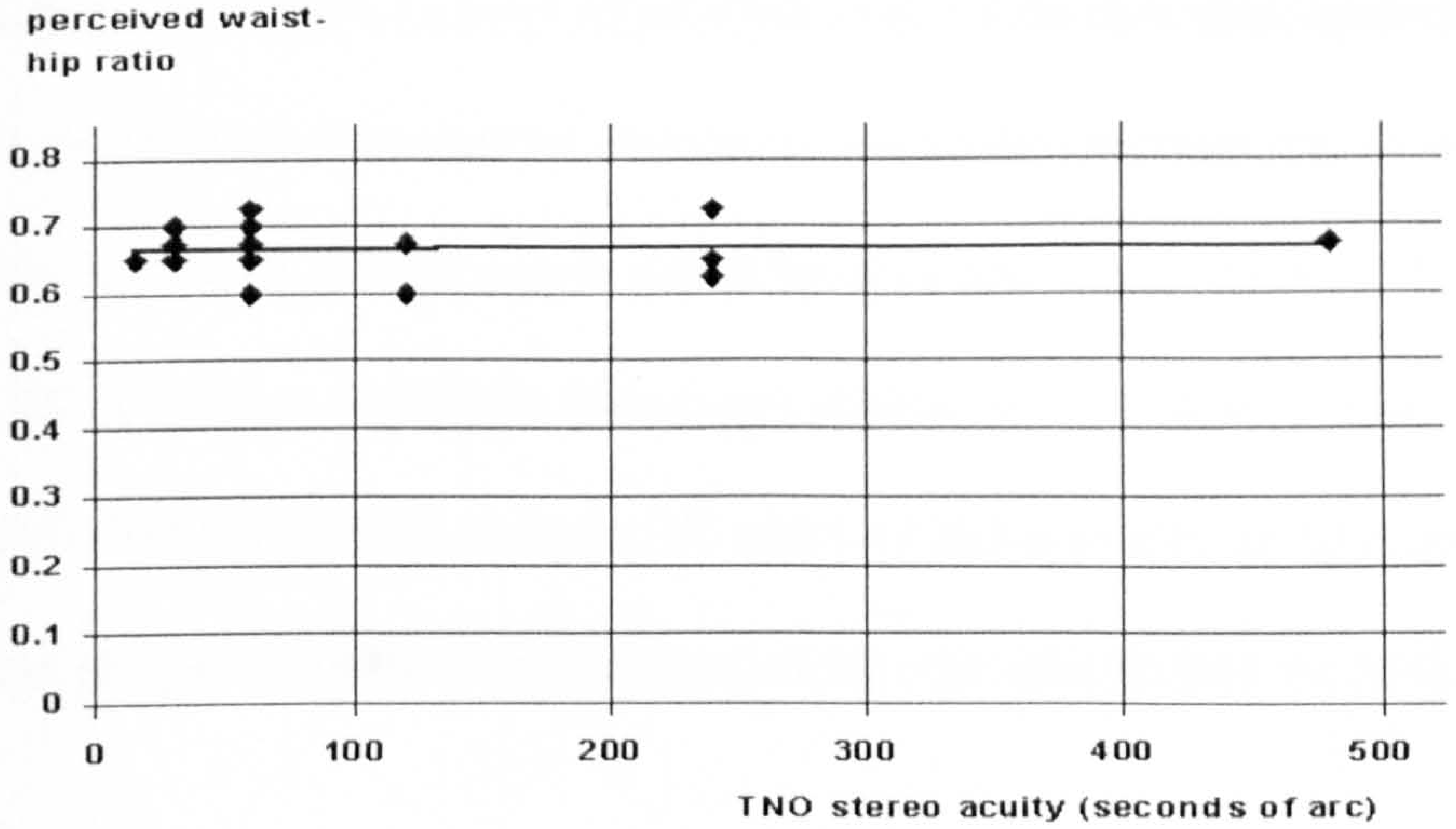


Figure 5.14. The relationship between perceived waist-hip ratio in the 65 mm, convergent disparity condition and the stereo acuity of the individual participants in Experiment 3.

There was no correlation between the participants stereo acuity, measured with the TNO test, and perceived waist-hip ratio in the 65C condition ($r = 0.29$, $n = 20$, $p = 0.904$) (Figure5.14). Subdividing the participants into those with high (15-60 seconds of arc) and low (120-480 seconds of arc) stereo acuity and using a mixed design two-factor (acuity and disparity) ANOVA showed there was no effect of acuity on their performance in the size judgement task ($F(1,18)= 0.46$, $p = 0.506$). Neither was there an interaction between the effect of disparity on size judgements and the performance in the stereo acuity test ($F(3,54)=1.49$, $p=0.228$).

5.5 Experiment 4: The Second "Textured" Peanut Experiment

5.5.1 Introduction.

The results of Experiment 3 indicated that there were some unexpected outcomes from the revised methodology that would require further investigation in a subsequent experiment. While the relative slimming effect of binocular disparity was again demonstrated, there were a number of confounds to the data that needed to be resolved. It had been expected that when the data was analysed, the average perceived waist-hip ratio of the peanut shape would be at or above 0.7 WHR in the 2D card match. This is because synoptic images are almost always seen as larger than life-size (chapter 7). But presenting synoptic 2D and true stereoscopic images seemed to encourage people to make slimmer than reality size estimations for both types of stimuli.

Additionally, the TNO stereo acuity test gave data that was surprising. There was no correlation between the people's measured stereo acuity and their ability to see the slimming effect of binocular disparity. Participants who posted very low, or zero scores on the TNO test should not have been able to see the stereo detail of the 3D images and therefore made card matches that were 2D, and so should average out 0.7 WHR. However, their data were indistinguishable from those with the highest ratings of stereo acuity. This suggested that these people did indeed have functioning stereoscopic vision that the test was not identifying. Other concerns were that the participants were recording their own responses and may have made errors, that the light required to view the response sheet could affect their vision and that the pool of illumination from the table lamp could flood the projected image. More sophisticated statistical analysis might extract interesting underlying data from an enhanced experimental method and any new conditions adopted. For instance, gender information had not been recorded in experiment 3 and so any effect of gender and perception could not be ascertained from the data gathered.

5.5.2 Pilot Studies with Revised Stimuli and Methods.

The specific questions to be answered in experiment 4 were: Is there an underlying effect of size misperception that was not identified in experiment 3? For instance: Is the slimming effect robust? Is it an artefact of flawed methodology? Were the stimuli used enhancing or reducing the effects observed? Can another test of stereoscopic vision be added to the projected stimulus that will give more useful data about the participant's stereoscopic vision? Can any of the wider disparities out to 300mm of interaxial separation be used in a revised method? Finally, is there any interaction between the effects that has not previously been observed?

When the participants were debriefed after completing experiment 3, they were asked to give their opinions on how well they had performed the task. Typical comments were that most people thought the task was very difficult, their responses were not very consistent and that matching images on small cards to images on large screens was confusing. Additionally, some felt that the shape on the screen did not have a perfect match amongst any of the 13 cards and that they would have liked to have indicated an intermediate choice "between this card and that one." Once again, many were convinced that the image size was being changed on the screen and did not believe that the projected image had the same dimensions in every presentation. The data however showed that they were very good at the task and that most changes in the presentations produced predictable changes in perceived size. It was however possible that the size of the image on the card and its proximity could be causing the perceptual effect of smaller than expected size estimations for the synoptic condition. The way to resolve this was to present the varying waist-hip stimuli at the same size and distance as the varying disparity stimuli. This would ensure that the only differences between each of the images were of disparity and waist size.

An additional projection screen, projector stand and remote control Kodak Ektar projector were sourced. The 13 vari-waist stimuli from the previous experiment were rendered onto 35mm film using the previous method and were sized to be identical to the existing stereoscopic pairs. A difference between these images and the printed versions was that the images intended for projection had their file number printed in the bottom left hand corner. This was to ensure that as each WHR was presented, its identification could be clearly established by the experimenter. In the pre-study presentations, the new screen and projector were set up to the right of the existing screen. A new position for the viewing chair was established on the centre line between both projection screens. The new projector was loaded with the vari-waist images and its remote control cable was run to the right hand position on the armrest of the viewing chair. This was so that the participants could advance the projector with the 2D images similar to the way they used the cards in the previous experiment. The images were projected and found to be almost identical in overall height and width as the stereoscopic stimuli and moving the projector back a few millimetres gave an identical magnification to the image on the first screen. The overall brightness of the 2D image was brighter than the combined image from the two projectors. This was because although the stereo projectors were set to full power, the polarising filters absorb over 50% of the available light output. In order to match this level of brightness, a third polarizing filter was fitted to the 2D image projector (vertically oriented) and the brightness level was set to the lowest setting. This provided an image comparable to the 3D projectors when the synoptic images were projected.

During the pre-study demonstrations to a number of colleagues, it was agreed that the task of judging waist-hip ratio on the right screen and finding a match amongst the 13 images was easier than with the previous experiment. They found the remote control easy to use, as pushing the top button made the images fatter, and pushing the bottom button made the images slimmer. Only one person noticed that the small identification number in the bottom right corner was changing with each image presentation. However, this was a concern, as some people may still find the task difficult and opt to choose the same number each time rather than an image of matching size. It was agreed that a method needed to be found of ensuring that only the experimenter can identify the image selected. This was because the participants would be in near darkness (to improve image quality) and could no longer record their own responses. It was also suggested that the participants should be introduced to the variable waist stimuli before the experiment begins. This was so that they become accustomed to the subtle differences between the stimuli and were confident that most people can make fine judgements between seemingly identical stimuli. This familiarization task incorporated the 13 vari-waist cards and entailed putting them in order from slimmest to fattest. By recording the order, it was reasoned that it should be possible to identify any participants who have difficulty making fine shape judgements and allow their data to be excluded from the study.

The quest for a projected test for stereo-acuity met with failure. The three other tests investigated (the Randot, Fly and Frisby tests) were unsuitable for projection because they were not available in projected versions or could not be copied without giving monocular cues to depth. It was decided to use the peanut stimuli themselves as a test for depth perception. It had been observed that the projected disparities (from zero to 300mm) gave different virtual depth planes for each background. When viewing the projected images, the author could fuse all of the random dot backgrounds regardless of their disparity. At the most extreme disparities, the backgrounds behind the peanut shape were reproduced in virtual space well behind the projection screens and the wall at the rear of the laboratory. The projectors and screens were moved and eventually it was found that the virtual background appeared to be level with the back wall (using the 300mm disparity stimulus) when the screens were four metres in front of it. A modified arrangement of the two projector set-ups allowed a clear view to the lab wall to be seen between each screen. This allowed distance markers (in 0.5 metre intervals) to be set between the screens (figure 5.15) so that viewers could report how much depth the scene had in virtual space when compared with real distances. This was done by the experimenter moving a target backwards and forwards until the participant indicated that it was at the same distance as the virtual background (of the stereoscopic image shown on the left hand screen). This "screen distance" was to be recorded for each of the presented disparities.

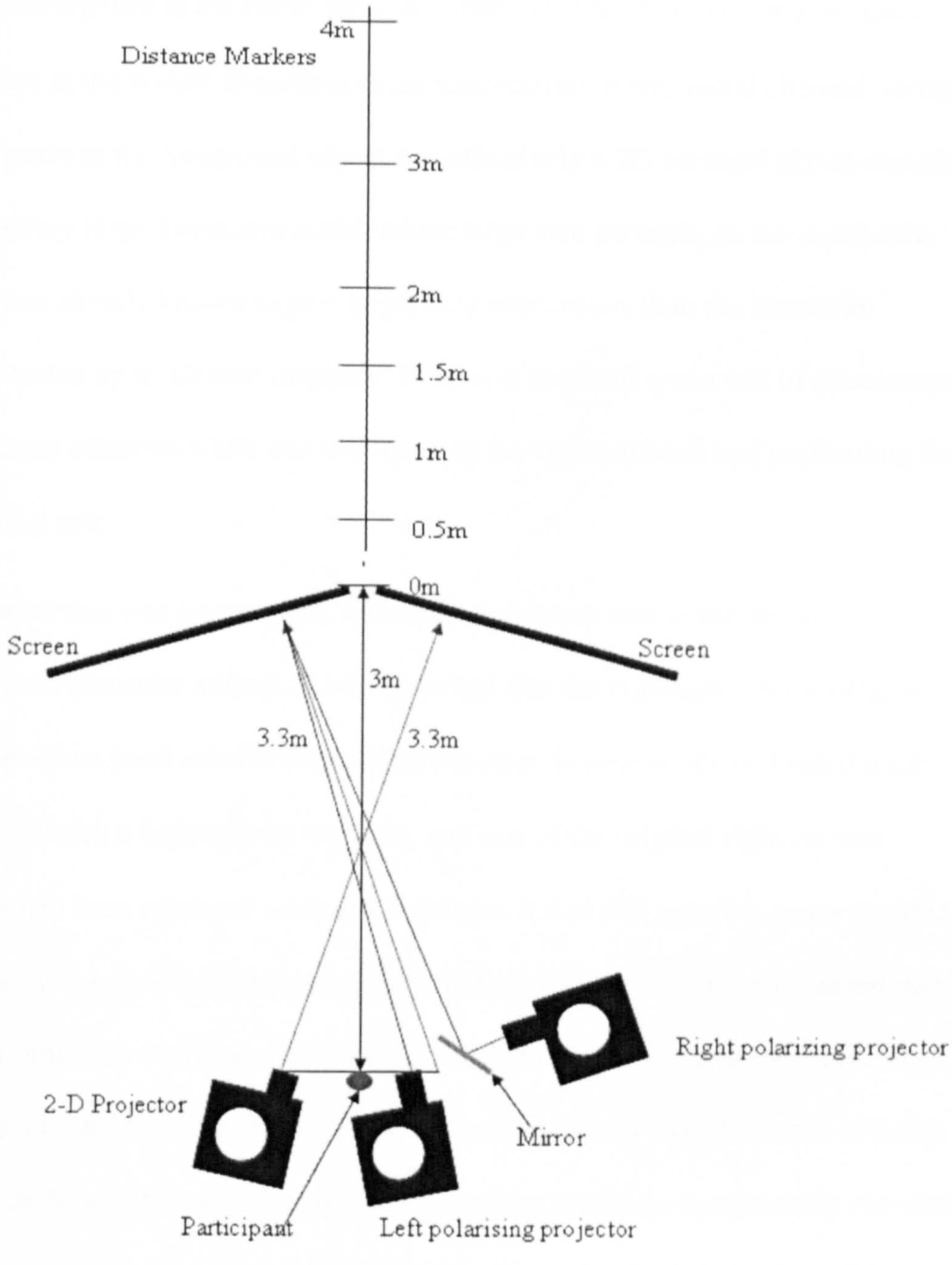


Figure 5.15. Plan view of the projector and screen set-up used in experiment 4. The viewpoint of each participant was positioned between the two sets of projectors at the centre line of both projected images. The stereo convergence angles are identical to those of the image capture configuration.

The pre-study phase showed that fusing the reversed disparities was possible for most people. Only those people who had very poor sight in one eye or had poor or uncorrected vision could not do it at all. The widest camera separations caused the subjects the most difficulty in fusing a stable stereo image. However little is known about what happens when one cannot fuse reversed background disparities when the foreground object is still observable and has very small disparities because it is the centre of convergence in the stereo field. A question we needed to answer is: Does fusion failure at the widest disparities cause suppression in one visual channel, so that when one gazes at the foreground object it is effectively a 2D percept? If that were the case, using very large disparities could induce large size percepts, as the monocular condition was already known to give larger size estimations than the binocular conditions tested up to 120mm disparity. What was required was a test of stereoscopic vision that was effective while one was viewing the stereo stimuli and performing the size matching task.

A proposal that was investigated with this experiment was to use mixed monocular and binocular stimuli. It was observed that the highlights on the original peanut shape gave good relative disparity depth cues. However, if one loaded a left camera image with a highlight on the front, and one of the original right camera stimuli that had been rendered without a highlight, it was still possible to see a stable stereo image when the highlight was seen by only one eye. It had been expected that any lack of binocular correspondence might cause retinal rivalry and the image would be unstable. Or alternatively, the highlight in the left channel would be seen as being behind the surface of the peanut shape, as its distance would be monocularly encoded as at the screen rather than in the virtual space in front of it. In fact, neither of these things happened. *The highlight appeared in to be in virtual space in front of the screen, even though it had no binocular correspondence.* Under these circumstances, the brain seems to accord this highlight its correct position in space on the grounds of ecological validity. This is because in real space, it is possible for an object's surface to reflect a highlight to one eye only. So the brain has to apportion a plane of reflection to this region from monocular cues alone.(Ref Howard). The perceptual process appears to be the same for virtual stimuli as it is for real objects. So the peanut shapes were re-rendered with three highlights (fig 5.16) in both the monocular binocular targets. The stereoscopic targets however had two of the three highlights rendered in one channel only. If the participants were seeing the peanut in stereo, they would be able to report seeing three highlights. If one channel was suppressed, they

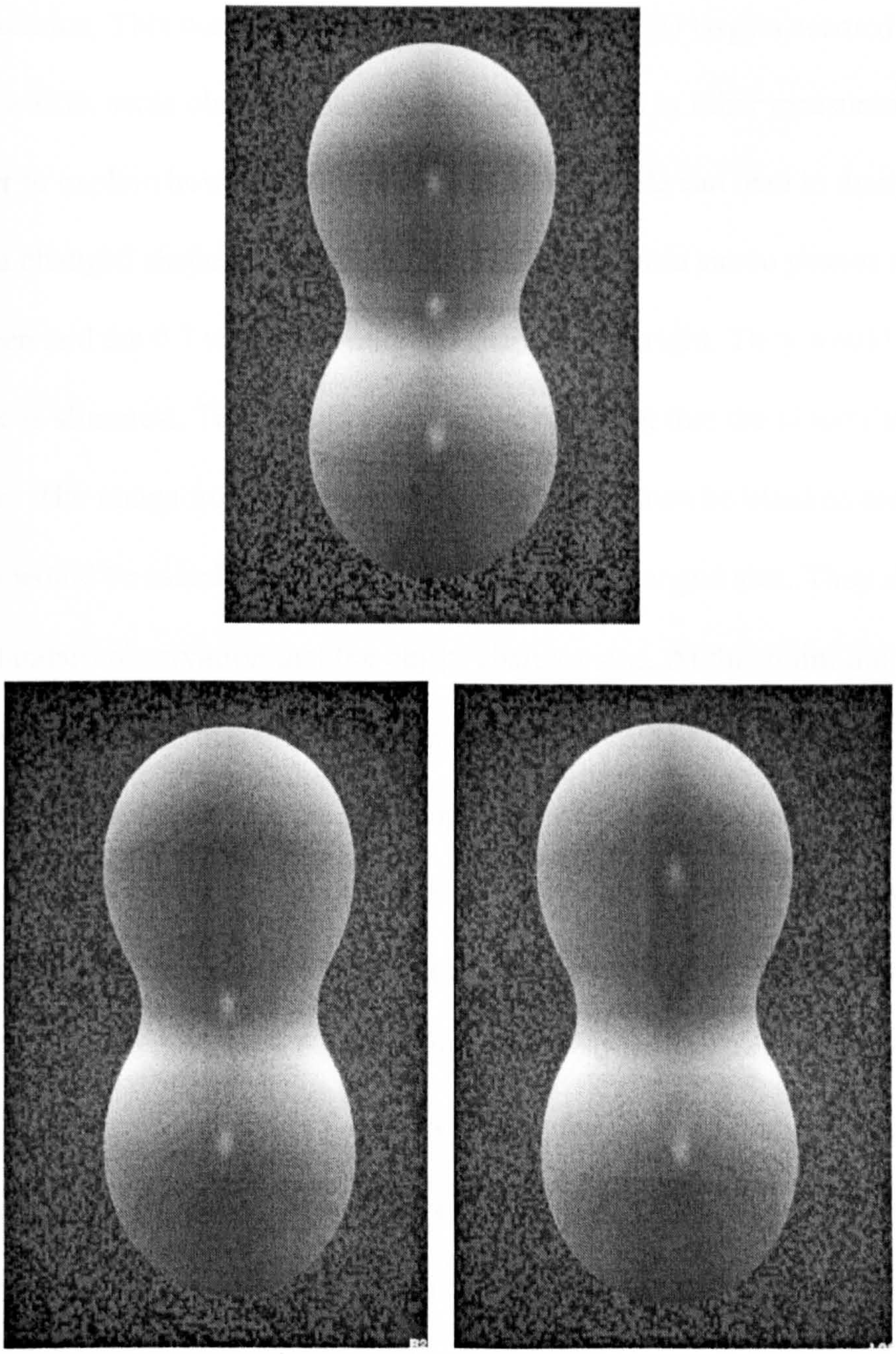


Figure 5.16. The first image is from the central camera and has all three highlights rendered. The lower images are the right and left stereo pair. Their central and highest highlights do not have corresponding stereo disparities. Observers who cannot fuse a stereoscopic image and suppress one channel will therefore report seeing only two highlights.

would report seeing two highlights. Suppression (if reported) could then be correlated to the reported shape, which should match to the standard 0.7 monocular match.

The final task for these pilot study participants would be to report their general impressions of the size and shape of the 2D and 3D targets during the normal debriefing session. This was because although the 2D and 3D targets seemed to vary most at the waists, some observers thought they differed to in other measurements too. In order to explain how projecting the same shape in 3D can lead to their perceiving a changed shape, they were to be shown the 65mm stereo peanut shape on the left screen and the 0.7 waist- hip ratio 2D shape on the right. They would be asked which shape is slimmest. This should lead to their reporting that the binocular stimuli was slimmer. The image from one stereo projector would then be blanked and the participants would be asked if the left screen image had changed size. They should report the standard observation that the object changes size. At this point the participant would be told that the objects on the stereo screen did not change the actual shape at any time, and the shape change they perceived was in the brain and not measurable on the screen. Once made aware of the effect, they were to be asked to look at all of the dimensions of the peanut and to report two observations about the objects: Is there a size difference between the objects? If they reported the right 2D object was larger for instance "left larger" would be recorded. The alternative possibilities were "right larger" or no difference. As similar question was to be asked about which object had the greatest volume and the three possible responses recorded. It was expected that these responses would establish if there was any variation between the participants in perceiving to the overall dimensions of the shape compared with the consistently slimmer responses given to the shape at the waist area.

The general principle adopted in designing this experiment was to gather as much data as possible from the participants without tiring them during an expected period of 25 minutes of viewing stereographs in a blacked-out room. A concern raised by the previous peanut experiment is that people would not try hard enough to make matches because they were confused by the task, or thought it was too difficult. This could lead to unacceptable levels of variation in the data, so it was agreed that all of the participants should be paid £10.00 for one hour of their time and that the experiment would be run at the most convenient time in their schedule. The card-ordering task should familiarise them with the vari-waist stimuli in daylight conditions, so this task and the TNO test were taken in daylight at a separate location before the main size-matching phase of the experiment was conducted. Additional data gathered in the pre-test session were about the vision of the participants and if they required correction in order to see the all of the stimuli. All participants were to be told to bring their spectacles or contact lenses, and that short sighted people should wear their correction at all times and that long-sighted people should wear them only for the pre-test session. The opportunity was taken during the pre-test session to give each participant a simple eye dominance test to get a more general picture of each participants visual capabilities and to report to them good news about their visual capabilities (such as TNO scores and accuracy in the card ordering task) before the main test phase. *This was to ensure that they were happy with their vision, their ability to do the task and had the opportunity to decline taking part in the size-matching phase if they had any concerns about the experiment.*

5.5.3 Textured Peanut Stimuli.

The pilot study reported in section 5.5.2 found that the previously observed slimming effect of binocular disparity was present. However it did not find the synoptic effect expected, in that the peanut shape was never seen as larger than the 0.7 WHR of the actual stimuli. Yet the experience of viewing targets with a synoptic viewer (chapter 7) suggested that participants should see any synoptically projected shape as larger. Two possible causes of this unexpected finding were identified: The participants may have thought there was no image larger than 0.7 that was a suitable match for a size estimation that was only very slightly larger than 0.7. If the participants had been able to choose from WHR's with smaller gradations, they might have produced matches closer to what they see. Also, it was noted that the lack of detail on the surface of the peanut was very unnatural. Most objects have a surface texture that allows the HVS to see subtle changes in surface shape. Perhaps adding a surface texture would make any changes in WHR more clearly seen by the participants.

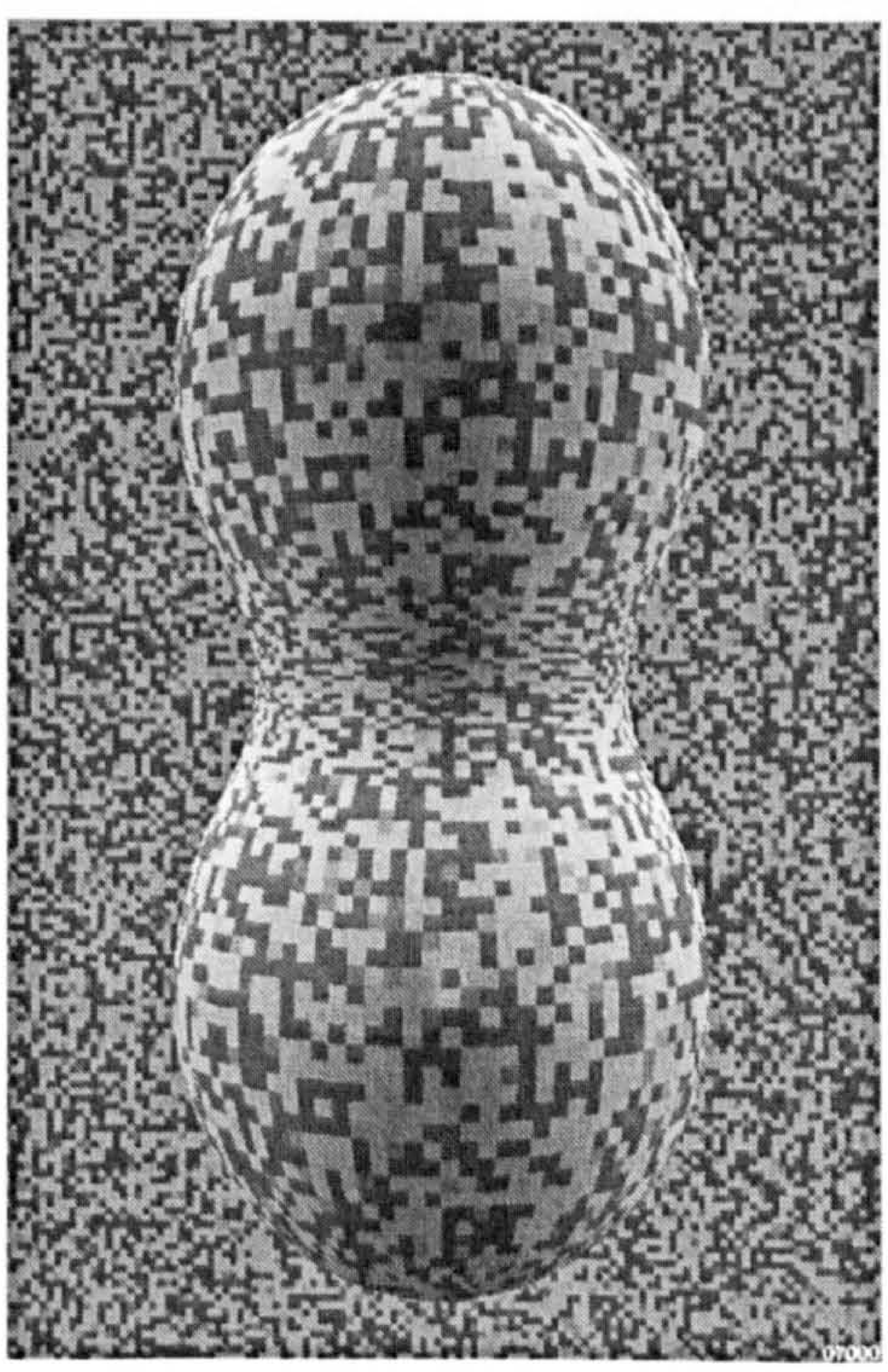


Figure 5.17 The random dot background used in all of the experiments was wrapped around the 0.7 WHR peanut shape to give it an enhanced surface texture. This master image was used to make all of the "textured peanut"stimuli in all subsequent presentations.

Improvements in computing power had meant it was now possible to test render the images with a surface texture. In order to avoid misleading patterns or the wire-frame effects of the peanut images in the feasibility study, it was decided to use the PhotoShop random dot background again and to wrap it around the object. All of the lighting and tonal qualities of the image were kept identical to the original peanut images and only the background texture was used to revise the new stimuli.

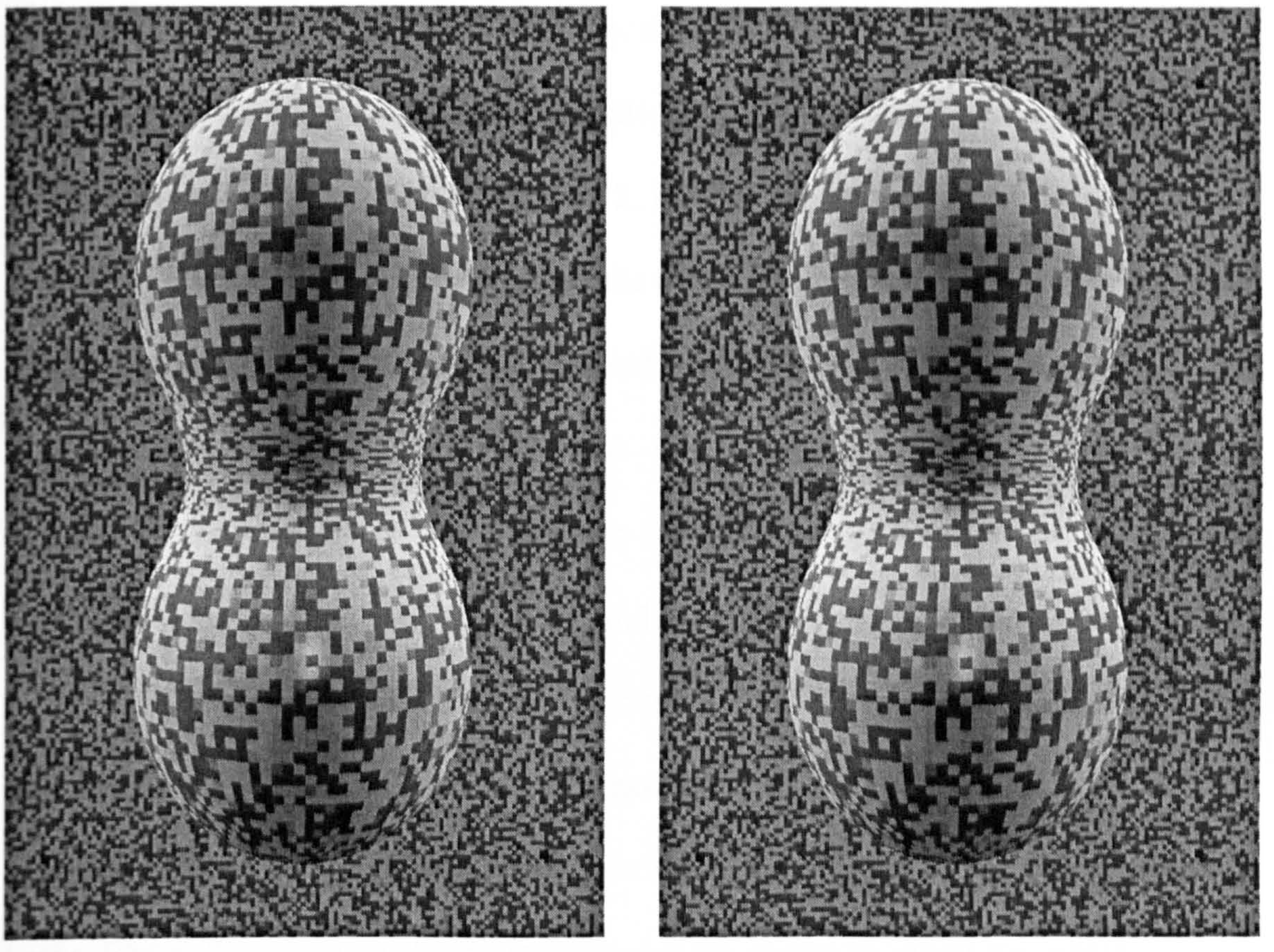


Figure 5.18 These are the right and left (60mm separation) stereo pair from the image set used in the next presentations. The highlights used as a stereo test in the experiment 2 images were removed from this set, as everyone who participated in the experiment reported seeing 3 highlights in all conditions. This meant that it could not be used a discrimination criteria that was independent of perceived shape change, and was therefore unnecessary in this version of the peanut experiments.

The new stimuli were labelled as "T" for textured to differentiate them from the white peanuts, which kept their original labels. Despite being 50% darker, due to the average mid-grey values of the pixels, the images appeared to be bright and have better sculptural qualities than the otherwise identical white peanut shapes.

Aesthetically, the new images appeared to be much more pleasing to the eye and easier to view than their predecessors. The images were inspected for pixelation artefacts from the proscribed viewing distance of 1.68 meters. It was found that almost all of the surface rendered random dot pixels had corresponding brightness's in each of in each of the channels. However, a close inspection revealed one area of sub-pixels that were bright in the right channel and dark in left. Despite causing some binocular rivalry when viewed in close-up, the effect was away from the waist area and too small to see from the viewing position. It was also found that, as with the monocular-only highlights in experiment 2, it did not cause false depth perception and seemed to simply be a surface characteristic of the image. As correcting this sub-pixel would require that the whole set be rendered at a much higher resolution, these images were retained without modification. Another decision made at this time was to not use any disparity wider than 120mm in the second peanut experiment, as it would take to long to run the widest disparities and little new information could be gathered by their inclusion. These stimuli were then added to the previous set so that the effect of adding texture to the target could be measured.

(A short summary of the results of the pilot study will be placed here and a reference to the data in the appendix. It will indicate that the slimming effect was again observed, that the textured peanuts were seen as slightly slimmer than the previous stimuli in all conditions except the synoptic.)

5.5.3.1 Stimuli with Corrected Vertical Disparities.

The textured peanut stimuli gave the opportunity to see effects that had not been previously observed. When the synoptic pair of textured peanut shapes were projected in the best possible registration, it was clear that (apart from the edge of the peanut shape) there were no areas with perfect registration. This is because when one attempts to superimpose identical images using converging projectors, trapezoidal distortions will not allow the identical images to have pixel-accurate matching. The convergences will cause clear vertical disparities on the edge of the overlapping images that when viewed stereoscopically could give a convex distortion to what should be a flat background plane (see figure 5.19).

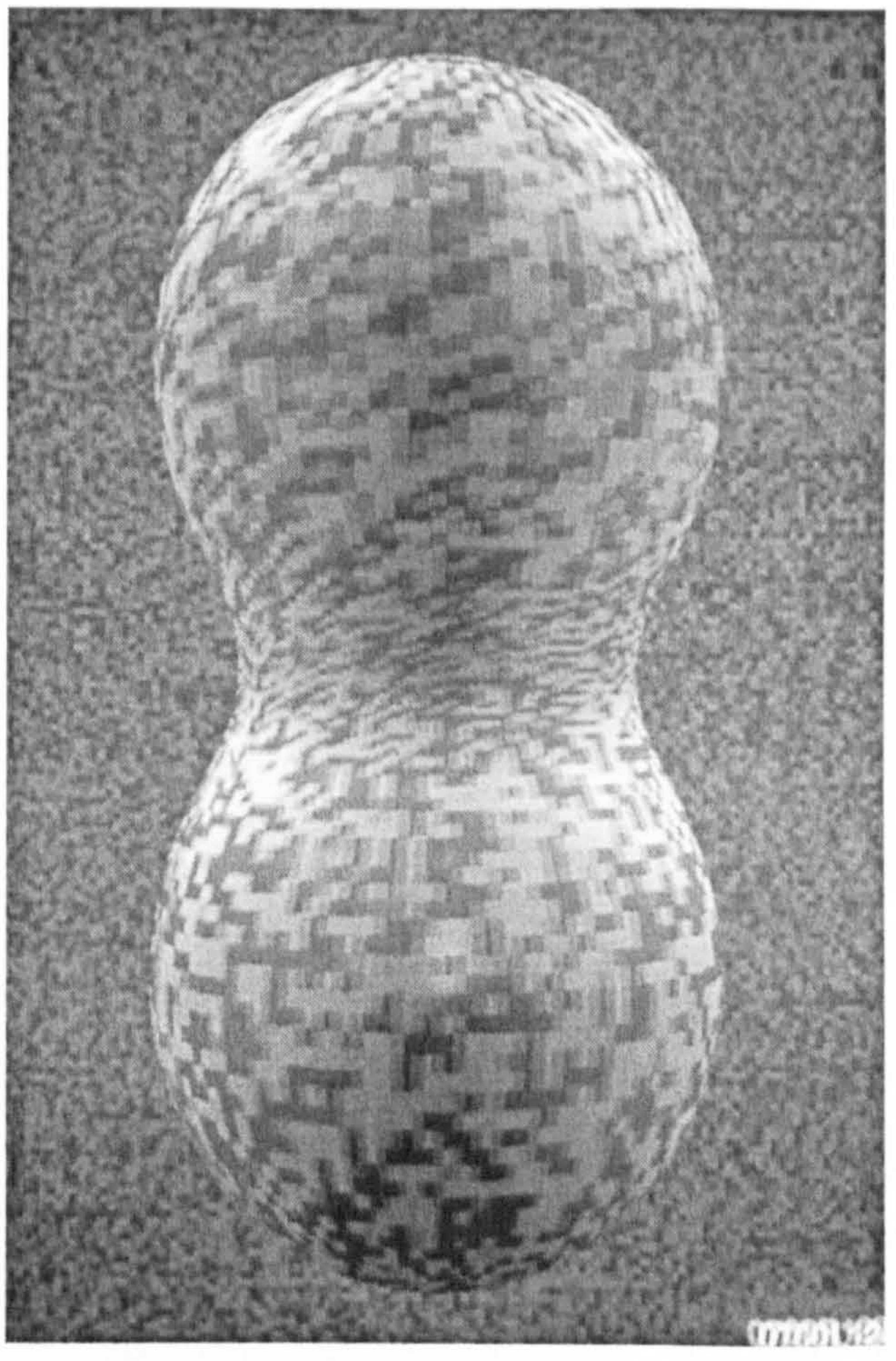


Figure 5.19 This image shows an exaggerated example of vertical disparities observed (without polarizing spectacles) when viewing identical images projected with the best possible horizontal registration.

The addition of texture to the peanut shape made it clear that misregistration was evident on the target too. Perhaps these incorrect vertical disparities were causing a misperception that was disguising the size difference between the synoptic and 65mm stereo stimuli?

The previous peanut experiments relied on the fact that computers could generate images from virtual camera positions that did not exist in reality. The projector's convergence was set to match the virtual cameras at 60mm of separation, which was 2.04 degrees. All of the other virtual camera positions (synoptic, 30mm, 120mm etc) gave images that could not be perfectly matched by the projectors unless they were moved during the presentation. As this impractical, it was decided that virtual projector positions could be simulated in software in a similar way to the virtual camera positions. A programming rule was devised (see appendix E) to trapezoidally distort the synoptic images to compensate for fact that the fixed projectors positions in the experiment were correct only for the orthostereoscopic stimuli.

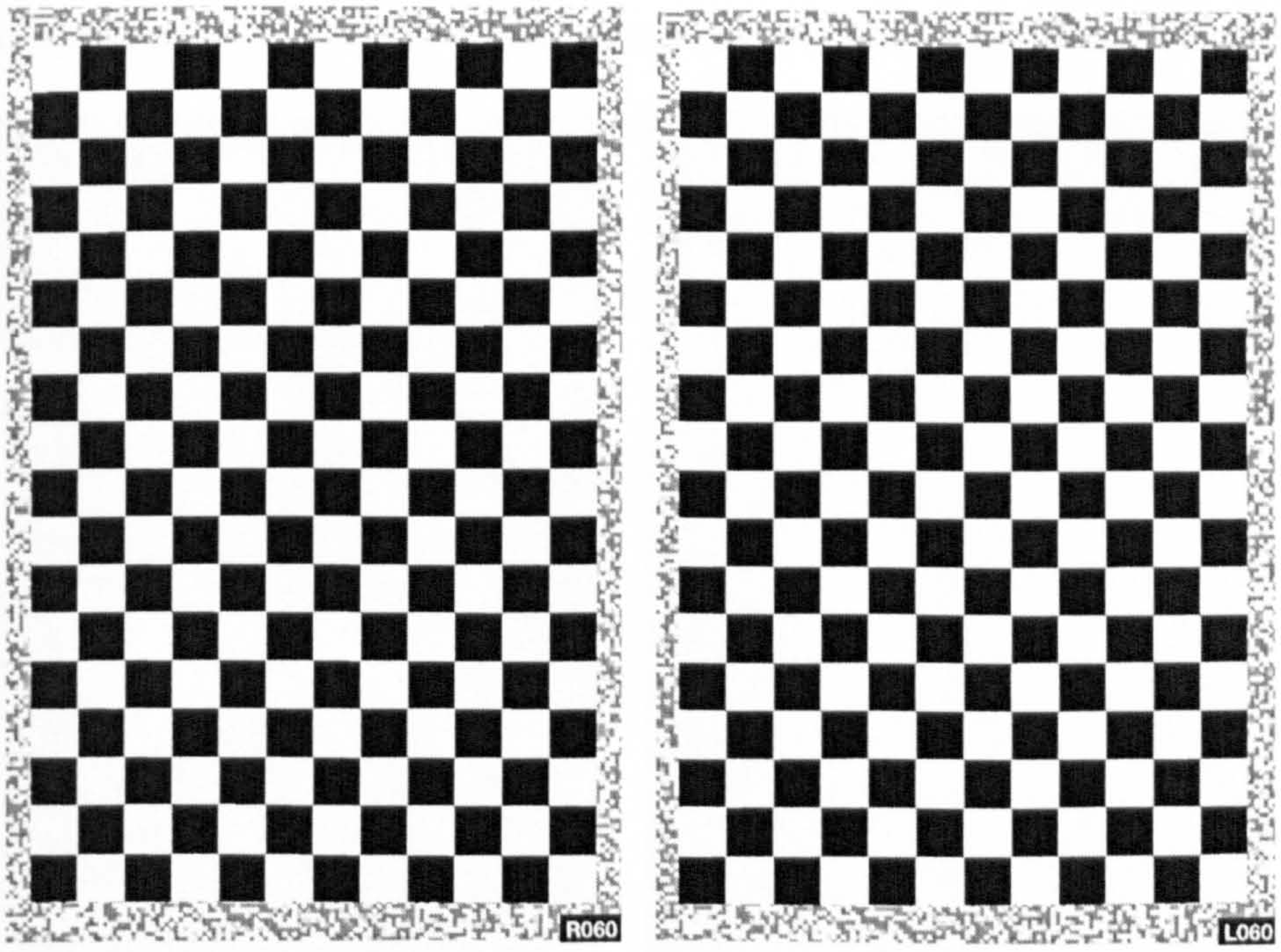


Figure 5.20 This "chess board" was rendered in PhotoShop and inserted in the StrataVision model at the same camera to subject distance as the peanut. It was then photographed using the virtual cameras at 60mm separation and converging at 2.04 degrees. They were then printed on to 35mm film.

Test images of a virtual chess board were projected using the same converging projectors used previously. The images were found to be close to perfect registration on first inspection. When aligned, the chess boards were measured and found to have no distortions. Additional registration marks, consisting of four black pixels were inserted into the test image using the revised geometry rules. The registration marks were designed to be trapezoidal on the computer screen, in that none of the angles were of 90^0. However on projection from the converging projectors, they would appear to be emanating from a central projector position and have perfect right-angles at each corner. These "correcting distortions" reversed the measured keystoning of the original synoptic pairs to give a rectangular appearance of the registration marks and target when measured from the screen. The same rule was used to crop the edges of the images to a black border, so that the edge of the transparency mounts did not keystone the outer edge of the projected image which too was measured as having right-angles at each corner. The rule was applied to the 00 (centre camera) image so that a zero disparity left and right test transparency could be generated (figure 5.21).

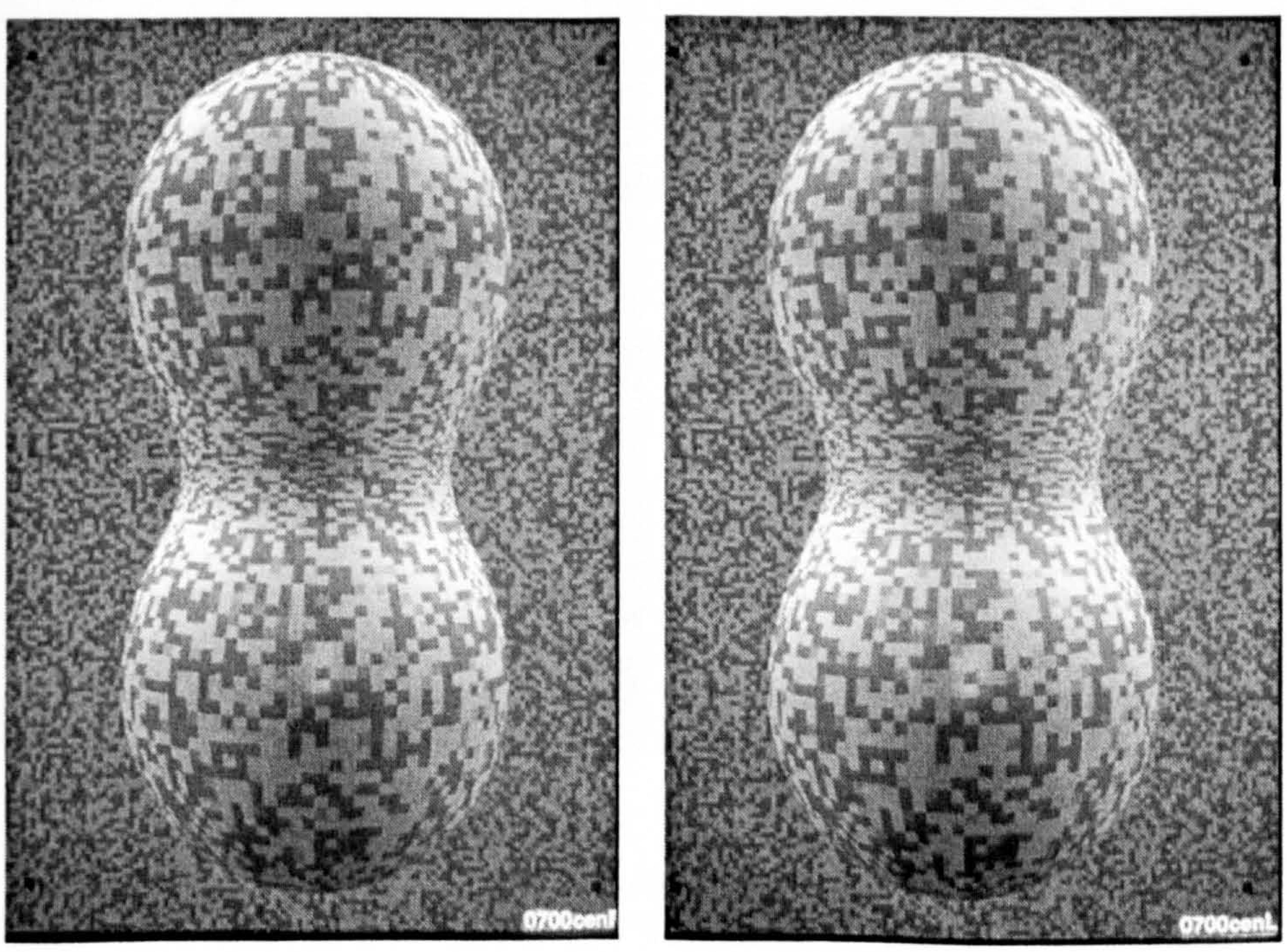


Figure 5.21 Right and left textured synoptic stimuli with corrected vertical disparities.

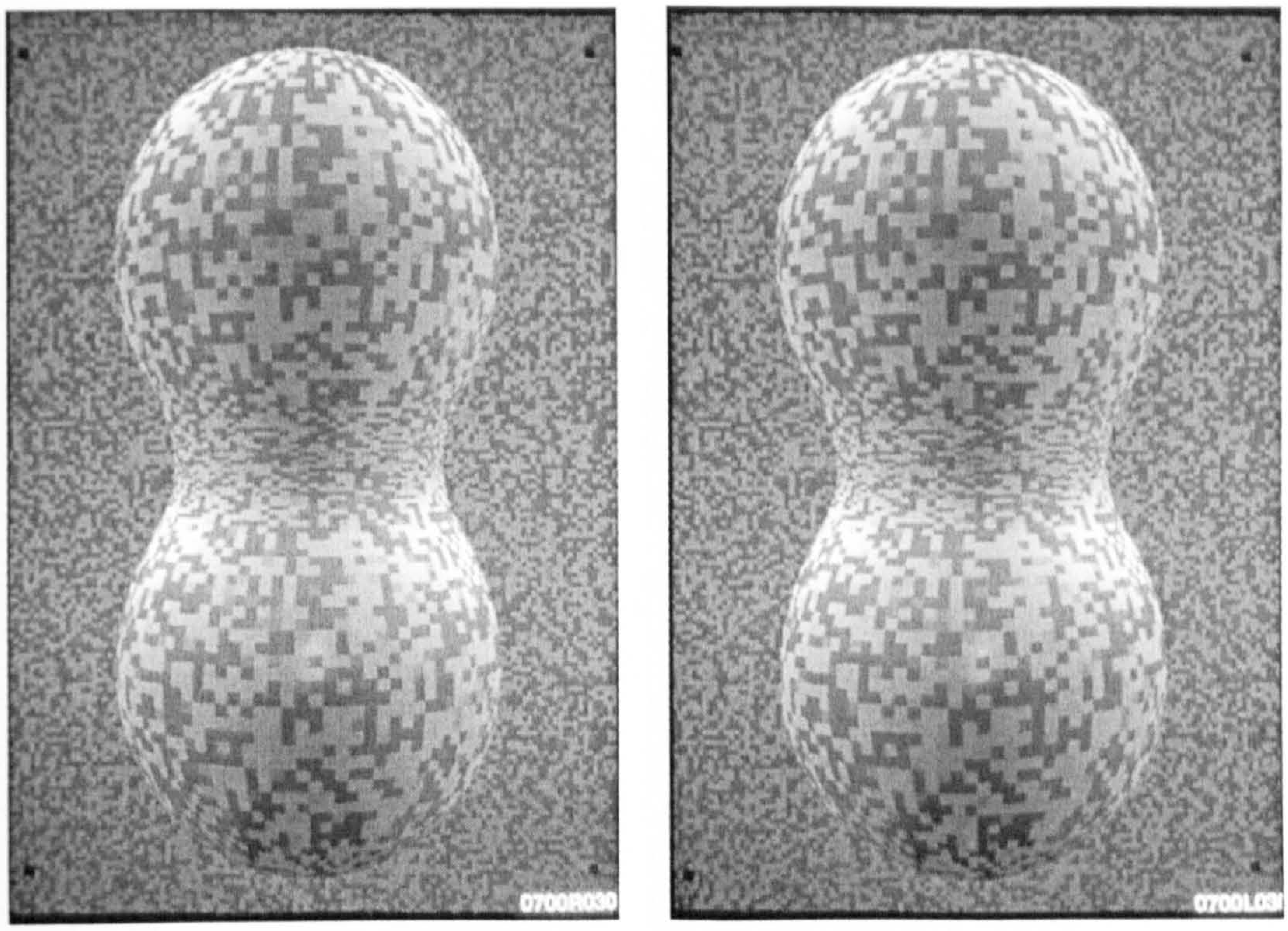


Figure 5.22 The right and left projector image renders with 30mm of stereoscopic disparity. These images have the corrected vertical disparities used in experiment 2.

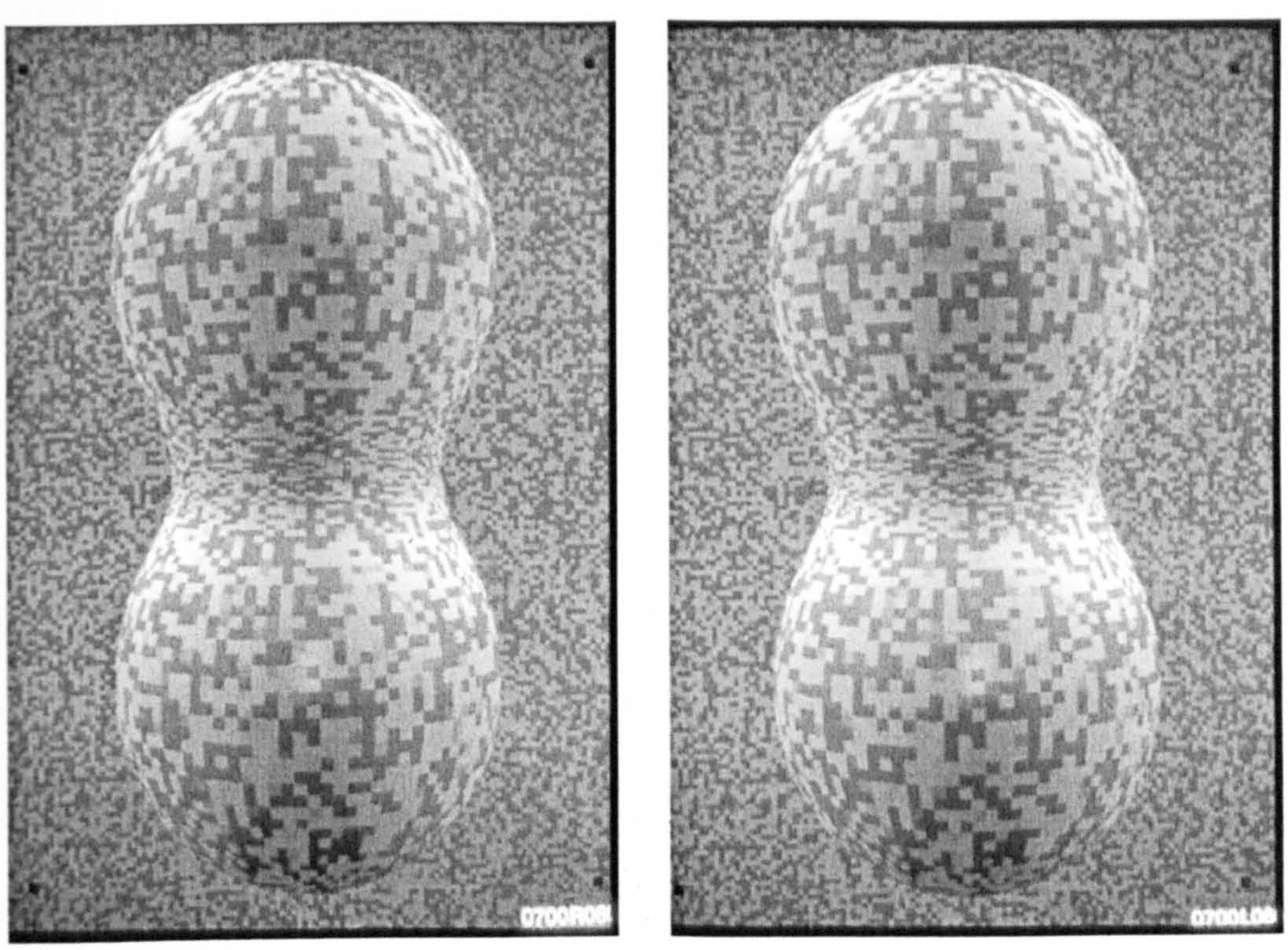


Figure 5.23 The right and left orthostereoscopic pair used in experiment 2 were rendered without any of the trapezoidal distortions applied to the other stimuli used.

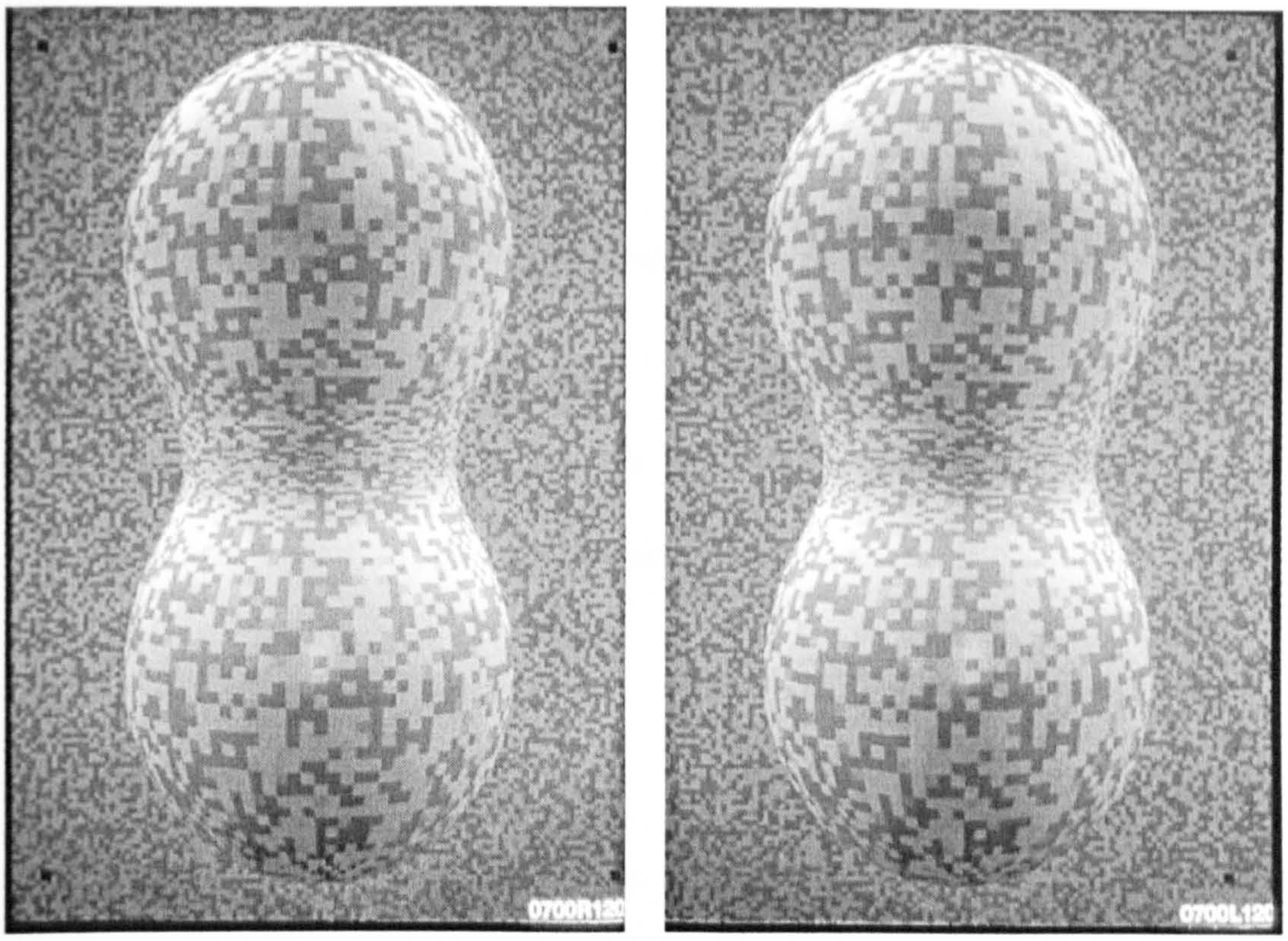


Figure 5.24 The120mm disparity right and left projector images. These image were different from the synoptic and 30mm pair in that the trapezoidal distortion was applied in the opposite direction (see appendix E). These corrections were also applied to the untextured images from experiment 1 so that they could be compared with the final versions.

When correctly aligned, the left and right synoptic images gave seemingly pixel accurate registration on the screen. *When viewed with or without the polarising spectacles, the synoptic image appeared to be almost identical to the single projector image of the 0.7 peanut shown on the adjacent screen.* When comparing the new synoptic pair to the previous design, the image was superior in quality and the revised vertical disparities brought the image forward in virtual space. The rules underlying the rendering of the new synoptic pairs were then applied to the other images from the original experiment 2 set (including the original untextured images) with the exception of the 60 mm converged pair which were already matched to the projection set-up. The revised disparities were projected in the same manner as before. All appeared to be easier to fuse than their predecessors, but more importantly, the random dot background was free of any curvature of field in all of the disparities.

5.5.4 Method

5.5.4.1 Participants

Twenty-one participants from The University of Liverpool took part in the study.

The sample consisted of 10 males and 11 females with a mean age of 22.57 years, ranging from 18 to 39 years of age. All had either normal or corrected-to-normal vision and were naïve about the hypothesis being tested. Each participant was tested individually.

5.5.4.2 Procedure

Before entering the laboratory, participants were given a number of preliminary tests. They began by taking the TNO test for stereoscopic vision (TNO, 1972), and then The Frisby Stereotest (19). Their scores for both of these tests were recorded. This was followed by a card-ordering task. The participants were presented with one of the sets of 13 WHR cards (either textured or white), which had been randomised, and were asked to place them in order from slimmest to fattest. This was repeated with the remaining card set and any mistakes were noted. Scores for the card-ordering task were later calculated depending on both the number of errors and the position of the error. If two adjacent cards (say 7 and 8) were translocated (so that the order read ...5, 6, 8, 7, 9, 10...) then a score of two was given because each individual card had moved one place from its intended position. If two cards were translocated (say 6 and 8, so that the order read ...5, 8, 7, 6, 9, 10...), then a score of four was awarded, because each card had moved two places, and so on.

The participants were then shown into the laboratory, and were asked to sit in the viewing chair. All lights were switched off, except for the lamp illuminating the distance markers. Each participant was briefed about the task by way of a pre-

prepared script (see appendix F). They were told to "hunt" for a matching image among the comparison stimuli, using the remote control to scroll through the images until a match was found. They were encouraged to always go past the point at which they thought there was a match, and then return to what they perceived as the matching image. They were also encouraged to voice their thoughts (e.g. "the image is too thin", "the image is too fat", "that is the match", etc.). After briefing, participants were given a short training session, using the images with larger gradations, to accustom them further, and familiarise them with both textured and white 3-D images. They were given four practice trials, consisting of two textured images at 0 and 60mm of disparity, and two white images at the same disparities. The results of their choices were noted. After this training session, the participants were asked to estimate the distance of the background at each level of disparity, using the distance markers as a guide. The main part of the experiment commenced only after these preliminary stages were completed.

Participants were presented with the first image of the sequence (at 120mm of disparity) under the relevant surface condition, and were asked to pick a match from the comparison set of stimuli. This was repeated with the remaining three images (at 60mm, 30mm, and 0 disparity). To avoid order effects and to give two opportunities to make a size match at each disparity, the whole sequence was then repeated again in the same surface condition (0, 30mm 60mm and 120mm disparity) and the mean of the two measurements taken. The participants were then shown the images in the alternate surface condition and repeated this double sequence. After the experiment, each participant was debriefed as to the purpose of the study.

5.5.5 Results

To determine the effects of disparity and surface texture on size judgement, the participants' perceived 2-D match to each polarized image was recorded, at each level of disparity (0, 30mm, 60mm, and 120mm) and in both the textured and white surface conditions. Figure 5.25 shows the frequency distributions of participants' matches for the four levels of disparity and under both textured and white surface conditions.

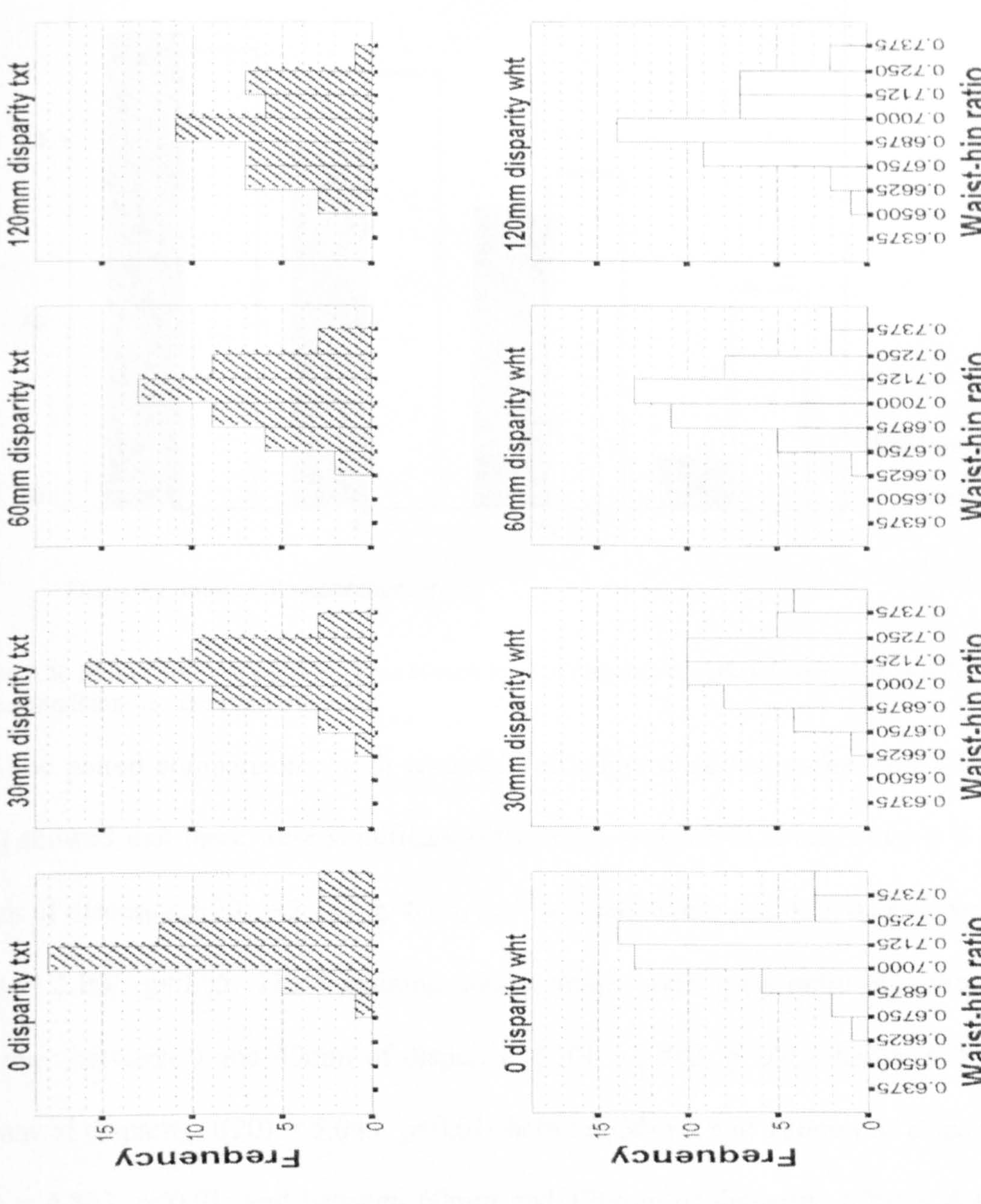


Figure 5.25. Shows the frequency distribution of participant matches to the 0.7 waist hip-ratio shown synoptically or stereoscopically. Cross-hatched bars show the textured peanut condition; white bars show the white peanut condition.

Figure 5.26 shows the overall group means for these choices. A two-factor (disparity and surface) within subjects ANOVA revealed an overall effect of disparity on size judgement, $F(3, 60) = 19.126$, $p<0.01$, but no significant effect of surface on size judgement, $F(1, 20) = 3.081$, $p>0.05$. The interaction of disparity and surface was not significant, $F(3, 60) = 1.506$, $p>0.05$.

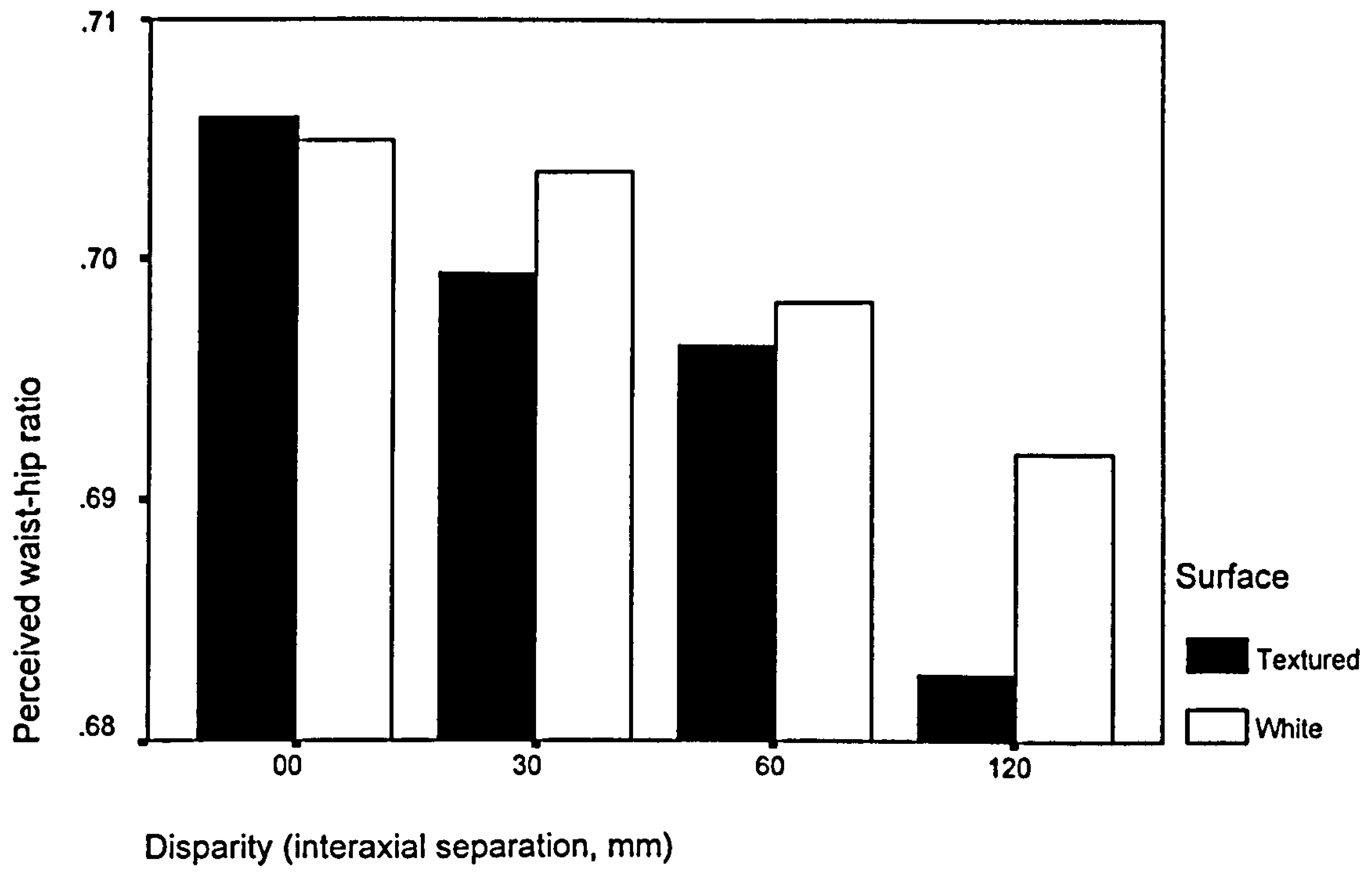


Figure 5.26 Mean perceived waist-hip ratio at each level of disparity (interaxial separation) for the two surface conditions textured and white.

Post-hoc paired comparisons (with textured and white conditions combined (figure 5.27) showed that there were significant differences at the 0.05 level between 0 and 30mm of disparity, $t(20) = 2.187$, $p<0.05$, and between 30mm and 60mm of disparity, $t(20) = 2.190$, $p<0.05$. The remaining combinations were even more significantly different: between 0 and 60mm of disparity, $t(20) = 3.882$, $p<0.01$; between 0 and 120mm of disparity, $t(20) = 5.043$, $p<0.01$; between 30mm and 120mm of disparity, $t(20) = 4.803$, $p<0.01$; and between 60mm and 120mm of disparity, $t(20) = 4.396$, $p<0.01$. Despite a non-significant result for the effect of surface on size judgement, by

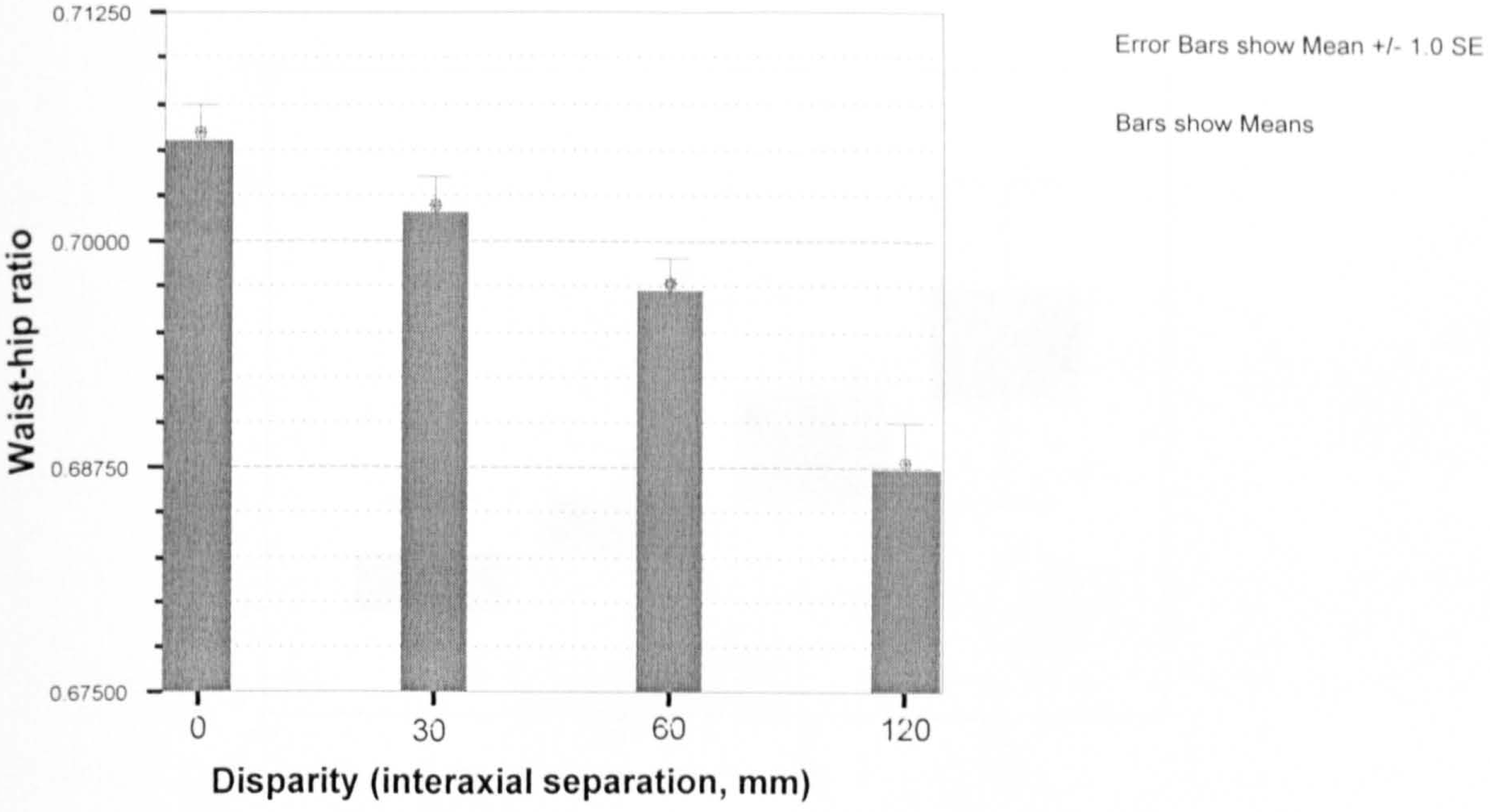


Figure 5.27. Mean perceived waist-hip ratio (textured and white conditions combined at each level of disparity

examining figure5.26 it seemed clear that there was a difference between the textured and white conditions at 120mm of disparity. A post-hoc paired comparison supported this observation by finding a significant difference between the textured and white conditions at 120mm of disparity, $\underline{t}(20) = -2.223$, $\underline{p}<0.05$. No significant differences were found between the textured and white conditions in the three remaining levels of disparity.

The level of disparity, therefore, affects size judgement. With increasing disparity (from 0 to 120mm of interaxial separation), smaller matches were chosen by the participants. Overall, surface texture seems to have no effect on size judgement. However, at 120mm of disparity, smaller matches were chosen in the textured surface condition than in the white surface condition. The participants were also asked to judge how far from the screen the background in the polarized images appeared to be, at each level of disparity. Figure 5.28 shows the results. The mean estimated distances

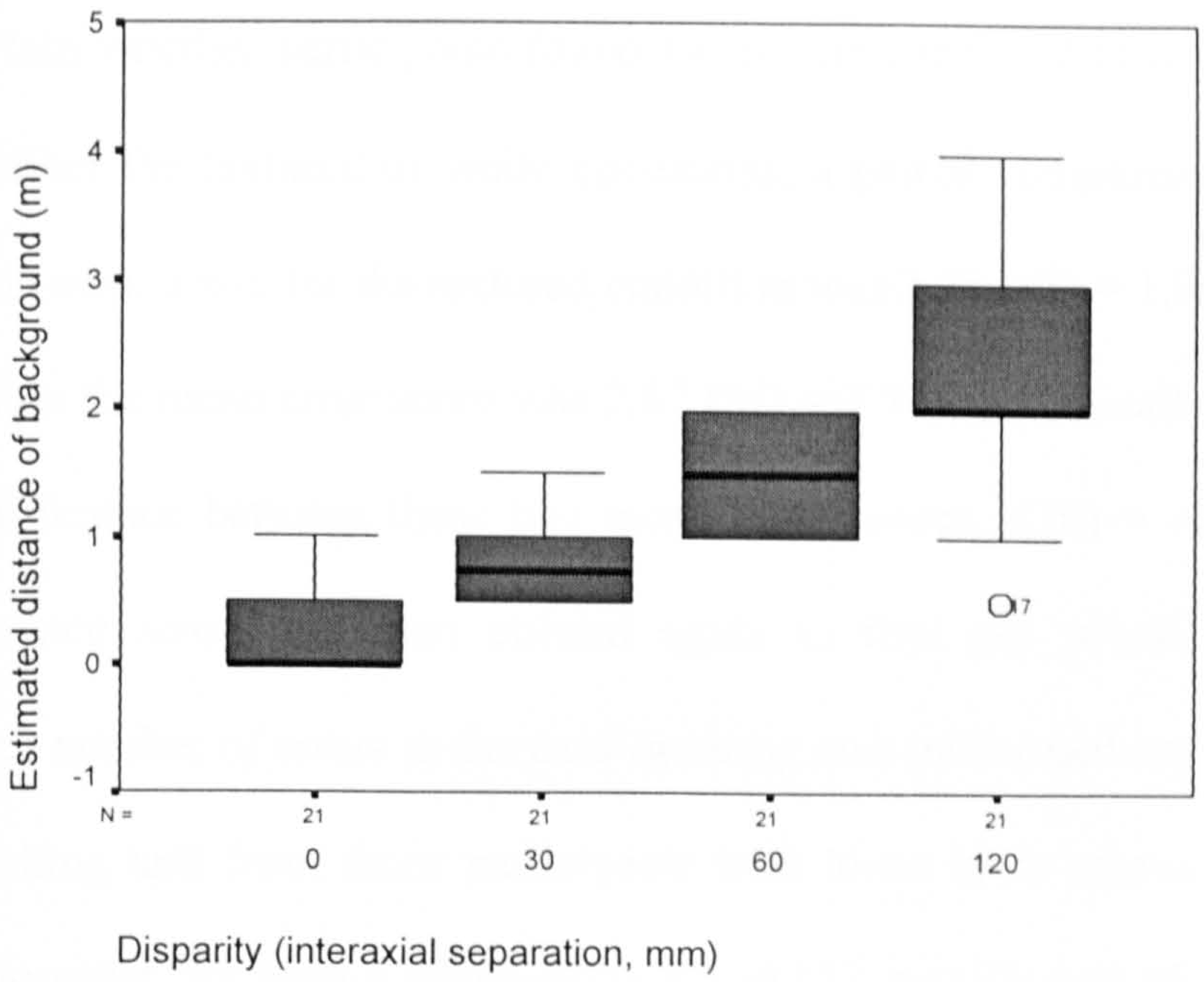


Figure 5.28 Estimated distance of background at 0, 30mm, 60mm, and 120mm of disparity.

at 0, 30mm, 60mm, and 120mm of disparity were 0.24m, 0.77m, 1.44m, and 2.40m, respectively. A correlation between the perceived distance of background and disparity proved significant, $r = 0.835$, $n = 84$, $p<0.01$. Thus, as disparity increased, so did the perceived distance of the background in the stereoscopic image. The background of the stereoscopic image at 0 disparity should be on the screen (0m) because this image is effectively 2-D. To discover whether individuals estimating the background to be further back from the screen performed any differently when choosing size matches, a score for performance in the size-matching task was required. The performance score was gained by subtracting the mean (that is textured and white combined) perceived waist-hip ratio at 120mm of disparity from the mean perceived waist-hip ratio at 0 disparity, for each participant. There was, however, no significant correlation between the estimated distance of background at 0 disparity and the performance score, $r = 0.053$, $n = 21$, $p>0.05$.

To ascertain whether participants found the preliminary card-ordering task more difficult in either the textured or white conditions, a paired comparison was carried out. The mean error score for the textured condition was 2.38 (SD = 1.96), and for the white condition the mean error score was 2.67 (SD = 2.71). Statistically there was no significant difference between these two mean error scores, $t(20) = -0.404$, $p>0.05$. The performance score was then utilised again to find out whether participants gaining a high number of errors in the card-ordering task performed any differently in the size-matching task from those participants with lower error scores. No evidence was found, however, for such a relationship, $r = -0.137$, $n = 21$, $p>0.05$. . In the card-ordering task, the presence, or lack, of texture on the image seems to have little, if any, affect upon the amount of errors made by the participants. Also, there would appear to be no relationship between performance in the card-ordering task and performance in the size-matching task.

The 'TNO test for stereoscopic vision' and 'The Frisby Stereotest' are both designed to measure stereoscopic acuity, on which the size-matching task is based. To find out whether there was any relationship between the scores gained on the separate tests and performance in the size-matching task, a correlation was determined between both the TNO and Frisby tests, and the performance score. The obtained correlation between the TNO test and the performance score was not significant, $r = 0.228$, $n = 19$, $p>0.05$. However, the correlation gained between the Frisby test and the performance score showed that there was a statistically significant relationship between the two, $r = -0.455$, $n = 21$, $p<0.05$ (2-tailed). Figure 5.29 shows the relationship between the Frisby test results and the performance score. Thus, though the TNO test is not a good predictor of performance in the size-matching task, the Frisby test can predict performance.

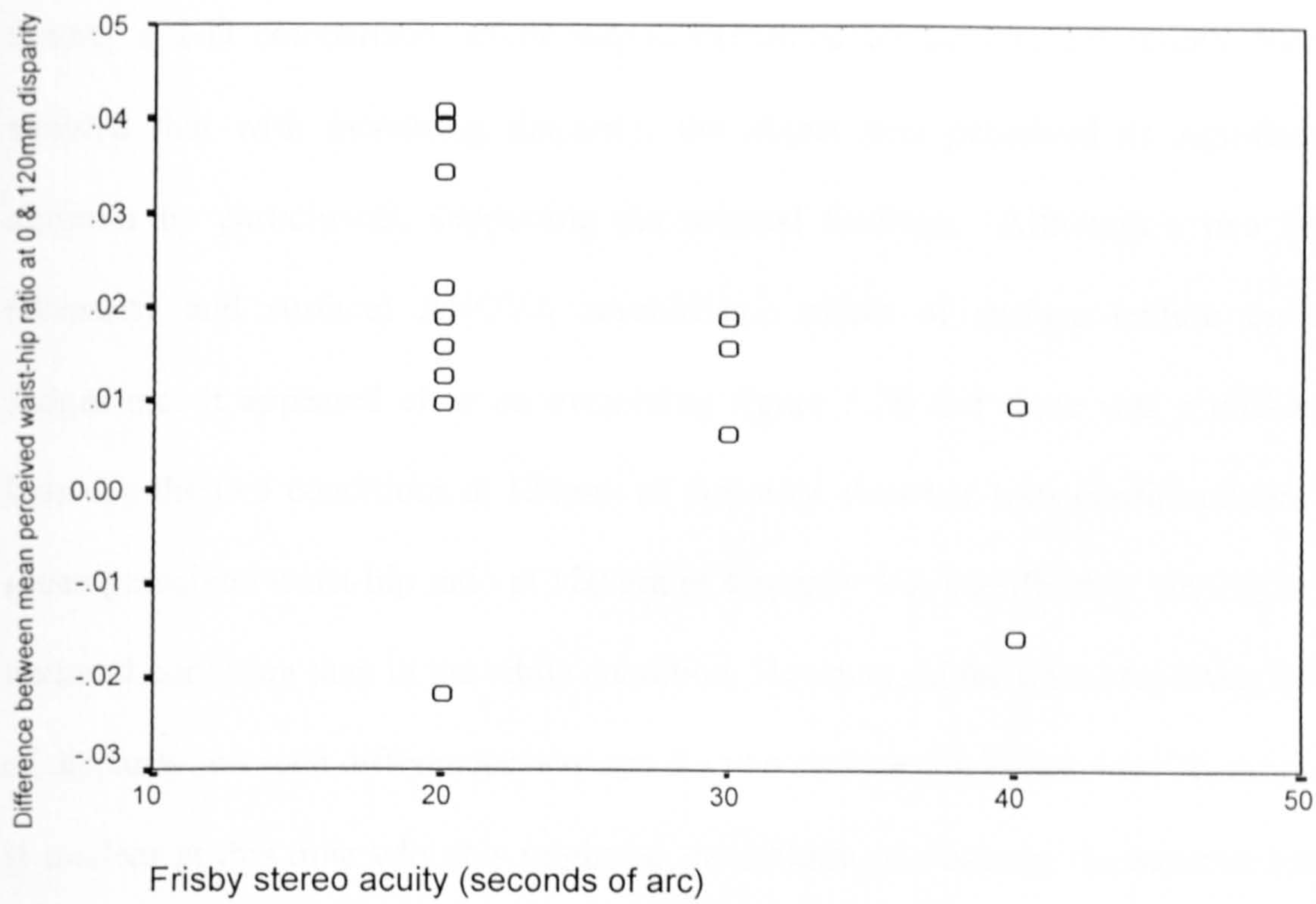


Figure 5.29. Results of 'The Frisby stereotest' for each participant, plotted against a score gained from subtracting the perceived waist-hip ratio at 120mm of disparity from the perceived waist-hip ratio at 0 disparity (performance score) for each participant

Finally, to establish whether gender had any effect in the size-matching task, a three-factor (disparity, surface, and gender) mixed ANOVA was carried out. No significant interaction was found between disparity and gender, $F(3, 57) = 0.578$, $p>0.05$, or between surface and gender, $F(1, 19) = 0.153$, $p>0.05$. There would seem to be no difference between the sexes in their perception of 3-D stereoscopic information.

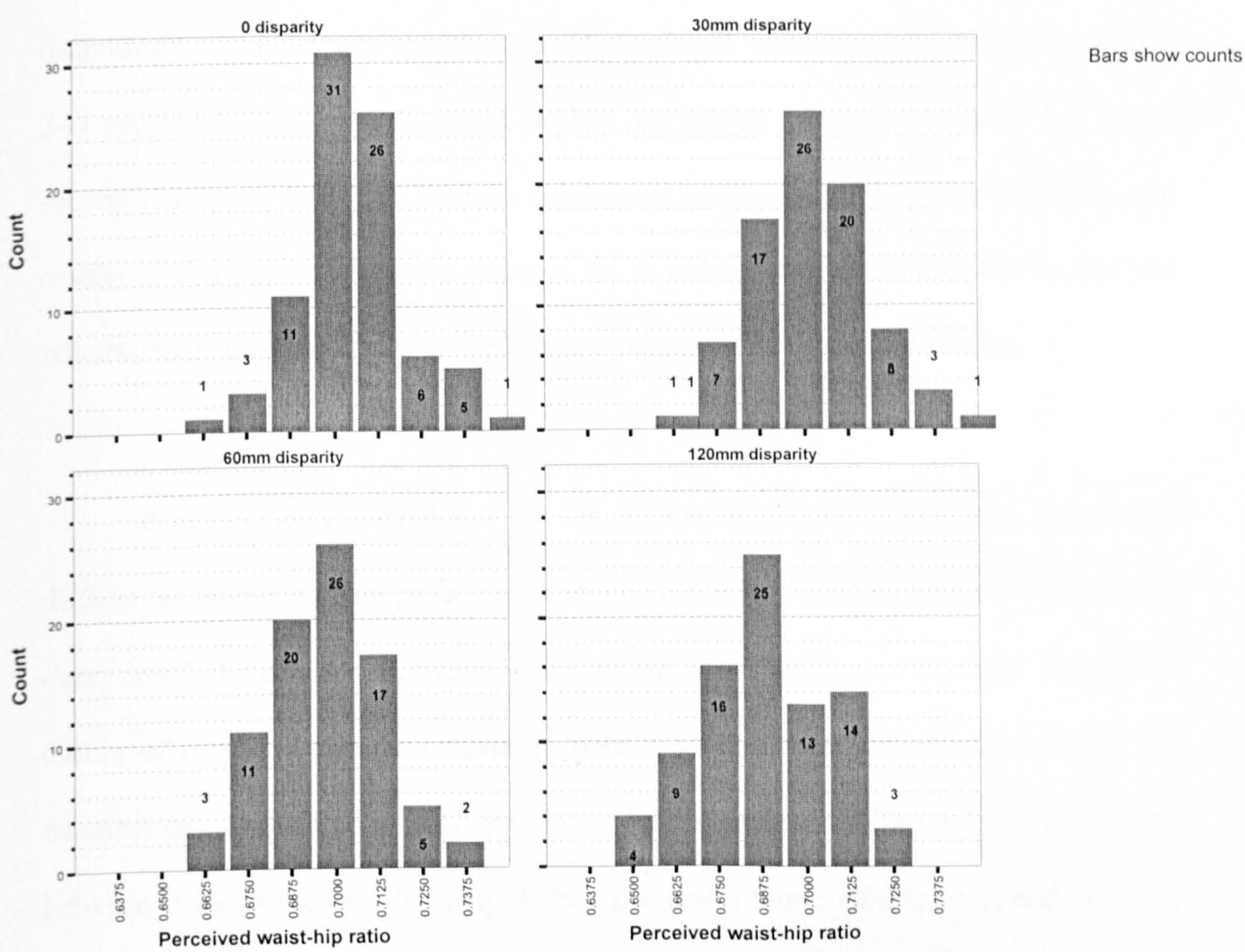
5.5.6 Discussion

The second peanut experiment investigated size perception as a function of disparity (interaxial separation) and surface texture in 3-D images. A 3-D peanut-shaped computer model was generated and rendered for polarized projection. Identical images were created in a series of widening disparities with 0, 30mm, 60mm, and 120mm of virtual interaxial separation, and were projected orthostereoscopically onto a screen. The surface of the object was rendered either with or without texture. Participants were presented with each image and were asked to find a match from among a 2-D comparison set of stimuli presented on an adjacent screen. Results showed that with increasing disparity, the object was perceived as significantly slimmer by participants, supporting the original findings. Although a two factor (disparity and surface) ANOVA revealed no effect of surface texture on size judgement, it appeared clear on examining figure 5.26 that there was a difference between the two conditions at 120mm of disparity. Post-hoc tests confirmed that the mean perceived waist-hip ratio at 120mm of disparity was significantly smaller in the textured condition than in the white condition. However, in the three remaining levels of disparity, no such differences between the two surface conditions were found and it is unclear at this time why this might be. At 120mm of disparity the separate image channels may have been difficult to fuse for some participants. Perhaps the presence of texture, and the resulting greater number of stereo cues present on the object, aided this fusion, and caused participants to choose smaller matches.

In experiment 1, the zero disparity stimuli were matched correctly to the corresponding 0.7 waist-hip ratio comparison images by more than half of the participants. The present study found that only 37% of matches were correct at zero disparity. Although no significant effects were found relating surface texture to size estimations, it is interesting to examine the frequency distributions in figure 5.25 in which subtle differences in spread between the two surface conditions can be seen. In the textured condition at zero disparity, for instance, 43% of matches were correct, compared to 31% in the white surface condition. In fact, a slightly higher percentage (33%) chose a fatter 0.7125 waist-hip ratio match in the white condition. Despite these differences, the means of the two conditions were very similar (0.7060 and 0.7049 for the textured and white conditions respectively) with the bulk of the data lying towards the right hand side of the graph (the "fatter" end). By combining the data from these two conditions, it was found that the mean perceived waist-hip ratio of the group at zero disparity was 0.7054 (i.e. larger than the actual 0.7 waist-hip ratio of the image). This finding was expected, even though as the earlier study had reported an average of 0.694, which is just below the actual waist-hip ratio of the object. The result is consistent with previous findings that synoptic viewing can give larger than life size estimations.

The discrepancy could also be linked to optical disparity cues that were present in the original peanut stimuli but were taken into account and controlled for in the present study. At 30mm of disparity the mean perceived waist-hip ratio of the combined data was significantly less than at zero disparity (0.7015) and was closest of all the levels of disparity to the actual waist-hip value, despite still being slightly

Figure 5.30. Combined frequencies of perceived WHR at 0, 30mm, 60mm, and 120mm of disparity



larger. In figure 5.30, a definite peak can be seen at 0.7, yet the data is still clustered towards the right hand side of the graph. It is not until 60mm of interaxial separation that the average drops below 0.7 to 0.6973. As virtual interaxial separations widen from 0 disparity, through 30mm of disparity, on to 60mm of disparity, it is clear that the "weight" of the data is moving from right to left – from the fat end of the graph to the slim end – though dominant peaks occur at 0.7 in each case (figure 5.30). The mean perceived waist-hip ratio drops still further to 0.6874 at 120mm of disparity. However, at this point the data becomes more widely spread (figures 5.25). In figure 5.30, although there is a dominant peak at 0.6875, an emerging peak can be seen at 0.7125, which could indicate a bimodal spread. Because of the large interaxial separation, twice that of the average human inter-ocular distance, some participants may have been unable to fuse the two images and diplopia, or the suppression of one image, may have occurred (Howard & Rogers, 2002). If this is so, then these individuals will have seen the object as if viewing it with only one eye - effectively a 2-D representation - which would explain the larger matches that have led to the second emerging peak at 0.7125. In subsequent studies, it might prove useful to gain results at 90mm of disparity in order to fill in the missing data point and to find out whether this separation would yield similar results to the 120mm disparity.

Post-hoc tests carried out on the data revealed that there were significant differences between mean perceived waist-hip ratios (textured and white combined) at every level of disparity. Between 0 and 30mm of disparity, and between 30mm and 60mm of disparity, the differences in perceived size were statistically smaller, but nonetheless, a difference was present. More significant differences were found between 0 and 60mm of disparity, between 0 and 120mm of disparity, and between

60mm and 120mm of disparity. However, despite highly significant differences between the mean perceived waist-hip ratios at 0 and 120mm of disparity, and between 60mm and 120mm of disparity, the use of 120mm of interaxial separation in stereo image capture may cause some viewers to shut down one image channel, thus effectively making their view a 2-D one. As in previous studies (Yamanoue, Okui, & Yuyama, 2000), the present findings support the use of interaxial camera separations similar to the human inter-ocular distance. In agreement with other published studies in this area (Yamnoue, 1997), results show that as disparity increases so too does the perceived distance of the random dot background in the polarized images, with significant differences between each level of disparity. Figure 5.28 shows a relatively uniform increase in the mean estimated distance as the disparity increases. The spread of the data, too, remains consistent in the 0, 30mm and 60mm levels of disparity. At 120mm of disparity the spread of data dramatically increases, which again could be linked to diplopia in some participants.

Because a zero disparity image is effectively 2-D and therefore conveys no real depth information, the background of the image at this level of disparity should be on the screen (i.e. 0m). Some viewers, nevertheless, saw the background very slightly behind the screen. Participants who saw the background as being behind the screen at zero disparity might therefore have performed differently in the size-matching task to participants who saw the background on the screen. No correlation was found, however, between the estimated background at zero disparity and the calculated performance score for the size-matching task. Error scores gained from the card-ordering task revealed little difference between the textured and white cards. Mean error scores did not differ significantly from each other. It was hypothesized that participants gaining a high number of error scores overall might perform differently on the size-matching task from those with lower error scores. This would not appear to be the case, however, as no correlation was gained between the combined mean error score and the performance score for the size-matching task. It would seem that surface texture has no effect upon the amount of errors made by participants in the card-ordering task, and that combined errors cannot predict performance in the size-matching task. However, the card-ordering task is essential in preparing the participants for the main experiment (the size-matching task) so cannot be omitted. Pilot studies have shown that results gained from the size-matching experiments are far more variable if the participants do not undertake the card-ordering task beforehand.

Experiment 1 found no correlation between participants' stereoacuity measured by the 'TNO test for stereoscopic vision' and performance in the size-matching task[4]. The present results support this finding. However, a significant correlation was found between the performance score and stereoacuity measured by 'The Frisby stereotest'. Figure 5.29 shows the relationship between these two variables. How can two tests that measure stereoacuity yield such wildly different results? The TNO test uses a parallel stereo-image capture technique for its random-dot anaglyph plates. These anaglyph disparities are rendered to indicate the limit of a subject's ability to fuse red-green "double images". Experiment 1 had found that parallel image capture geometry produced far more varied results than convergent image geometries. Also, in individuals who suffer from red-green colour blindness the test would be useless. The Frisby test, by contrast, uses real depth cues. The test consists of square plates of Perspex of three different thickness. On one side of each plate are printed four

[4] In the first study a performance score was not employed. Instead, stereoacuity was correlated with perceived waist-hip ratio in the 65C condition (65mm of interaxial separation with converging lenses).

seemingly identical patterns, which are located in the four sections of the square. However, one of these patterns has a hole in its centre, and only appears to be the same as the other three patterns because of a printed circle on the other side of the Perspex, which is visible through the hole. The circle therefore appears to either protrude outwards or recede back from the pattern, depending on which way the test is oriented. The task for the participant is to pick out the pattern containing the circle. As the plates become thinner, the more difficult this task becomes. Stereoacuity is calculated from the distance at which the participant can no longer see the circle and at what particular plate thickness this occurs. The present study always oriented the plates so that the circle could be seen to recede back behind the main pattern (preliminary trials showed that this might be easier for the participants, even though the instructions for the test recommend that the experimenters should simply use a consistent orientation). If the plates were to be oriented randomly so that the circle could be seen to protrude outwards from the pattern as well as recede, even more significant results might be gained because a more wide and varied selection of stereoacuities would result. It would appear, then, that the TNO differs considerably from the Frisby test in some important respects. The Frisby stereotest is a good predictor of performance as it relies on real depth cues and principles closely related to those used in orthostereoscopic projection.

Finally, a difference between genders was considered. It was suggested that perhaps females would be better at perceiving the differences in waist-hip ratio than males. Females are often more aware than males of beauty, fashion and body shape etc, and (it is popularly believed) are more concerned with comparing themselves to others. The results of the present experiment, however, found no difference between the sexes in their perception of 3-D stereoscopic information. This may have been due to the very small gradations used in this experiment, and indeed the relatively small sample, though Menon, Bansal, & Prakash (1998) also found no significant effect of sex on stereoacuity with 20 subjects

In summary, the results confirm that increasing disparity results in smaller size-matches. Although texture appears to have no overall effect on size perception, at 120mm of disparity there were significant differences between the two conditions, with smaller matches for the textured stimulus. The reasons for this are unclear but may be related to object surface stereo cues aiding fusion of the two image channels. The study found that with increasing disparity the perceived distance of the background retreated further behind the screen, but that estimates of background distance at zero disparity did not correlate with the performance score for the size-matching task. No differences were detected between the error scores gained from the card-ordering task, nor was a correlation established between the combined mean error score and the performance measure.

Stereoacuity measured by the 'TNO test for stereoscopic vision' did not correlate with performance in the size-matching task. However, performance in the size-matching task was found to correlate with stereoacuity measured by 'The Frisby stereotest'. No effects of gender were revealed. The study therefore supports the use of orthostereoscopic techniques for stereo image capture and projection, using interaxial separations close to the inter-ocular distance of the human eyes (60-65mm). Results also point to the use of 'The Frisby stereotest' instead of the 'TNO test for stereoscopic vision' for predicting stereo performance, probably because it uses the perception of actual depth targets. This confirms Heron et al (1985 cited in Rogers and Howard, 2002 p.147) who found that the Frisby test had the best test/retest reliability of the available clinical tests.

CHAPTER SIX

EXPERIMENTS 5 & 6

The Effect of Changes in Focal Length on Perceived Body Weight and Age

6.1 Experiment 5: Portraits made at varying camera to subject distances I

6.1.1 Introduction.

The following two experiments use orthostereoscopic imaging (see chapter 4) to investigate the effect of changes in lens focal length coupled with changing camera to subject distances on perceived body weight in photographs of people *using head and neck images only.* Experiment 5 was reported in the MIT Press journal, Presence: Teleoperators and Virtual Environments. Entitled *Cyclopean Vision, Size Estimation and Presence in Orthostereoscopic Images by Bernard Harper and Richard Latto,* it appeared in Presence Vol. 10, No. 3, June 2001. (A copy is included in appendix A)

The primary questions addressed in this section are how to revise the method used previously to demonstrate statistically that the fattening observed was a real effect. The 2D pilot study suggested that perception of body size in 2D photographs varied in relation to the focal length of lens used. However, it was realised that the fattening effect observed (with the focal lengths of 50mm and greater) was not caused by distortions of the lenses themselves. The lenses used were optically corrected to ensure that curvilinear and anamorphic distortions of shape were kept well below observable levels. The lenses most commonly used in photography are longer than 50mm in focal length and tend to encourage portraiture and figure study photography to be at a distance of greater than 2 metres. So it is camera to subject distance that is the key variable causing the fattening effects of the most commonly used lenses, rather than the design of the lenses themselves. In figures 6.1 a & b, (overleaf) most of the portraits were taken with telephoto lenses from distances where the model's face could be obscured by a thumb held up at arms length. Yet when one holds a photograph taken at the same distances between the finger and thumb, one is never aware from the perspective that the printed image subtends a visual angle that is much lager than the models face. This means that a models faces is always conveyed with inadequate information in 2D photography. In the perception of real close-up objects, there is more information than the perspective, proximity and magnification cues than can be conveyed by a 2D photograph. *Direct vision would normally have stereoscopic depth and monocular cues to depth (with accommodation, saccadic visual angles and motion parallax) at these distances to produce an accurate perception of size.*

The vari-focal feasibility study also indicated that the focal length of lenses normally used in portraiture did strongly correlate to perceive bodyweight. In a follow-up study, the eleven photographs taken in the first pilot study (from super wide-angle 19 mm lens close-up portraits out to telephoto images taken with a 400mm lens over five metres away) were re-used in a new way. The printed images were sized so that the models inter-pupilary distance was approximately the same on each print. The prints were trimmed to identical sizes and then randomised. The participants were asked to place these images in order of slimmest to fattest. When the task was completed (normally taking about 1-2 minutes for each of the two models) the photographs would be turned over to reveal that the order selected correlated almost identically to the increases in camera to subject distance and its related focal lengths. (figure 6.1)

Figure 6.1a.
11 portraits taken with these lenses and distances:

19mm lens, 32cm distance
24mm lens, 45cm distance
28mm lens, 55cm distance.
35mm lens, 67cm distance.
50mm lens, 85cm distance
80mm lens, 118cm distance
100mm lens, 154cm distance
135mm lens, 205cm distance
200mm lens, 298cm distance
300mm lens, 418cm distance
400mm lens, 538cm distance

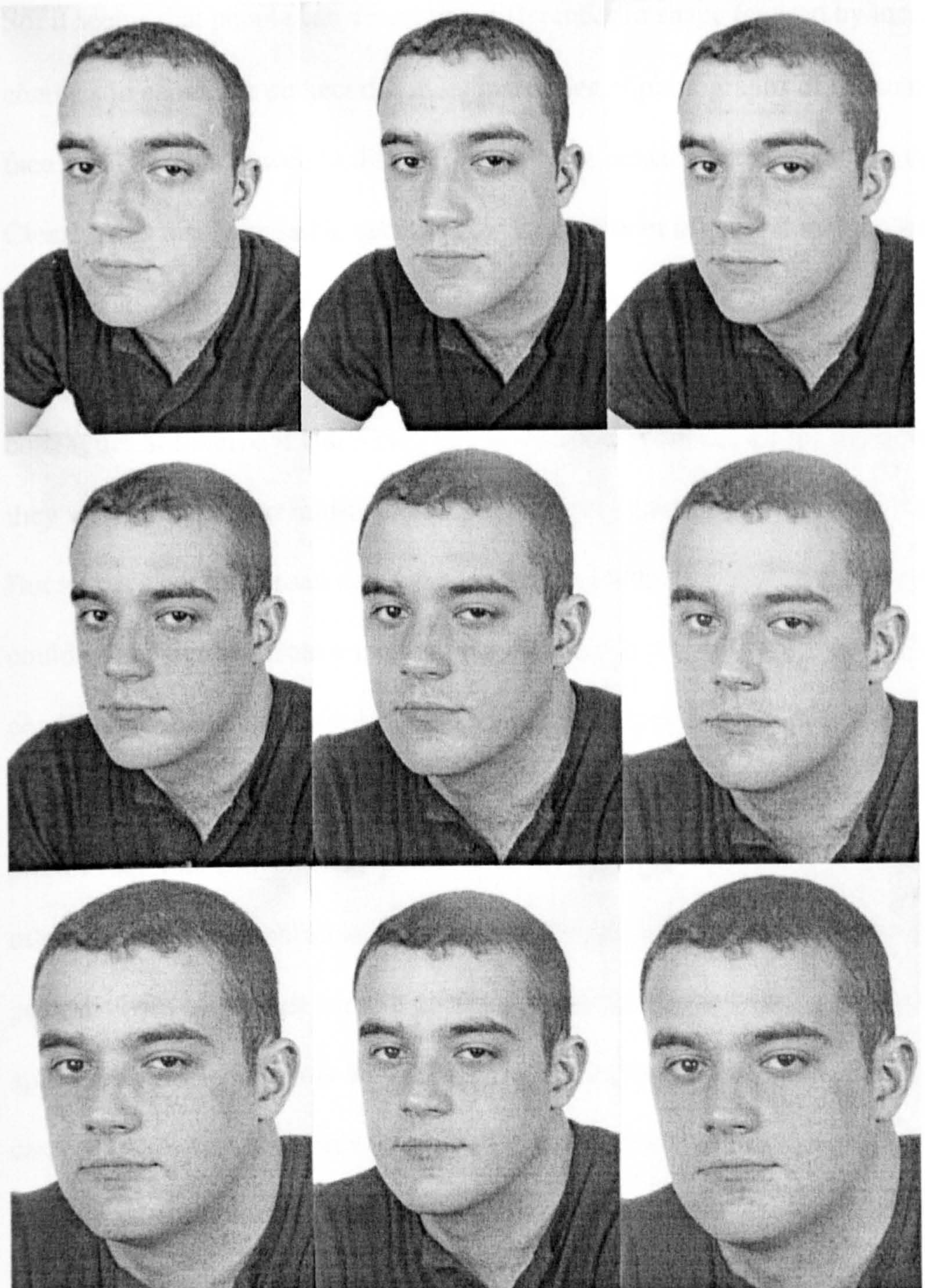


Figure 6.1b.
11 portraits taken with these lenses and distances:

19mm lens, 32cm distance
24mm lens, 45cm distance
28mm lens, 55cm distance.
35mm lens, 67cm distance.
50mm lens, 85cm distance
80mm lens, 118cm distance
100mm lens, 154cm distance
135mm lens, 205cm distance
200mm lens, 298cm distance
300mm lens, 418cm distance
400mm lens, 538cm distance

So, it seems that people can detect tiny differences in shape (caused by incremental changes in camera to subject distance) in a series of photographs of the same human face and infer a bodyweight that correlates to the actual camera to subject distance. Clearly, this task is possible because the visual system is good at making small incremental differentiations in photographs of otherwise identical human stimuli. The vari-focal pilot study featured images of one male and one female model. Both were colleagues at Liverpool University and were asked to model for the study because they were both in their mid-twenties and of approximately average weight and height. But it was clear that broad conclusions about the fattening effect of photography could not be derived from a method that included just two subjects. A further confound to using these images in any experiment was the ¾ "semi-profile" portraiture technique used. This technique was adopted as an attempt to make the photos look like conventional portraits, rather than the type of full-face "mug shot" often used in psychophysical studies. However, the super wide lenses used for the portraits with the closest camera proximity gave the faces an unusually stretched appearance that could only be corrected by photographing the models face on to the camera (Appendix H figures 6a & 6b). However, most photographs of people rarely offer viewers a choice of proximities, or a known object in the image that one can scale the target against to make a more accurate size estimation. So it was important that in any experiment to investigate the effects of different camera to subject distances, the participants must only see one photograph of one model at one focal length and to not be able to compare the face with any other image of the same person.

To gather enough data from one participant, a set of photographs were presented to the participants that included images from a wide range of camera to subject distances, with a different person in each image. For instance, if five models were photographed at five different camera to subject distances, then one would make up five sets of photographs, each one with a five different combinations of faces at different focal lengths. By ensuring that the same person is never seen more than once, the underlying effect of camera to subject distance can be identified separately from any individual characteristic of the models.

The first vari-focal study used two people photographed with eleven different lenses, giving twenty-two images in all. To incorporate more models with this number of lens changes would be an enormously complicated exercise both to photograph and present the images in an experiment. It was decided to reduce the number of photographic conditions to five camera to subject distances. These were taken with an extreme wide-angle lens (19mm) at 0.32 metres distance, a wide-angle lens (35mm) at 0.45 metres, a standard lens (50mm) at 0.71metres, a short telephoto lens (100mm) at 1.32 metres and a telephoto lens (200mm) at 2.7 metres. The lighting was the same as in the first stereoscopic photography and was used for the same reasons of consistent, shadow-free illumination with similar brightness for the foreground and background planes.

Six models (three male, three female) within an age range of 20 to 23 would be the minimum required to get useful data for this study. The choice of model size was important, so one male and one female was found for each of three groups: A correct weight group, a slightly underweight and slightly overweight group. The criterion used to ascertain which group a model should be placed in their Body Mass Index (BMI). Each model was weighed and their height measured. The figures were then compared to a BMI Ready Reckoner chart obtained from the dieticians at the Royal

Liverpool Hospital. By using weight and height to derive a range of useful figures, the chart indicated that the BMI figure for the "correct" weight pair would be 22. An approximate figure of BMI 26 would indicate a person on the borderline of being overweight, and a figure of approximately BMI 19 would indicate a person on the borderline of slightly underweight. Six models of the required BMI range and ages were identified from students of Liverpool University and asked to attend a single photographic session in the Department of Psychology. They were to be paid for their participation and informed that the photography was a one-off event (due to the difficulty in arranging for the required individuals and equipment to be available at the same time) and could not be re-staged.

The photography for the vari-focal lens experiment was completed in a large room in that could black-out all daylight. This was so that the lighting was solely from the flash units and was consistent from shot to shot. A white background was chosen, as the professional "cloudscape" photographic background used previously could give distance and size cues from its patterns of colour and brightness that would be distinct to each of the lenses used. A ceiling height of over three metres was required to hang the white photographic background on the end wall, as the wide-angle lenses would require a far larger area of uniform white behind the subject. A chair was placed close to the background and on a centre line marked on the floor. It ran at 90 degrees from the plane of the background and was marked with the distances in centimetres. These were set to run from a vertical marker that was used to align the face of each model so that each face would lie in the same plane of sharp focus at every distance. This technique was used to avoid changing the focus (and therefore the magnification) with each model as they were posed in the correct spatial position.

Image 1. 19mm super-wide lens
Subject distance: 0.32 metres

Image 3. Standard 50mm lens
Subject distance: 0.71 metres

Image 5. 400 telephoto lens
Subject distance: 2.7 metres

Image 2. 28mm angle lens
Subject distance: 0.45 metres

Image 4 135mm telephoto lens
Subject distance: 1.32 metres

Figure 6.2a Shows five of the twenty five images used in experiment 5. Each image was printed to 5-7 inch cards and was scaled to retain the same magnification (measured by the models inter-ocular distance) in all photos in each group.

Image 1. 19mm super-wide lens
Subject distance: 0.32 metres

Image 3. Standard 50mm lens
Subject distance: 0.71 metres

Image 5. 400 telephoto lens
Subject distance: 2.7 metres

Image 2. 28mm angle lens
Subject distance: 0.45 metres

Image 4 135mm telephoto lens
Subject distance: 1.32 metres

Figure 6.2b Shows another five of the twenty five images used in experiment 5.

6.1.2 Method.

6.1.2.1 Stimuli. Two males and three females were photographed in identical poses using zoom lenses in a series of five focal lengths from a wide-angle to telephoto range. The method made it possible to record the facial features of each model to the same magnification and resolution at the film plane from distances of 0.32 m, 0.45 m, 0.71 m, 1.32 m and 2.70 m. Prints were made from the portraits and five sets were made up, each containing one photograph of each model at one of the five focal lengths. Examples from the range are shown in Figures 6.1 & 6.2

As photographic set-ups are very difficult to replicate if the session is interrupted, it had been decided that there would be one attempt to take all of the photographs with this group of subjects. However, one model (the slightly overweight male) did not appear and so it was decided to complete the photography with the five people who did manage to arrive at the correct time. Attempts to re-shoot these photographs with a complete group were unsuccessful. This meant that the experiment would need to be run with five sets of photographs, each with five different people portrayed at the five different camera to subject distances. It should be noted that quoting focal lengths in millimetres can be misleading. Lens calibrations can offer different image magnifications depending on the camera used. For instance, a 50mm lens on a 35mm SLR is a standard lens. On a 6x6 camera it is a wide-angle lens. On a video camera it would be a telephoto lens. For Experiment 5, the independent variable reported is therefore camera to subject distance while maintaining a same-size image, since this is repeatable regardless of the camera system or lens design used.

... The prints from these images were trimmed to the same size and separated into the five separate groups. The proposed presentation method was to show only five the photos to each of 25 participants. Each photo was examined independently of the others and the participants' task was to apply one of the descriptors in the seven point Likert scale as used in experiment 1 (and copied below).

Descriptor	Abbreviation	Likert value
VERY OVERWEIGHT.	V.O.	7
OVERWEIGHT.	O.	6
SLIGHTLY OVERWEIGHT.	S.O.	5
CORRECT	C.	4
SLIGHTLY UNDERWEIGHT.	SU.	3
UNDERWEIGHT	U.	2
VERY UNDERWEIGHT	V.U.	1

However, on this occasion the participants were asked to write the number (next to the appropriate descriptor) on a response sheet that most approximated their estimation of the bodyweight of each of the models. They were encouraged to not feel "kind or considerate" of the feelings of the people portrayed and use the extremes of the range if the felt it was necessary. They were also encouraged to act as "visual weighing scales" and impartially record which descriptor conveyed the most accurate size estimation.

6.1.2.2 Participants. Twenty Liverpool University undergraduates were tested individually.

6.1.2.3 Procedure. One set of five photographs of five individual models was shown to each of four groups of five participants. Unlike the examples in Figures 6.1 & 6.2, they were never shown the same model photographed at more than one focal length. Each participant was asked to place the five different model portraits in a rising order of apparent body weight using the same seven point Likert scale as in Experiment 1 (see Section 3.1.1.3), and to apply a number from 1 to 7 to each image. A number higher than four was given to people who appeared to be overweight and numbers less than four to people who appeared to be underweight. The most overweight was to be given a score of seven, the most underweight a score of one.

6.1.3 Results.

Figure 6.3 shows that as camera to subject distance (and focal length) increases, a higher score was given on the Likert scale ($r = 0.824$, $N = 5$, $p < 0.05$, one-tailed). A one-factor (camera to subject distance) ANOVA found an overall effect of distance on size judgement ($F(4,76) = 8.858$, $p < 0.001$). Planned comparisons of the scores at each distance with the mean score at 0.45 metres, using two-tailed t-tests, showed that the wide-angle, close proximity images (0.32 m) showed underweight estimations ($t(19) = 4.073$, $p = 0.001$). The standard lens image (0.71 m) showed a slight but not significant overweight estimation ($t(19) = 1.097$, $p = 0.287$). The telephoto distance images (1.32 m and 2.7 m) showed overweight estimations ($t(19) = .2.101$ & 5.101, $p = 0.049$ & < 0.001).

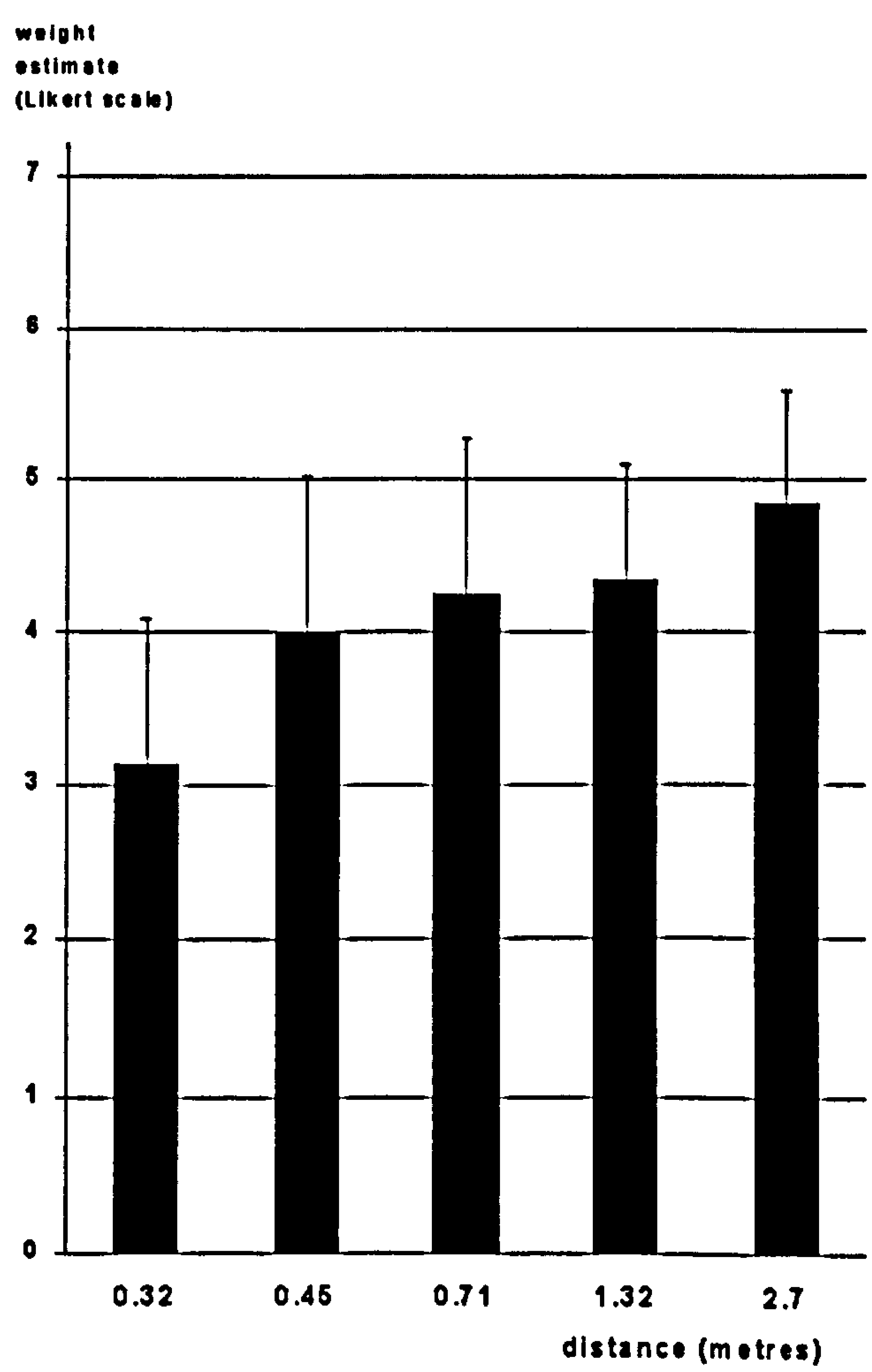


Figure 6.3. The mean perceived body weight for the five different camera to subject distances (in metres), and therefore five different lens focal lengths, used to photograph the models in Experiment 4.

6.1.4 Discussion.

Because of the limitations of the photographic location and lenses available, it was not possible to test if extending the range of focal lengths would show a continuing positive relationship between focal length and perceived bodyweight. It is likely however that the focal lengths used in this experiment cover the range where the strongest effects could be demonstrated. Extreme wide-angle distortions at one end of the scale and proportionally smaller changes in the depth compression of telephoto lenses at the other would probably act to curtail the extent of the effect. This would mean that for images made at less than 0.3metres, the slimming effect observed would be countered by the "ballooning" distortions of the face caused by the widest angle lenses (see Appendix H figure 7). Conversely, portraits taken from further away than 2.7 metres may have very small geometric changes at the increments possible with conventional lenses. It may be that one would need to double the distance with each new condition (2.7 m, 5.4m, 10.8m etc) and these would require photographic lenses that were astronomical in both their normal use and cost. Television zoom lenses however can give these magnifications, but traditionally have done so at the expense of resolution and curvilinear distortion.

6.2 Experiment 6: Portraits made at varying camera to subject distances II

6.2.1 Introduction

The first vari-focal experiment seemed to reveal three effects not observed previously. As expected, the fattening effect of telephoto lenses and long camera to subject distances was well supported by the data. It was however a surprise to find that there was also a significant slimming effect in portraits taken with wide-angle lenses (Figure 6.3). Also, when viewing the whole set of prints, there seemed to be an obvious gender effect that had not been noticed previously. It appeared that the fattening effect of telephoto lenses seemed to be larger on the female models than the males (figures 6.1 & 6.2). A final effect that was noticed is that the apparent age of the subject appeared to rise with the longer focal length lenses.

Post- experiment 5 discussions revealed an important point which had not been previously considered: Some respondents and colleagues were unsurprised at the fattening effects demonstrated. This group seemed to regard photography and portraiture as innately untrustworthy and an unreliable method of recording the human face and form. Interestingly, this group seemed to be almost all female. To ascertain if this was merely anecdotal or an actual gender bias, these opinions were to be polled in a pre-experiment study with a simple forced-choice paradigm: Two very similar statements (below) were to be printed on the response sheet and participants were asked to underline which one agreed most closely with their own opinion.

"Photographs are generally a reliable and accurate method of recording the appearance of people and normal scenes."

"Photographs can be an unreliable and inaccurate method of recording the appearance of people and normal scenes."

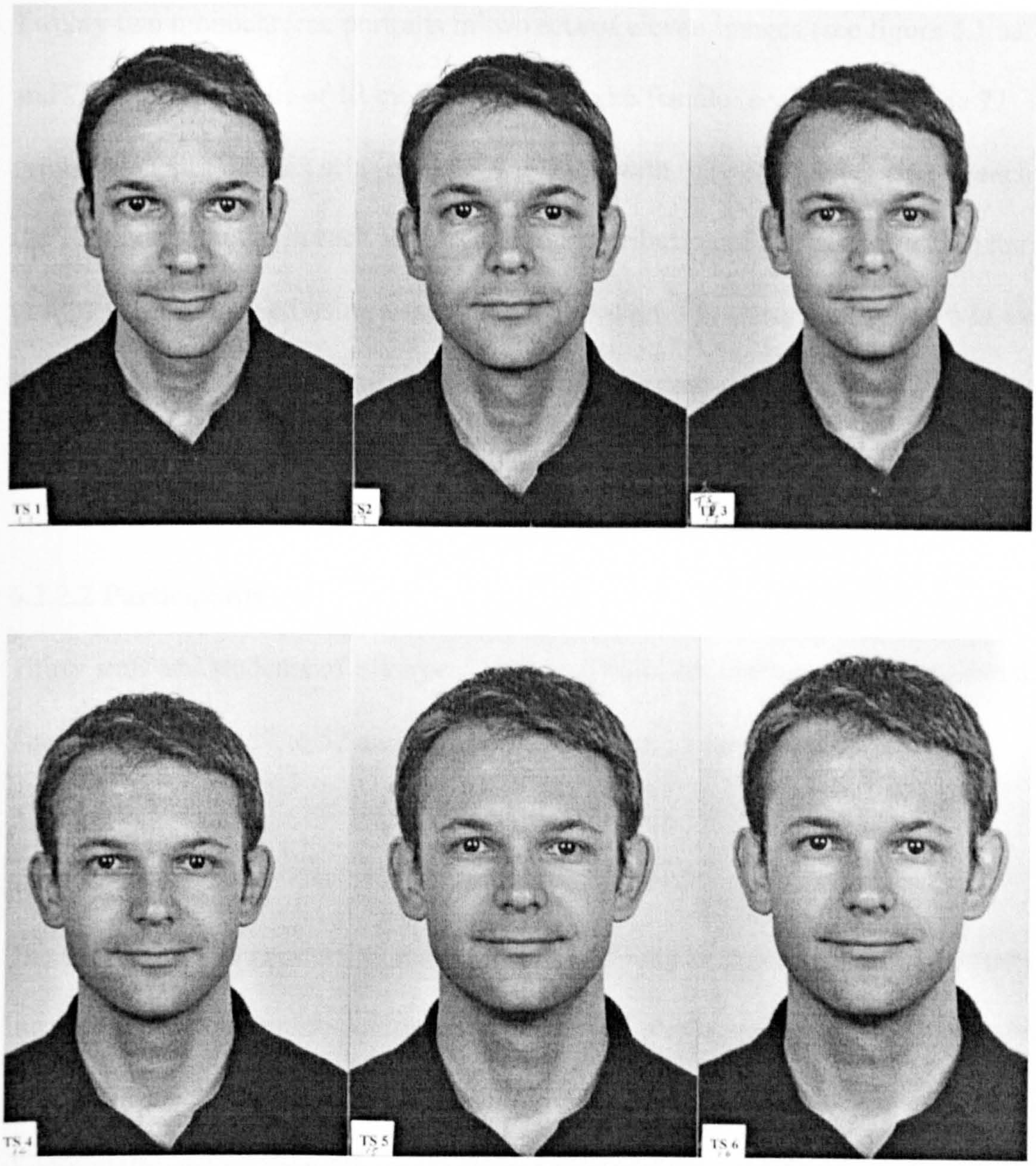
Experiment 6 was designed to confirm the result of the previous experiment and gather more data about the new observations and to quantify the fattening effects, so that a chart of perceived weight gain in kilograms for each of the lenses could be derived from the data. Clearly, the five models in the first experiment (three female, two male) was too small a group to extract reliable data about gender effects. And as their ages were not recorded and their heights were self-reported rather than measured, it was decided that a completely new set of photographs, models and measurements were required. It was estimated that the minimum number of people required to gather gender data in a similar experiment would be six males and six females, all of approximately the same age. Additionally, it would be desirable to find pairs of people that were slightly underweight, correct weight and slightly overweight to establish if the effects observed were generalisable across a range of normal body sizes. The effect of apparent fattening or aging and camera to subject distance would also be investigated using an additional condition: The furthest camera to subject distance used in the first experiment was 2.7 metres. However, this is still well within the normal range by which professional portraiture and moving images are usually made. If a new condition at double this distance was investigated, it would include almost all of the normal distances used in generating professional still and moving images.

The photography for the new stimuli was completed with a method similar to the previous study. However the camera to subject distances were revised to include more telephoto lens images and the new furthest distance. The revised distances (previous distances in brackets) are: 35cm (32cm), 70cm (45cm), 100cm (71 cm), 140cm (132cm), 280cm (270cm) and 560cm. Fourteen models were invited to be photographed. The larger than required number of twelve models was to allow for a small number of dropouts and the option of rejecting any models images for unforeseen reasons. Thirteen of the people who had been invited previously (and had an approximate Body Mass Index to the requirements for the experiment) arrived to be weighed, measured for height, record their age/gender and be photographed in each of the conditions. The models' BMI was calculated and confirmed as being four of them being slightly underweight, five normal-weight (three female, two male) and four slightly overweight, before the photography began. This was done as it was important to establish that a complete set of male and female BMI examples were available for the session without requiring the photography to be repeated for missing individuals at a later date.

The film used was a high-resolution professional quality 35mm colour negative material (Fuji NPS 160) and the images were all processed and printed on the same printing machine and paper. This ensured that the image colour varied by the smallest possible amount and that the print magnifications were the same for all of the stimuli. Any variation in colour could then be attributed to small differences between the design of each camera lens, and any variation in magnification would be due to slight differences in facial position between each of the models. If the photography had been split into two or more sessions, the opportunity to maintain control over these variables would have been lost because film and print processing varies in quality and optical constancy on a day-to-day and printer-to-printer basis. The technique adopted gives the least variation in image quality between each of the images and reduces the likelihood of uncontrolled variables in the stimuli. The stimuli were placed in order of camera to subject distance for each model and the variation in print quality was assessed as being small enough to be unnoticeable by an untrained viewer. At this stage, one of the three sets of normal-weight female images was discarded by a random draw to ensure that there were equal numbers of each model in all of the likely presentations. The photographs were then ordered into six sets of twelve images. Each set contained one randomly chosen photograph of each of the twelve models. Each of the six distances were therefore represented by two models in each set. A BMI matched male and female set is reproduced in figures 6.4 a & 6.4 b.

Figures 6.4 a

Shows all six portraits of a model photographed from the revised distances (previous distances in brackets) of : 35cm (32cm), 70cm (45cm), 100cm (71 cm), 140cm (132cm), 280cm (270cm) and 560cm.



Figures 6.4 b

Shows all six portraits of a model photographed from the revised distances (previous distances in brackets) of : 35cm (32cm), 70cm (45cm), 100cm (71 cm), 140cm (132cm), 280cm (270cm) and 560cm.

6.2.2 Method

6.2.2.1 Stimuli.

Twenty two monochrome portraits in two sets of eleven images (see figure 6.1 a&b). and 72 colour portraits of 12 models, 6 male and 6 female (see 6.4 a&b). The 72 prints were split into six sets (marked A to F) of with only one photograph of each of the 12 models shown in each set. A balanced distribution of the images within the six groups was established using a random block design. The design was chosen to ensure that each set of 12 portraits would contain a representative distribution of genders and weights at all six camera to subject distances.

6.2.2.2 Participants

Thirty staff and students of Liverpool University and six visitors were interviewed. The age range was 17 to 52 and the gender split was 15 males and 21 females.

6.2.2.3 Procedure

The procedure was necessarily more complicated than in experiment 5. In addition to the new pre-study opinion poll on the reliability of photographic portraits, there were now six different camera to subject distances and twelve model pictures to be judged. Additionally, the participants were required to assess the age of each of the models and estimate the weight for an individual model.

. After recording their name, age sex and basic eyesight details (20/20, short-sight corrected/uncorrected etc), the participants were asked to choose between the two statements described in the previous section in order to assess their beliefs regarding the reliability or unreliability of photographs in recording the appearance of people and normal scenes.

When the preliminary data was recorded, a standardised script (appendix F) was used to ensure that the instructions given were the same for all participants. The participant's first task was to order the 12 photographs in their set (one of six balanced sets) from slimmest on the left to fattest on the right. They were encouraged to think of themselves as a "visual weighing scale" and to use the extremes of the range of descriptors as much as possible. It was also made clear that they must not be "kind" or give "the benefit of the doubt" for any of the models images and to be as dispassionate as possible in their size estimations. They were then shown the seven body-weight (below) descriptors and told their task was to apply one of the descriptors to each model.

Descriptor	Abbreviation Code
VERY OVERWEIGHT.	V.O.
OVERWEIGHT.	O.
SLIGHTLY OVERWEIGHT.	S.O.
CORRECT	C.
SLIGHTLY UNDERWEIGHT.	SU.
UNDERWEIGHT	U.
VERY UNDERWEIGHT	V.U.

Each photograph had a unique identifier code and after completing their first task, participants were asked to record each code on their response sheet from slimmest to fattest. The photographs were then randomised and re-presented. Participants were now told that their task was to look at each photograph individually and again give an estimated weight using the descriptors above and also to give an actual weight estimation for each model in plus, minus or zero kilograms. To help them with this difficult task, 12 individual BMI scales were prepared to accompany the photograph of each model. The seven weight-descriptor abbreviations were written on a caption card and the BMI weight range for the model being assessed was printed beneath each descriptor as a guide. This was to ensure that any under or overweight estimations would be within the range possible for each individual model and derived from their actual weight and height measurements. Each participant assessed the weight of each of the 12 images in their set. The experimenter wrote the individual weight estimation underneath the code previously written for each photograph and asked the participant to confirm that the figure was correct before moving on to the next image.

When this task was completed, the photographs were randomised again and the participants were now asked to put the photographs in order of apparent age: The youngest were to be place on the left and the oldest on the right. The order was recorded and the participants were then asked to give an actual age for each of the models based on their facial appearance alone. Each age was recorded under the code for each photograph.

The final task for each participant was to view each of the 22 monochrome portraits from the first pilot study (figure 6.1a&b). The photographs were randomised and the task was to place them in order of slimmest to fattest from left to right. Both orders were recorded and the participant was given short debriefing on the purpose of the experiment and asked not to repeat it to any possible participants.

6.2.3 Results

Figure 6.5 shows the raw weight estimations for all models in all conditions in relation to the average weight for all models. It shows that as camera to subject distance (and focal length) increases, a higher rating of weight in kilograms was given. In this respect the graph repeats the pattern shown in figure 6.3 of rising weight estimation using a Likert scale of bodyweight descriptors alone. Additionally, it shows that mean weight estimations across all models were under their actual weight with the wide-angle lenses used for photographs taken within one metre distance. Overweight estimations rise with distances beyond 1 metre, though the effect appears to level-off with all images made beyond 3 metres.

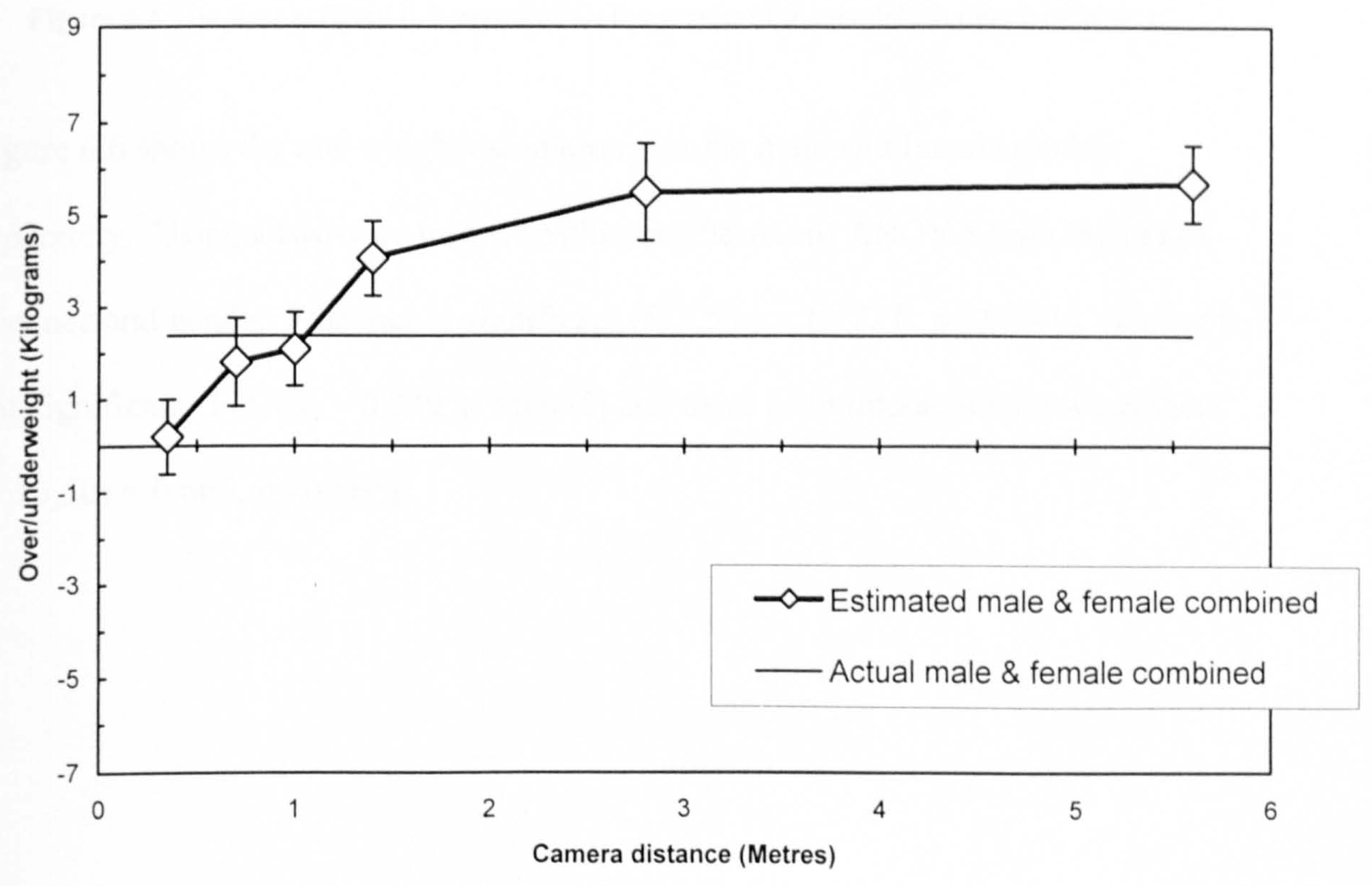


Figure 6.5 The combined weight estimates of both genders (in + or – kgs.) at all distances and the actual measured weights for both groups

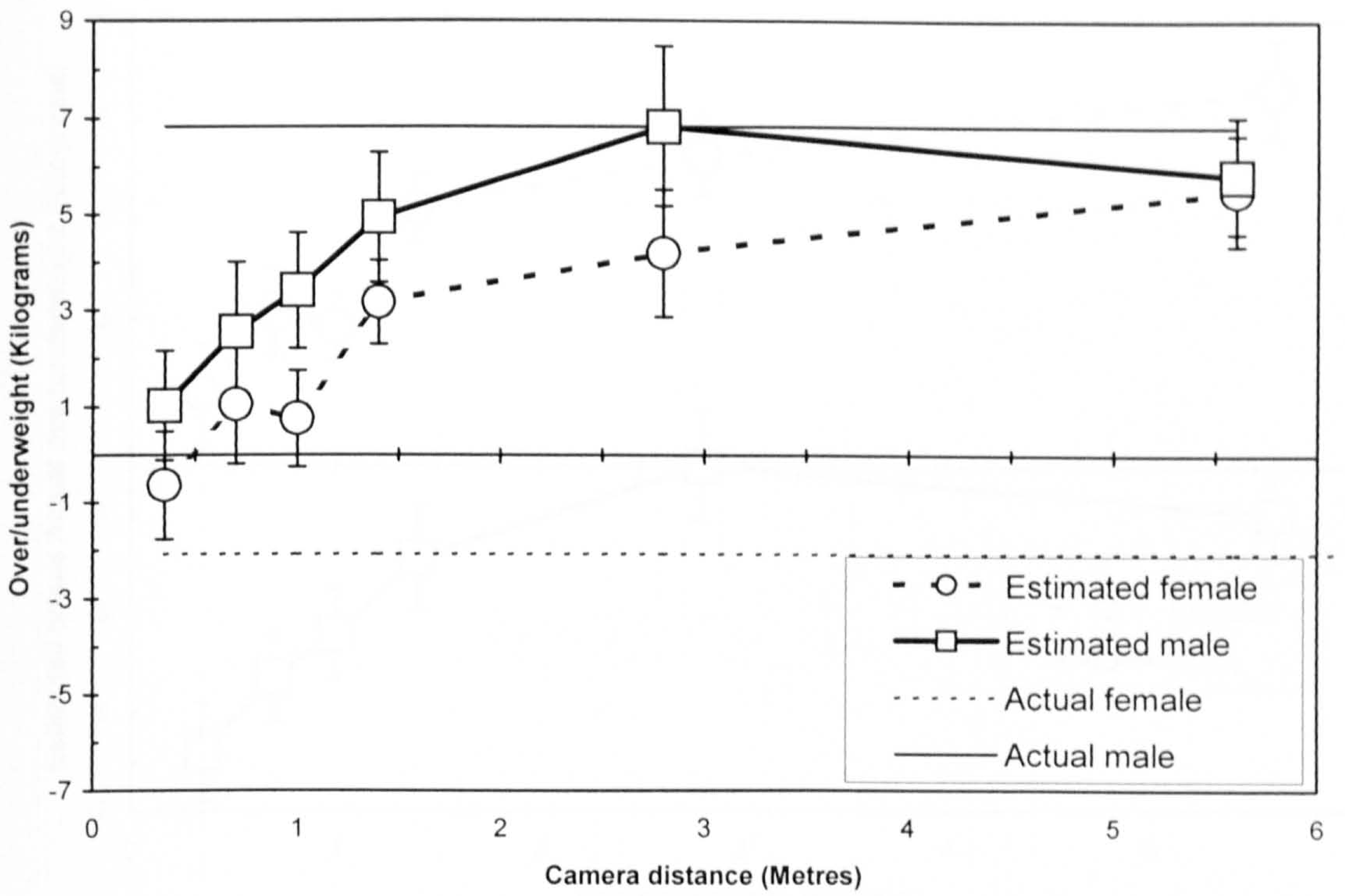


Figure 6.6 The data in figure 6.5 separated by the gender of the models in the photographs.

Figure 6.6 shows the raw weight estimations for the male and female models separately. Using a two-way mixed (within and between) ANOVA with factors of distance and gender: Distance is significant (F(5,50) = 16.727 , p <0.001). Gender is not significant (F(1,10) = 0.240, p =0.635) and there is no interaction between them (F 15,50) = 0.665, p <0.665).

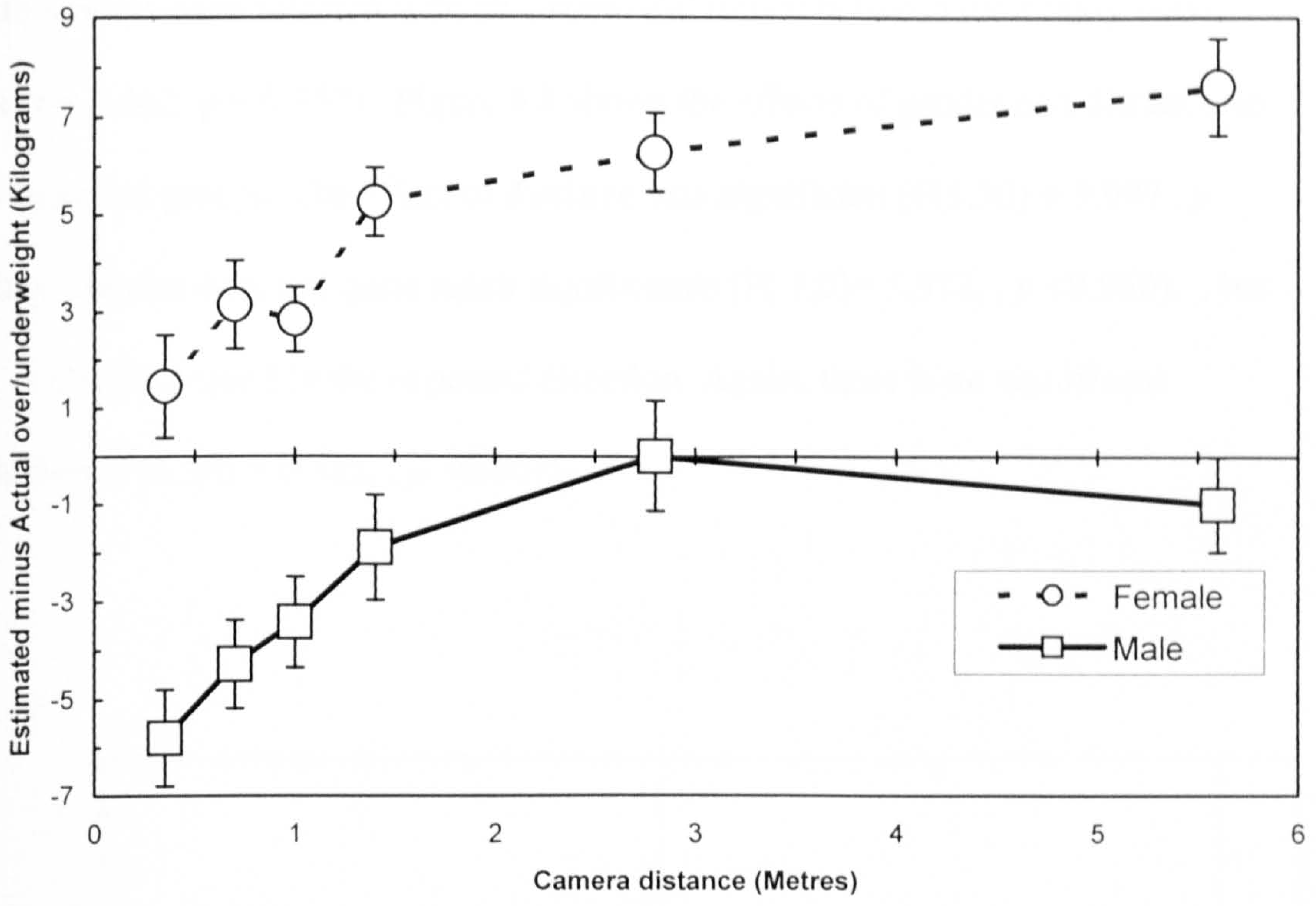


Figure 6.7 The estimated weight minus the actual weight for male and female models

As figure 6.6 shows, the actual weights of the men and women were very different and relative to these the estimated weights do show a gender difference, with the women being over-estimated and the men under-estimated. In order to allow for this, figure 6.7 shows the estimated weight minus the actual weight in kilograms. Using a two-way mixed (within and between) ANOVA with factors of distance and gender the effect of distance is again significant (F(5,50) = 16.727 , p <0.001). However, gender is also now significant: (F (1,10) = 13.775, p <0.004). Again, there is no interaction between them (F 15,50) = 0.647, p <0.665).

It is possible that this difference between the genders is due to the difference in their mean weights, with underweight models being over-estimated and overweight models being under-estimated irrespective of their gender. So a sub-set of four male and four female models were selected with no overall difference between their body mass index (t = 0.062, p = 0.952). Figure 6.8 shows the effects of gender and distance on these matched groups. The effect of distance was significant (F(5,30) = 9.997 , *p* <0.001). Gender does not quite reach significance (F(1,6)= 5.372, , *p* <0.060). , but there is clearly a trend in the expected direction. Again, there is no significant interaction (F(5,30) = 0.312, , *p* =0.902).

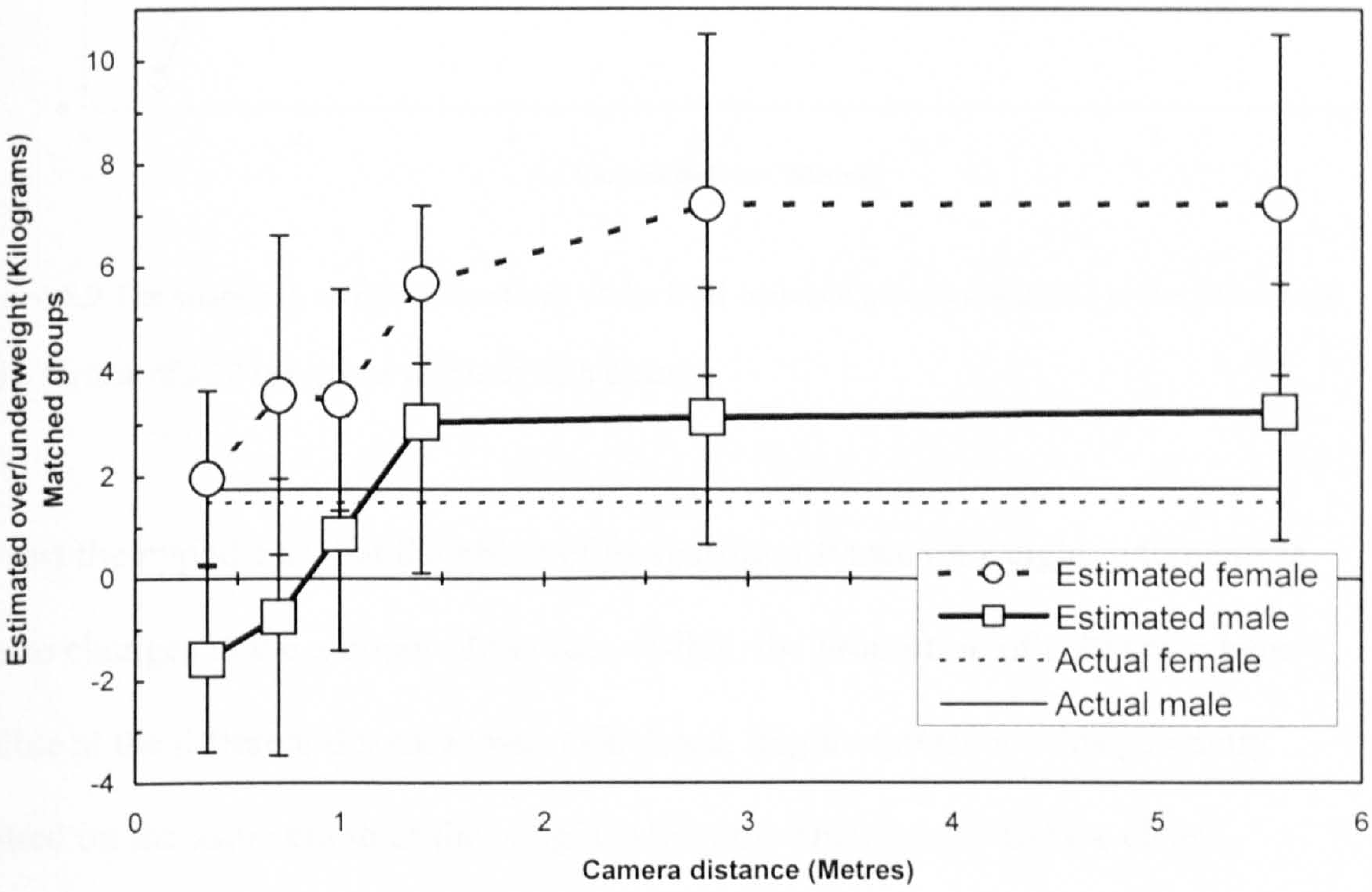


Figure 6.8 BMI matched gender groups.

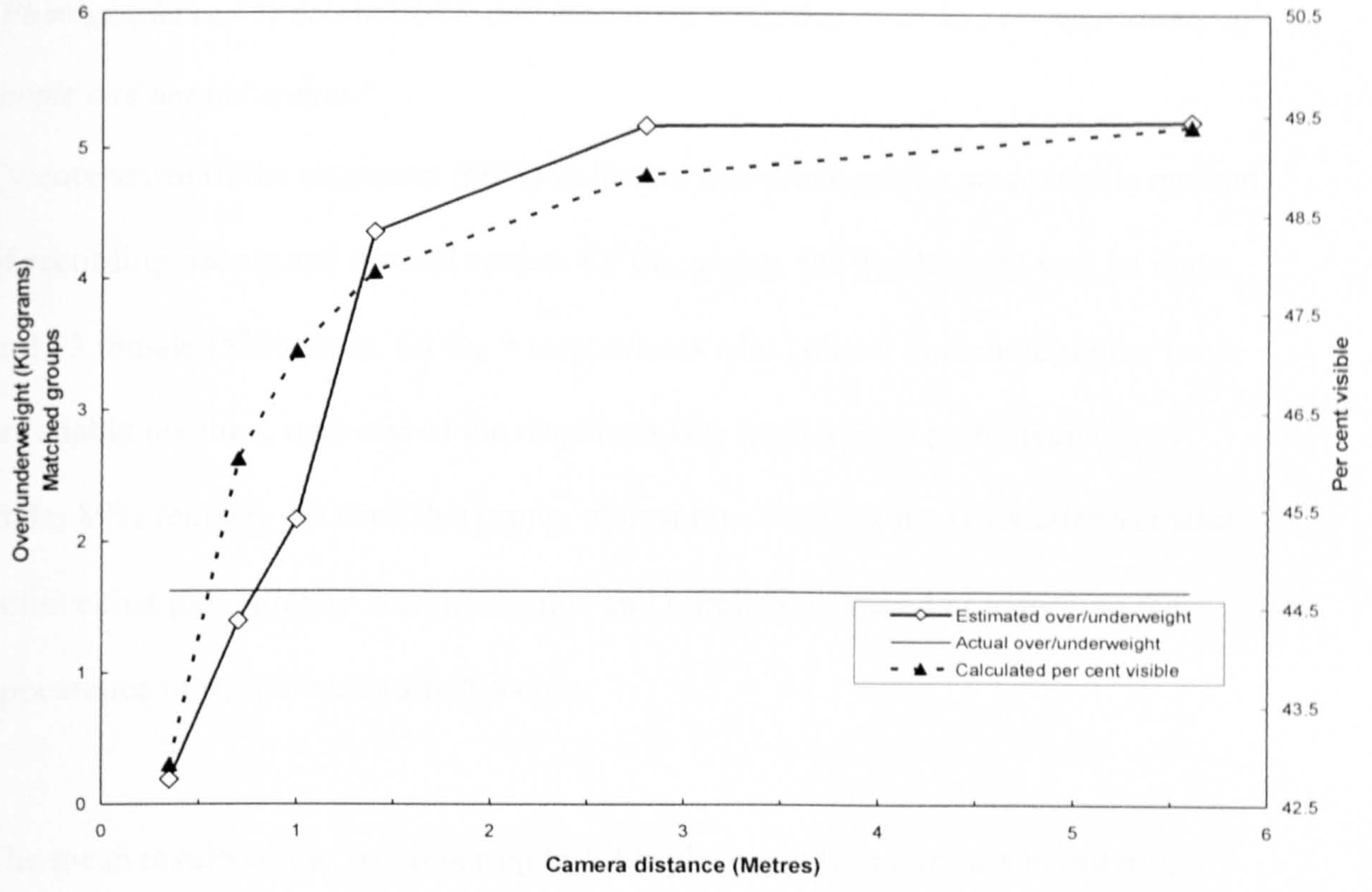


Figure 6.9 The change in weight estimations of the BMI matched groups compared to the percentage of the surface of a 20 cm sphere visible at each distance.

To test the hypothesis that the effect of increasing distance on weight judgments is due to changes in the amount of the face visible, the proportion of a 20 cm sphere visible at the different distances was calculated. Figure 6.9 shows this geometry plotted on the same graph as the weight estimates. There seems to be a clear relationship between estimations of bodyweight at each distance and calculations of the visible surface area of a sphere from the same distances.

All 36 participants (15 males and 21 females) were asked to to look at the following statements and underline which one agreed with own opinion:

"Photographs are generally a reliable and accurate method of recording the appearance of people and normal scenes."

"Photographs can be an unreliable and inaccurate method of recording the appearance of people and normal scenes."

Twenty seven of the responses (66%) indicated that photography was reliable method of recording people and general scenes. Of this group, the gender split was 14 male and 13 female (52%/48%). Of the 9 respondents who believe that photography is an unreliable medium, only one of the responses was from a male participant (11 % male, 89% female). So from this group, almost nine times as many females as males believe that photography is an inaccurate and unreliable method of recording the appearance of people and normal scenes.

The mean results of the size ordering task (of placing eleven portraits in order of thinnest to fattest) shown in table 6.1 had the expected reliability of responses. Two randomised sets (male and female) of the 11 portraits shown in figures 6.1 a & b were presented. Without any knowledge of the original order, four of the participants place one set of photographs in this order (widest angle to longest telephoto) using the criteria of slimmest to fattest images placed left to right. 15 participants ordered one set with one transposition of position. 17 participants made two or more transpositional errors. The most common transpositional error was to place photograph 7 in position 6 (and vice versa). This may have been caused by the similar perspectives of the 80mm and 100mm lenses used. Judgements may also have been affected by a slight size error in the printing that made image 6 slightly larger than image seven. The same effect is probably responsible for the size ordering error in the female portraits at positions 10 and 11.

Distance portraits taken from in cm. (wide-angle to telephoto)	32	45	55	67	85	118	154	205	298	418	538
Mean relative position for Male photos (slimmest left, fattest right)	1	2	3	4	5	7	6	8	9	10	11
Mean relative position for Female photos slimmest left, fattest right.	1	2	3	4	5	7	6	8	9	11	10

Table 6.1 The mean perceived size order of a male and female series of portraits (figure 6.1)

Table 6.2 shows how the most commonly used telephoto lenses used in portraiture, movies and television affect body weight estimations and gender. While only 12.5 % of weight estimations were overweight for the average BMI male models, 93% of all responses indicated that average BMI females were over their actual bodyweight in images made using these lenses. Table 6.3 shows the mean apparent age of each model at each camera to subject distance.

Percentage of total Weight Estimations above actual bodyweight for:	All Lenses & Distances	Minus Widest Angle Lens	From All Tele-photo lenses
All models	56%	57%	56%
All Males	31%	35%	24%
All Females	80%	78%	87.5%
Ideal BMI Males	23%	27%	12.5%
Ideal BMI Females	87%	84%	93%

Table 6.2 The total number of size estimations above the actual weight of models in relation to the telephoto lenses typically used in photography, film and television.

	N	Minimum	Maximum	Mean	Std. Deviation
Estimated Age of Male at 0.35M	42	19	38	25.64	4.977
Estimated Age of Male at 0.7M	42	18	36	24.24	4.515
Estimated Age of Male at 1.4M	42	17	38	25.88	5.209
Estimated Age of Male at 2.8M	42	18	42	24.90	5.595
Estimated Age of Male at 5.6M	42	18	45	25.62	5.939
Estimated Age of Male at 8.4M	42	18	35	24.26	4.494
Valid N (listwise)	42				

Table 6.3a Mean age estimations for male models at the six distances photographed

	N	Minimum	Maximum	Mean	Std. Deviation
Estimated Age of Female at 0.35M	42	18	36	24.02	4.901
Estimated Age of Female at 0.7M	42	18	39	23.62	4.813
Estimated Age of Male at 2.8M	42	18	42	24.90	5.595
Estmated Age of Female at 1.4M	42	17	35	23.64	4.908
Estimated Age of Female at 5.6M	42	17	34	23.17	3.602
Estimated Age of Female at 8.4M	42	16	38	24.02	4.871
Valid N (listwise)	42				

Table 6.3b Mean age estimations for female models at the six distances photographed

6.2.4 Conclusions.

The results of significantly increasing weight estimations with portraits made from increasing camera to subject distances in experiment 5 (figure 6.3) were confirmed in experiment 6 (figure 6.5). Figure 6.6 showed that the combined gender data in figure 6.5 concealed a gender effect of underweight estimations for all males in all camera conditions and overweight estimations for all females in all conditions. The use of Body Mass Index as a method of establishing whether each participant was over or under ideal weight showed that the male group mean was +7 kg overweight. This was mainly due to two males who between them were 31 kg. overweight and only one male who was slightly under weight. The female group had two slightly overweight and three underweight females, so the mean for this group was -2kg. So the result of overweight estimations for female models confirms the prediction made after examining the data from experiment 5. However, the results of underweight estimations for the male models was not predicted (figure 6.7).

Two matched BMI subgroups were derived from the BMI data (figure 6.8). There were six near ideal BMI models (+0.5 BMI), 3 of each gender to produce large matched groups so that ideal weight statistics could be derived. The matched groups were approximately 1.5 kg.s overweight, so it was difficult to produce statistics that were generalisable to normal bodyweight ranges. Nevertheless, the gender data was almost significant ($F(1,6) = 5.372$, , $p < 0.060$), and showed a trend in the expected direction.

Figure 6.9 shows the change in weight estimations for matched BMI gender groups compared to the percentage of visible surface of a sphere at each distance. The combined data illustrates the relationship between geometry and perception of curved objects such as a sphere and the human face. There seems to be a clear relationship between estimations of bodyweight at each distance and calculations of the visible surface area of a sphere from the same distances. Geometric effects like these are discussed in chapter eight and indicated that the human visual system can detect minor changes in apparent size and assign weight estimations based on these changes.

The result of the size ordering task was expected. It had been determined using informal trials with fellow photographers that the task was very difficult if one tried to order the images from widest angle to longest telephoto, but comparatively easy if one ordered them from fattest to thinnest. It is predicted that if there were consistent spacing between lens focal lengths, fewer conditions and better control of print magnification then most participants would find this size ordering task as very easy. However, attempting to do the same task by estimating focal length alone would be very difficult (even for photographers) and a correct order almost impossible for a non-photographer.

Table 6.2 shows that the telephoto lenses used in portraiture will produce overweight images of almost all females and will induce relatively higher weight estimations for the slimmest models. Males however will be portrayed with much more accuracy (or less unflatteringly) as only 12.5% of the average BMI male portraits induced overweight estimations. A possible cause for this previously undescribed gender effect is discussed in chapter nine. Tables 6.3a & 3b show that no differences in apparent age are revealed by the method used in this experiment.

The results from these experiments suggest that observers are able to judge bodyweight from images of the face and neck only. It seems that this ability is similar in nature to the phenomenon of *boundary extension* (discussed in section 3.4.2) and that the size of the unseen body can be directly affected by the camera-to-subject distance used to make the image.

CHAPTER SEVEN

EXPERIMENTS 7 & 8

The Synoptic Effect

7.1. Introduction.

In October 2000, the author presented a demonstration at the Presence 2000 conference at the Delft Technical University in Holland. The demonstration used stereoscopic projection to show how some virtual objects (later called "peanuts") appeared to become slimmer when one switched the display from conventional 2D to stereoscopic 3D projection. The term "flattening and fattening effect" was used to describe the reverse condition, when stereoscopic information is lost in a conventional 2D photograph. In the audience was Dr. Pieter Jan Stappers, who was head of the Design Conceptualization and Communication in the Department of Industrial Design. At the end of the presentation, he asked the author if he was aware of a device called a "Synopter." He explained that it was an old viewing device designed to add the appearance of depth to paintings and photographs. But it could also be used to remove stereoscopic depth from binocularly viewed real scenes. When used to view people, the device conveyed a remarkable "flattening and fattening" effect that later proved to have great relevance to these studies and the subsequent experiments.

Dr. Stappers' department had constructed a version of the device, derived from 1907 Carl Zeiss patent drawings for use in experiments by the psychologist Jan Koenderink. The synopter had recently been returned to him and a few minutes after the end of his stereoscopic demonstration, it was being demonstrated to the author. The Koenderink synopter differed the later Zeiss device in that it used prisms, rather than semi-silvered mirrors to split one light path into two parallel beams 65mm apart (see figure7.1). To fuse each "bi-ocular" channel, the HVS has to diverge the eyes to view the target as if it were at optical infinity. This process does not break the link between accommodation and vergence, so each tends to focus at or near infinty, even though the target may be well within accommodation distance. The synopter had a mechanism for widening the interocular separations to match the varying inter-pupilary distances found in small children or large adults. In principle however, it was almost identical to the mirrored beam-splitter shown in figure 7.2 overleaf. Dr Stappers aligned the device by asking people to look at him while holding the device as one would use a pair of binoculars. From his position, he could see one prism in the centre of the Synopter which appeared to show an image of two overlapping eyes. He gave instructions abut how to use the thumbwheel to adjust the I.P.D. until he could see both eyes perfectly aligned one behind the other. His viewpoint gave the impression he was being looked at by a person with a single eye centrally located above the nose. It is perhaps for this reason that the Synopter was given its alternate name, "The Cyclopter" See appendix page G for an appreciation of the cyclopter by the artist and perceptionist, Terry Pope.

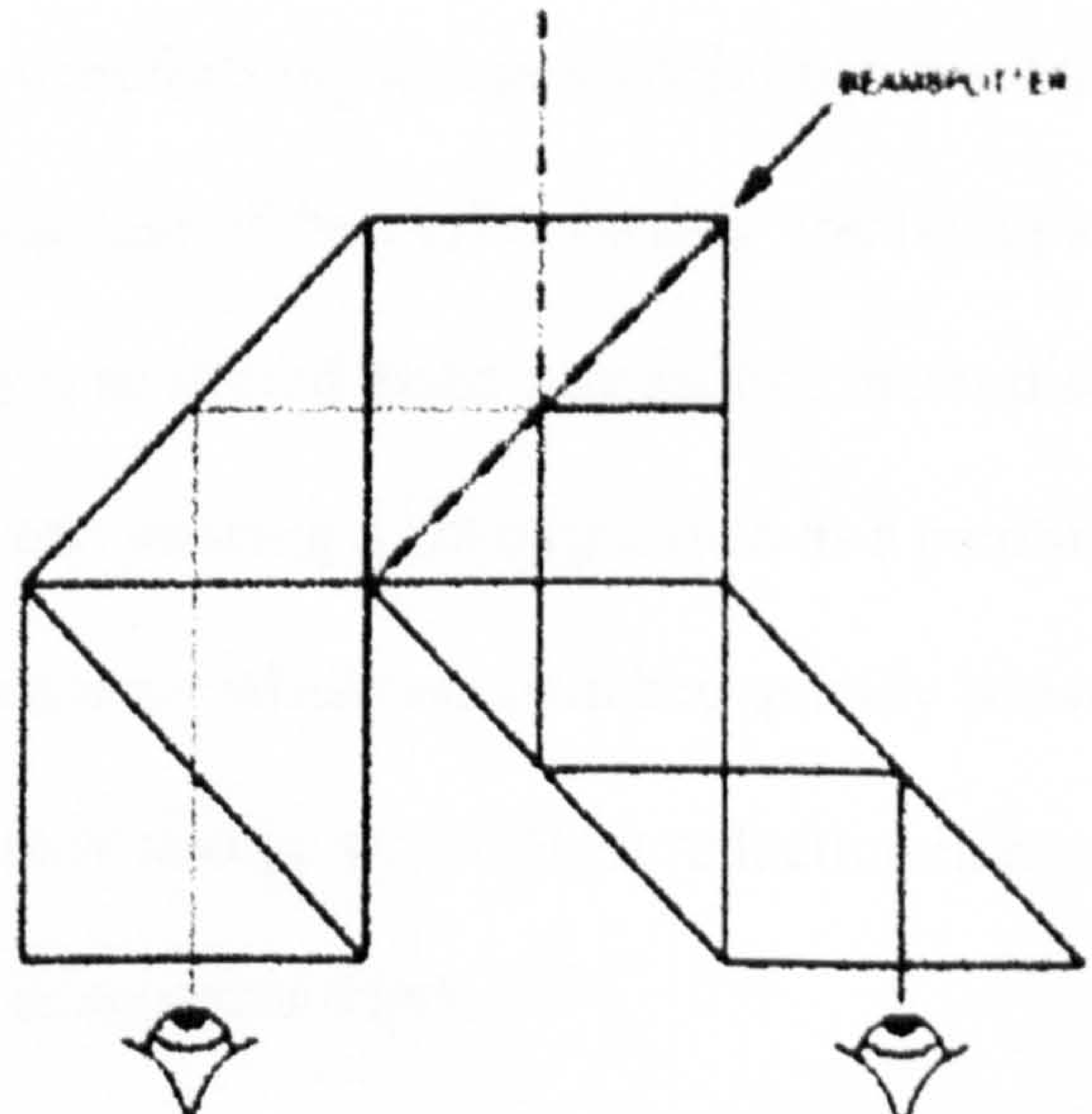
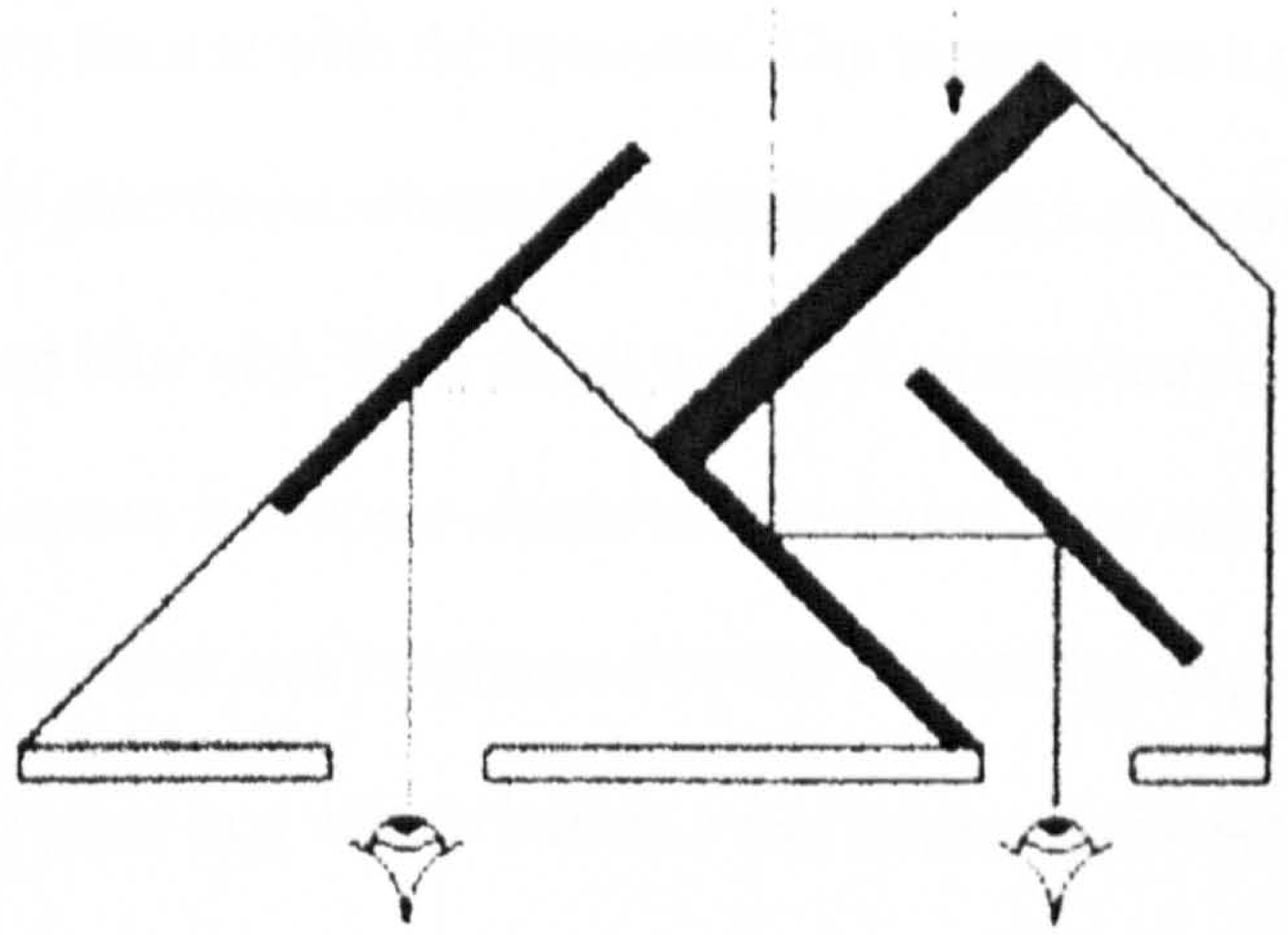


Figure 7.1: The optical system of the1907 Synopter, as used by Koenderink, et al. 1994



.**Figure 7. 2**. The principle of the optical system used in the1907 Zeiss mirror synopter

While Dr. Stappers was aligning the Synopter, he looked very different from how he appeared with direct vision. His face appeared to be artificially flattened and widened, while seeming to retain its vertical dimensions intact. It was almost as if his face had been squashed up against a pane of glass. The last time the author had seen this effect before was when he used autofocus binoculars to view people standing only a few metres away. When asked if there were any optical elements in the Synopter that could magnify the image, the author was informed that image is in fact smaller through the device than with direct vision. This is because the split light beams take a longer path to each retina, therefore any image seen through it subtends a slightly smaller angle of view. In the case of the Zeiss synopter, the light path is approximately 33mm longer, so should make a target one meter distant appear to be 3.3% smaller. However, when viewing a 2D target such as a painting, there is no apparent change in size constancy when one switched quickly between direct viewing and using the device. The only change was a slight reduction in apparent brightness, and a remarkable increase in apparent depth.

Dr. Stappers' lab had many posters on the walls and photographs deliberately chosen as targets for use with the synopter. The biggest was a poster of a holiday resort in a sunlit pine forest situated in a shallow valley an undulating forest floor and backed by a deep blue sky. With direct vision, it was an unremarkable scene. *But through the synopter, it became almost stereoscopic in its apparent depth and naturalness.* The effect was heightened by the advertising caption, which appeared to be printed on to glass in a widow before a real scene. There was what appeared to be stereoscopic separations between the all of the elements in the scene, starting with the text from the background, the nearest trees to the furthest and the furthest from the sky behind. The effect disappeared when one viewed the target directly and reappeared every time the device was used. Interestingly, the foreground details of the forest floor leading from the trees to beneath the camera were also enhanced. Viewed directly, the undergrowths was flattened and carpet like. But through the synopter, it appeared to have rolling undulations and relief that look like enhancements of the almost subliminal cues in the 2D image.

The synoptic effect did not work well with all of the targets, which were a mixture of photographs, paintings, graphic designs and text. The effect seemed to be strongest in photographs with strong or and natural-looking linear perspectives. In telephoto landscapes and aerial pictures, however, the effect seemed to be less obvious. The effect seemed to disappear completely when viewing text, graphics and impressionist-like landscapes. But paintings with detailed, natural looking perspectives and colours seemed to come to life with the synopter. Once one has seen this effect, it easy to understand why the earliest versions of the synopter were intended enhance the viewing experience of paintings in art galleries. When asked if he was aware of any experiments the synopters apparent ability to magnify 3D image, Dr. Stappers replied that there were none to his knowledge. This was despite the fact that it seemed to be a powerful and obvious effect to him and to everyone who saw it. He was asked if the author could borrow his synopter to run some size constancy experiments with naïve participants and return it to him as soon as possible. He agreed that it could be borrowed for a limited period, and the device was brought to the UK at the end of the conference.

The Delft synopter was a handmade device of robust construction. However, one became aware that it induced a sensation of eyestrain if one used it for longer than a few minutes at a time. Also when one asked others people to look through it, they reported a similar sensation. And any attempt to adjust the IPD setting by trying to merge the two virtual eyes into a single "cyclopean" eye was impossible because there appeared to be a horizontal misalignment of the optics. Tests using a laser shone directly into the synopters main prism confirmed that not only were the two output beams horizontally misaligned, they were also diverging by various amounts at various I.P.D. settings. The viewer was stripped down and inspected for wear. But it was found to be in good condition and the misalignments were caused by a design fault of having too much play in the adjustment mechanism. After consulting with Dr. Stappers, it was decided to realign the optics temporarily using metal shims and plastic wedges. The alignment of the output beams was fixed parallel at 65mm as this would only narrow the angle of view by a negligible amount for the few users whose I.P.D. was not near to the average measurement. The resulting alignment was reported as very comfortable by all subsequent users and it was decided that the Delft synopter was now ready to use in an experiment into size constancy between synoptic and direct viewing conditions.

7.2.1 Experiment 7184. Changes in apparent size of a Synoptically viewed target.

7.2.1.2 Introduction

When viewing a perspective photograph or painting through a synopter, many viewers report that the images have a 3D quality to them. (Koenderink et al 1994) have measured this depth perception and found it to be up to three times deeper than the measured depth of the original object. This experiment is designed to establish if a synoptic viewing device seems to present a magnified image to the visual system and to assess the stereo vision thresholds of each participant using two stereo acuity testing methods. A further test would be to ask the participants to view passers- by (using the synopter and then direct vision) and ask them to report if its image makes people appear to have a change in perceived bodyweight.

7.2.1.3 Participants

Twenty one undergraduate students from the Department of Psychology. N.B. One participants data has been dropped as she reported diplopia while using the synopter in all three conditions.

7.2.1.4 Procedure:

After recording the participant's identity, age and if they require corrected vision; the results of two types of stereo acuity tests was recorded (TNO and Frisby). The participant is then shown the synopter and the target. The target is an almost life-sized polystyrene head and was placed on a table if front of a coloured cloth background. They were asked to look carefully at the target from a fixed position. They are then informed that their task was report if the synopter is giving an enlarged image, reduced image, or no change in size compared to viewing the head directly. The task was repeated at a second distance and the two distances of 0.5M and 1.0M. were alternated for each participant to avoid order effects. In the de-briefing they were asked to view a randomly selected person from a nearby coffee bar (at a distance of approximately 1.5 metres) and to report if the synopter appears to make people appear to be fatter, slimmer or about the same size as they were with direct vision.

7.2.1.5 Results

Nineteen of the participants demonstrated good stereovision on the more modern of the stereo acuity tests (the Frisby test which features real depth targets). The TNO test uses anaglyph glasses to judge random dot stereograms and gives consistently lower scores. It indicated that six of the participants had poor stereovision but only one person had a zero score in both tests. There was no correlation between measured stereo acuity and performance in the experiment.

Of the sixty size judgements made, no respondent identified that the image of the target was in fact smaller when viewed through the synopter.

80% of the responses indicated that the object looked larger when viewed through the synopter.

Eleven of the participants indicated that the target looked larger in all three conditions.

Seventeen of the participants reported that people looked larger or fatter when viewed through synopter and the other three reported no change in size.

7.2.1.6 Discussion

The responses from this synopter experiment indicate that when it is used to view a 3D target head, the "flattened and fattened" subjective experience it gives is strongly supported by the data collected. It is reasonable to speculate that these 2D presentations of faces accurately simulate the flattening effect in photographic portraits of people using a standard (same size reproduction) lens camera. It also supports the data from Harper & Latto (Presence, June 2001 V.10 page 323) which showed that portraits taken with a standard lens gave overweight estimations when averaged across all subjects and indicated that in most conditions, the portraits gave the appearance of an increase in perceived bodyweight.

7.3.1 Experiment 8. Changes in apparent size of a Synoptically viewed target II

7.3.1.2 Introduction

The prismatic synopter (designed by Carl Zeiss to exploit the phenomena of "paradoxical monocular stereoscopy") used in experiment 7 described a highly significant fattening effect when used to view both real and dummy heads. Of the sixty size judgements made by twenty participants, no participant reported that the image of either target was in fact smaller when viewed through the synopter. 80% of the responses indicated that the object looked larger when viewed through the synopter. Eleven of the participants indicated that the target looked larger in all three conditions. Seventeen of the participants reported that people looked larger or fatter when viewed through a synoptic viewer.

A possible problem with these data is that some unknown magnification effect of prismatic beam splitting might be inducing increased sized estimations. As the device was returned to its owners (Delft University) in October 2001, it was decided to construct a replacement synopter using surface effect and semi silvered mirrors to further investigate these effects. The new device was constructed from engineering plastic and balsa wood and the mirrors were aligned using a laser. One semi-silvered mirror was mounted to allow 50% if the light to pass through to the first eye. The rest of the light is reflected to a surface silvered mirror that then re-directs the beam in a parallel direction to the second eye (see figure 7.3). Each mirror centre has a 65mm separation and was mounted at an angle of 45 degrees to the viewing axis . This synopter design is closer in principle to the folding devices issued to art gallery visitors at the turn of the century. It was considered to be inferior to the prismatic synopter as the surface effect mirror conveys an image that is approximately 10% smaller than the semi silvered mirror axis for an object viewed at 0.5 metres. This could produce diplopia in a close-up condition, as some people cannot fuse images with this amount of spatial separation or size disparity. In its original role, this was not considered to be a problem as most targets were further away than this. But in the first experiment, there was indeed a target placed at a distance of 0.5 metres. The design of this experiment was to reproduce the first experiment as closely as possible, but with the new device.

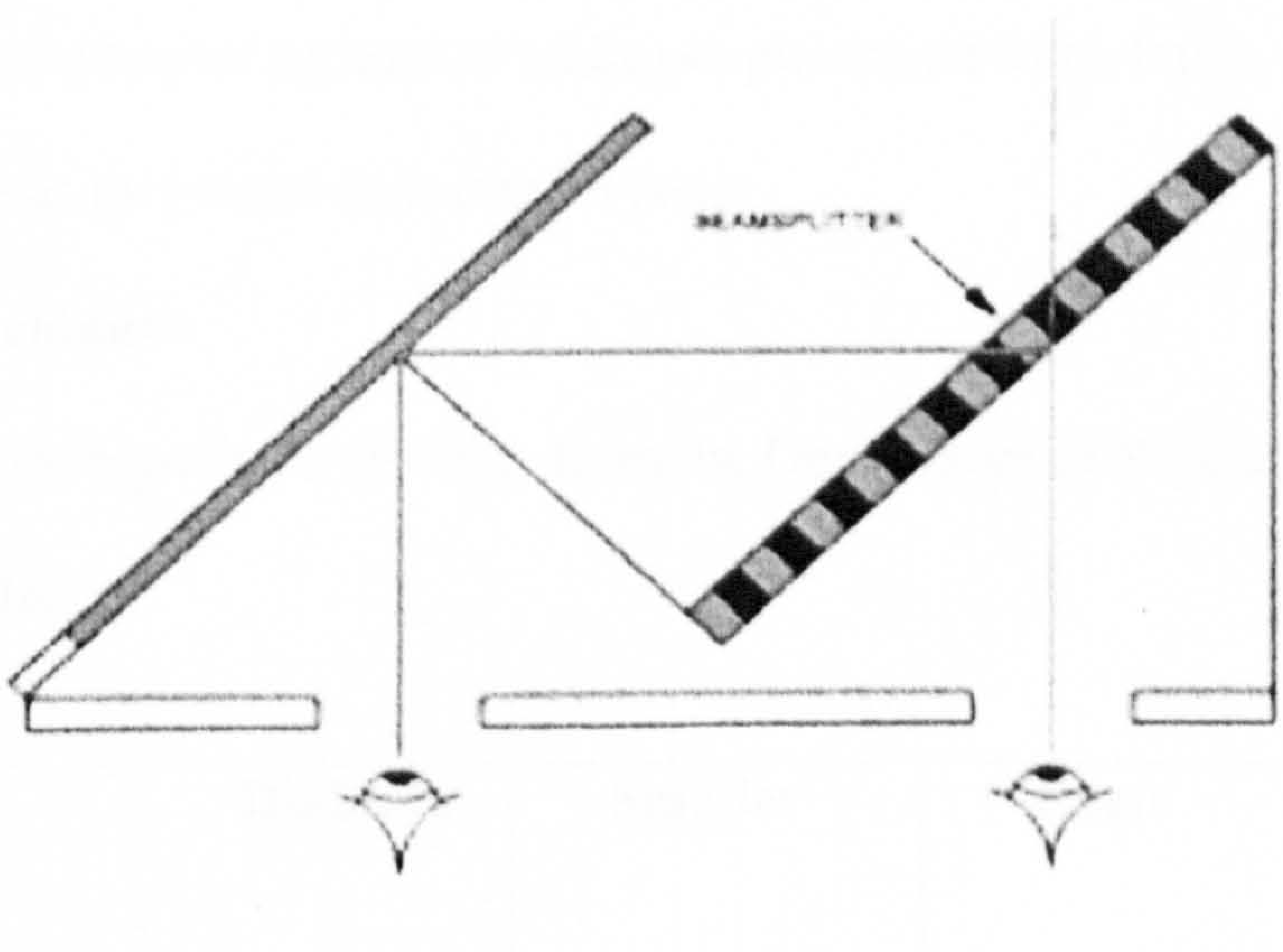


Figure.7.3 A Simple mirror and beam-splitter synopter with an IPD of 65mm. The 65mm illustration (by Terry Pope) shows the asymmetric beam path of this device and how the left hand reflected image is 65mm further away from the target than the image seen by the right eye in this orientation.

An obvious confound to this approach was that the mirror synopter has asymmetric optical paths. It was decided therefore to make judgments with the device in its normal, and in its inverted condition. This would double the number of responses recorded in each condition compared to the first synopter experiment.

7.3.1.3 Procedure.

After recording the participant's identity, age and if they require corrected vision; the participant is then shown the synopter and the target. The target is an almost life-sized polystyrene head and was placed on a table if front of a coloured cloth background. They were then informed that their task is to report if the synopter is giving an enlarged image, reduced image, or there is no change in size. The task was repeated at a second distance, with the two distances of 0.5m and 1.0m were alternated for each participant to avoid order effects. The synopter was also inverted in each viewing condition to balance the asymmetric optical paths of a mirrored device. In the de-briefing they were asked to view a passer by (at a distance of between 1.5-2.0m) and to report if the synopter appears to make people appear to be fatter, slimmer or about the same size as they were with direct vision.

7.3.1.4 Participants

Twenty two undergraduate students from the Department of Psychology.

7.3.1.5 Results

Object Condition	Diplopia	Smaller	Same	Larger
50 CM dist.	**11**	**1**	**4**	**28**
1 Metre dist.	**0**	**2**	**7**	**35**
Real person	**0**	**0**	**6**	**38**

Figure 7.4. Frequencies of responses under all conditions

7.3.1.5 Discussion

The responses from this second "mirror" synopter experiment indicate that when it is used to view a 3D target head, the flattened and fattened" subjective experience it gives is strongly supported by the data collected. Although the close-up condition was difficult to fuse on 11 occasions out of 44, the overall magnifying effect of the mirror synopter appears to have been reproduced with the more simple device at viewing distances of 1 metre and beyond. It also supports the data from Experiment 7 regarding the reported fattening effect on randomly selected people. Even though they were in close proximity to the subject of their judgement, most participants had no hesitation in saying that the synopter made the person appear to be fatter than when viewed directly. In the post experiment briefing, most participants expressed great surprise when told that image generated by a synopter gave a smaller image on the retina than with direct viewing. A possible explanation for this effect is discussed in chapter 9 in section 9.7.

PAGE NUMBERING AS ORIGINAL

CHAPTER EIGHT

The Effect of Geometric Changes, Subject Distance and Disparity on Perceived Body Size

8.1 Introduction

This thesis investigates Body Image Distortion in photographs of people. Yet it would be useful to demonstrate how any physical body can appear to change size or shape in different imaging conditions. For instance, photographers know that shadows rarely match the shape of the object that cast it (Figure 8.1) because the objects occlusion zone is illuminated by a close-up source. *Only objects viewed or illuminated from optical infinity can generate an occluded area that is the same size and proportions as the occluding object.* Light from the sun arrives from "infinity" in optical terms and is effectively parallel. Unless it is 2D, an occluders shadow (cast by the sun onto a parallel background) will have the same proportions as the illuminated object. In the following examples, occlusion zones are used to illustrate all of the perceived visual angles that can seen from the same position as a virtual point source of light.

Figure 8.1 Photographers and lighting designers use low lighting set close to the subject to create shadow shapes that are enlarged and distorted. (Photo source www.the-asc.org.uk/ About/abo_relevance.htm)

8.2 Proximity, Disparity and Occlusion.

A light source close to a 3D occluder will cast a shadow that will differ in proportion from the illuminated object when measured from the background. The 0.7 WHR peanut on the right (below) has cast a shadow with a WHR of 0.68 from a light 0.5 metres distant. The central light (1 metre distant) casts a shadow that is 0.69 WHR. The light at 5 metres will cast a shadow that has a WHR of 0.695. The shadow area

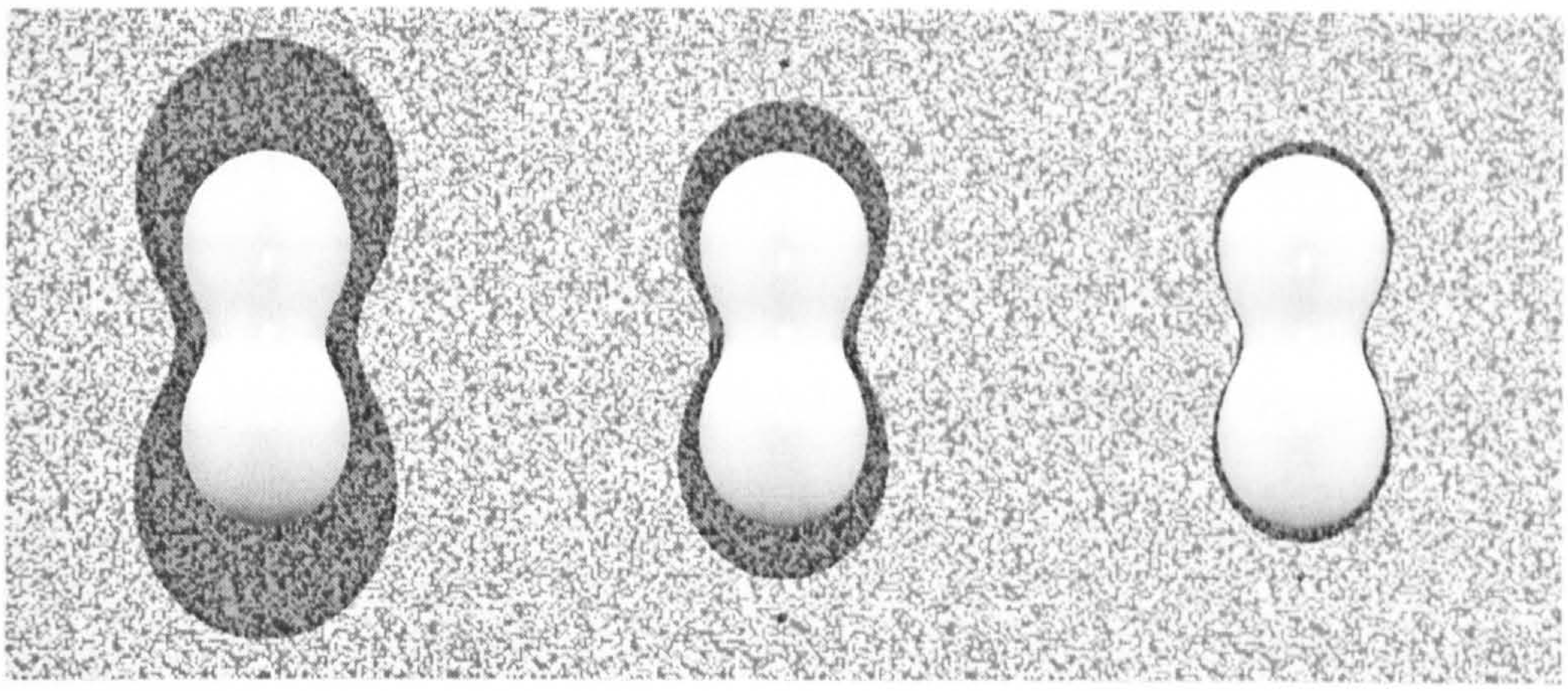


Figure 8.2 A light is moved away from the peanut in three stages and the shadow recorded at three distances: 0.5. 1.0 and 5 metres

changes its relative proportions because the waist is smaller than the widest part. The change is caused by the light ray being tangentially occluded at different parts of the circumference of the waist and hip areas. So the ratio between the measured background distance A-B (figure 8.3) and distance C-D is unique to this condition and will only match the occluders proportions when illuminated from a source at infinity.

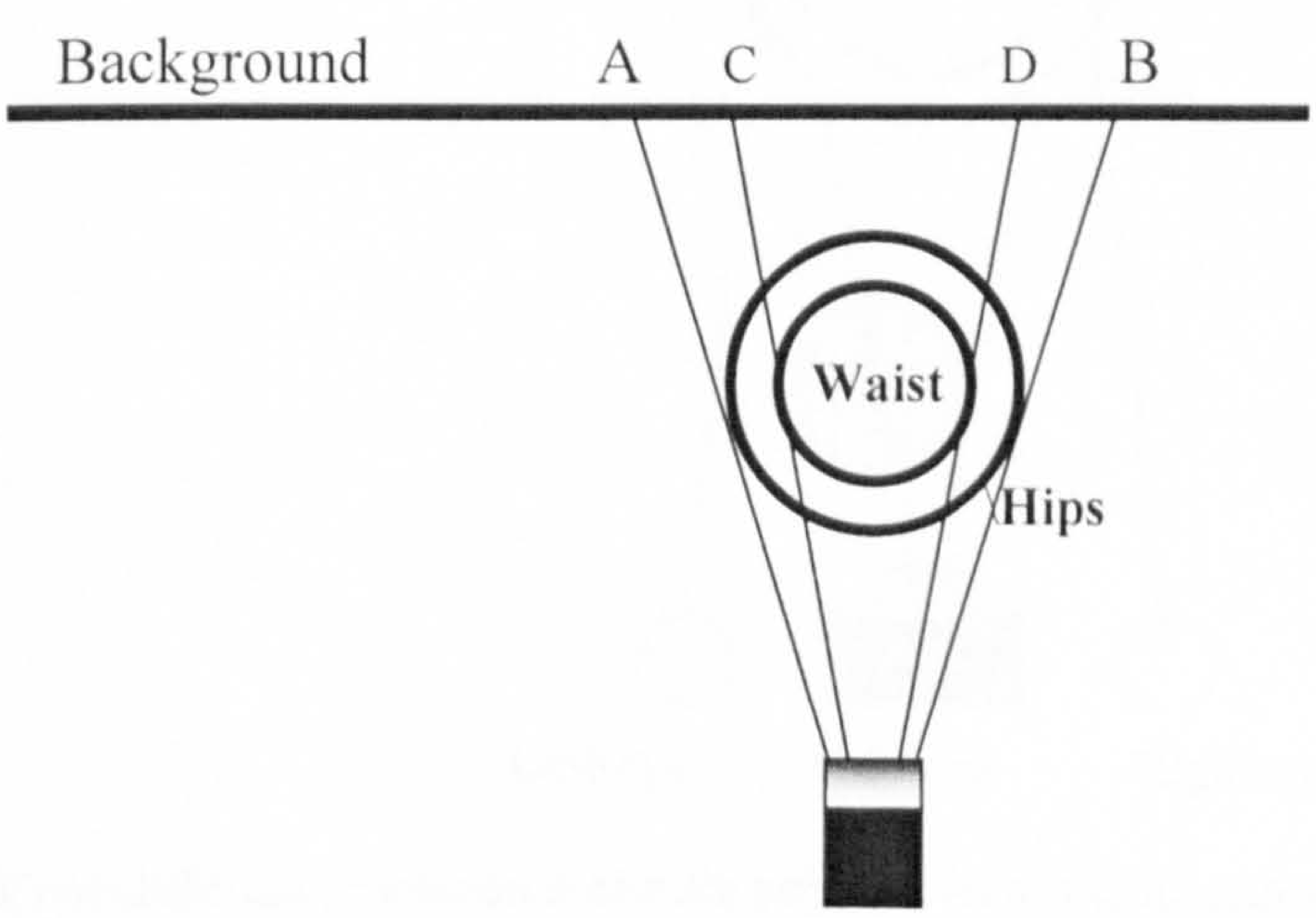


Figure 8.3 Plan of a lamp illuminating a background occluded by a multiple curvature peanut shaped object. The tangential point on the hips is nearer to the lamp than the same point on the waist, so the ratio of occlusions and visual angles changes with distance.

8.3 Disparity and Occlusion.

A possible explanation for the slimming effect of binocular disparity could be found by examining figure 8.4. An unscaled plan view is shown of an object imaged by a camera and stereoscopically from the same position. A larger area of the background plane is occluded from the camera position (A-B) than is seen stereoscopically (C-D). If one were to hold up a cup at arms length, a similar effect can be seen if one views it with either one or both eyes. Like the object in the diagram, the cup will occlude much more of the background when viewed monocularly compared to the binocular condition. *Perhaps the human visual system has a simple metric that governs core elements of size perception?* One possibility is that the HVS will assign a size percept to an object based on its perceived (monocular or binocular) visual angles coupled to the amount is appears to occlude the background. The principles of ecological validity would prevent us from seeing the cup change size as one viewed it monocularly and

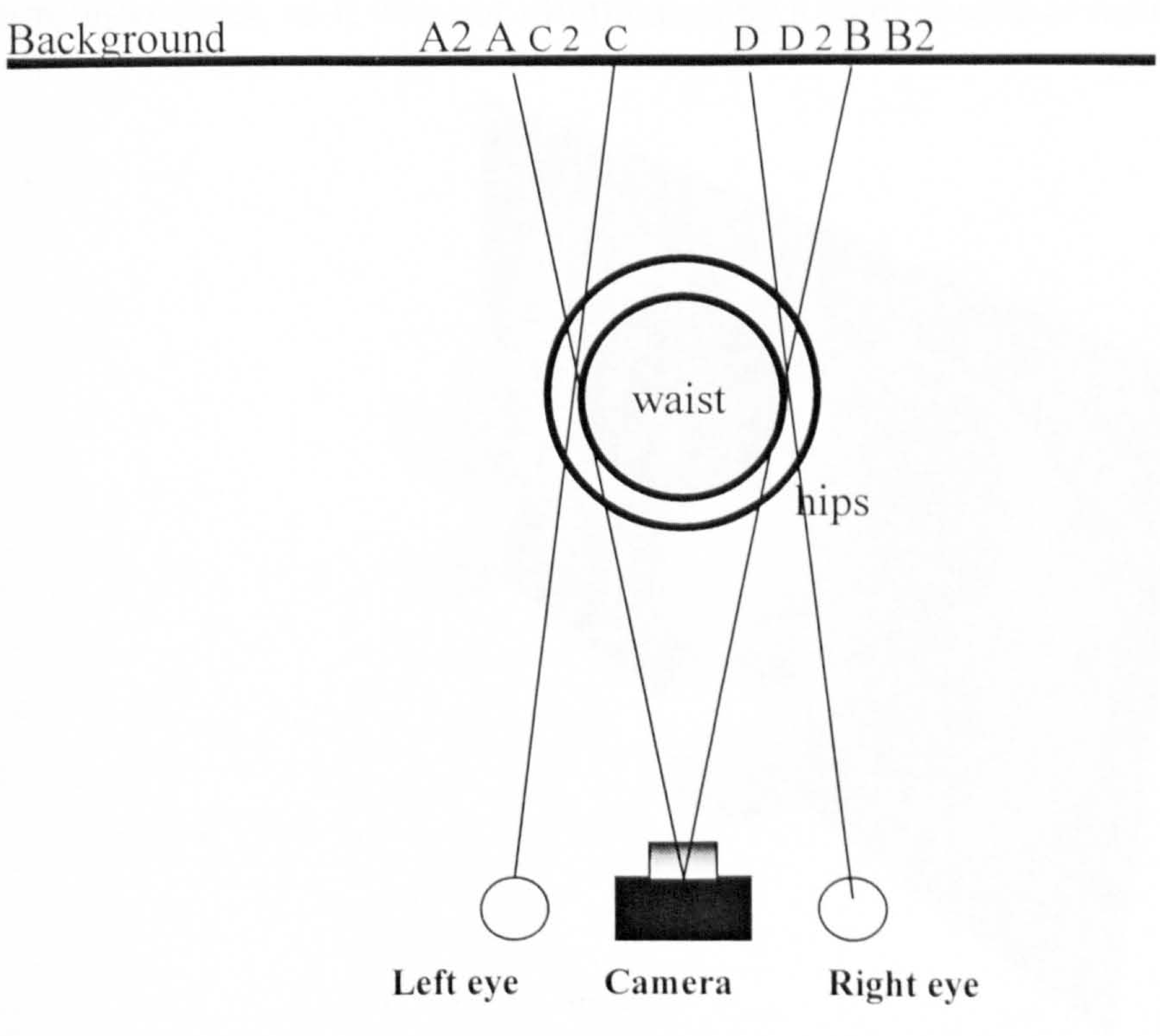


Figure 8. 4 The difference between a camera point of view and human stereo vision from the same position. A body occludes more of the background in a 2-D photographic perspective (A-B) than with stereo vision (c-d). The new tangential point on the hips is closer to the camera and eyes, so the ratio of occlusions (and visual angles) A2-B2 and c2 to d2 will differ even though the viewing position has not changed.

then binocularly. But if the targets were not directly comparable, then it may that a target viewed in 2D conditions may appear to be further away (due to its distance being beyond close-up accommodation feedback) and yet subtending a large visual angle on the retina. If an object appears to be far away and yet subtends such a visual angle, the visual system may assign a size percept that is larger because it seems to be far away (or at optical infinity) and must therefore be larger. However, when one views the same target binocularly it will have much more proximity information and reveal more of the area behind it. So a smaller size percept may occur due to the HVS perceiving its true distance and occlusion characteristics.

Figure 8.5 shows the peanut shape illuminated by single and paired light sources. It demonstrates the way in which the occluded area becomes smaller with disparity and its waist-hip ratio lowers with increasing disparity. It should be noted that the occluded area from the single light position is not equal to the 0.7 waist-hip ratio of the foreground object, as it was not illuminated by a light source at infinity.

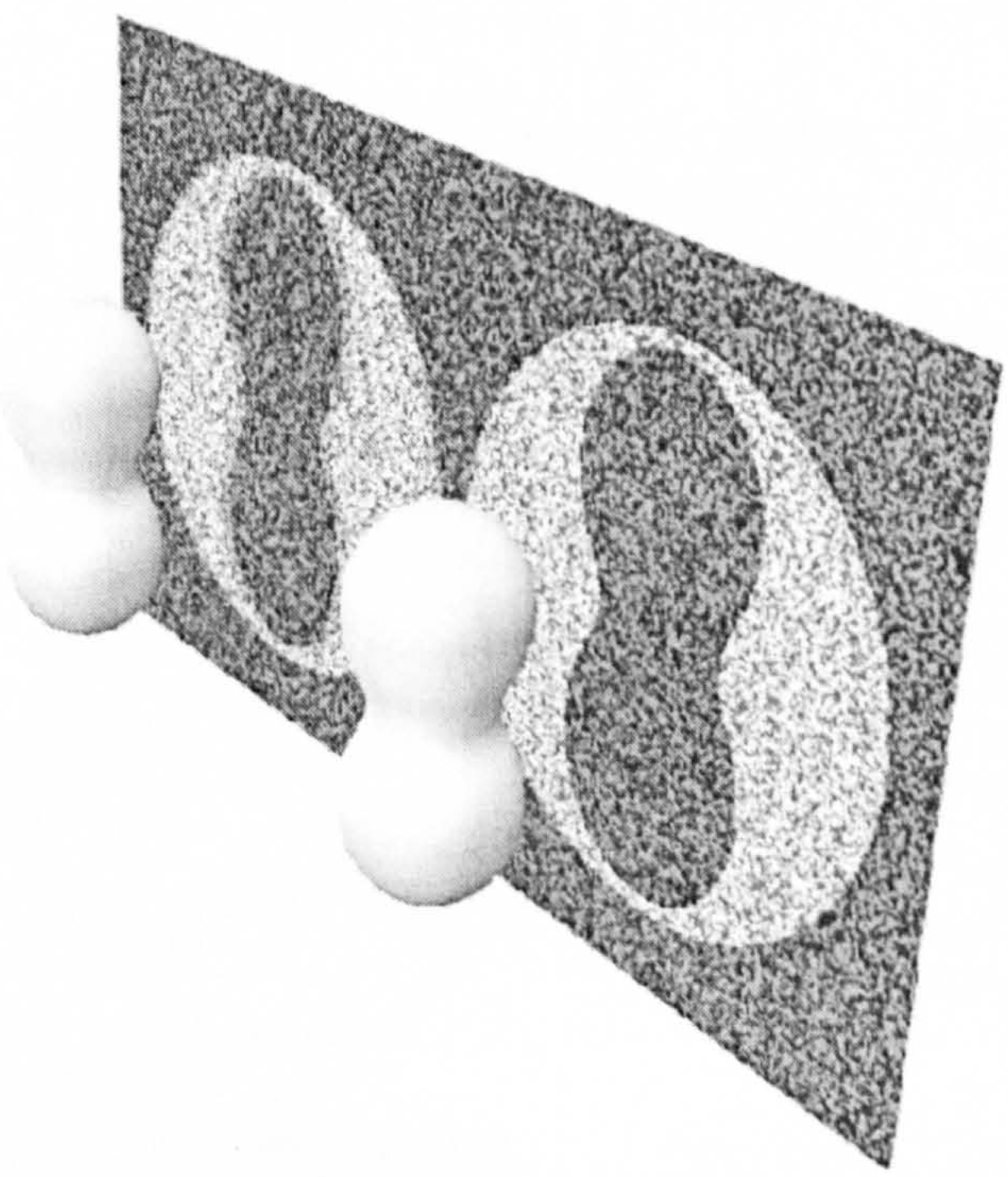


Figure 8.5a A representation of the size and shape of the occluded area behind the object shown in Figures 8.2 and 8.3 when illuminated by two lights on the left and one on the right 1.68 metres away.

Figure 8.5b shows the measurements of the occluded areas in 8.5a. They were derived from the original image files using tools in the StrataVision software. Measuring the waist-hip ratio of the occluded areas cast by a single light at 1.68 metres distance (the original camera to subject distance from the studies reported in appendix H and chapters 4-5), found that the occlusion zone waist-hip ratio was 0.692. When illuminated from the positions occupied by the orthostereoscopic cameras, the WHR reduced to 0.685. An additional condition not used in any experiments was measured. The light sources were placed 150mm apart and the occlusion zone WHR was measured at 0.668. This indicated that the changes in WHR with widening separations was consistent, but not linear. As these occlusion zones match the visual angles perceivable from the position of each light source, the relationship between disparity and occlusion was further investigated using a real target in section 8.3.

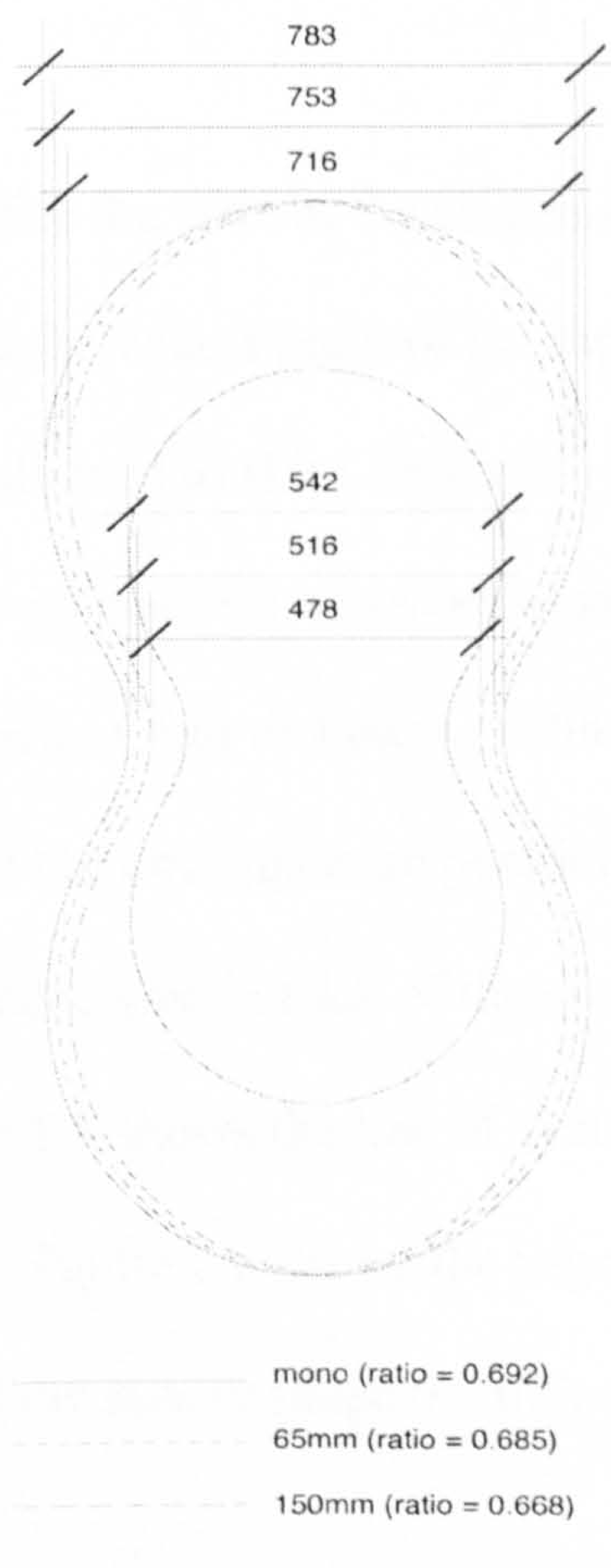


Figure 8.5b The size and shape of the occluded area behind the object.

8.3 Measurement and perceived visual angles of a full size waisted occluder

In order to support the contention that the virtual peanut stimuli changes its perceived proportions depending on distance and disparity, a simplified cylindrical occluder was made that matched the 0.7 waist-hip ratio and overall dimensions of the experimental stimuli. The object was then centralised onto a vertical pole that was placed 1 metre in front of a vertical white background to match the plan layout of the objects in the peanut experiment. The pole was placed above a central measurement line that was set at 90^{o} to the background plane. A small rig was made to support a laser pointer so that it could be accurately directed through each of three 1mm. wide apertures. The outer apertures were set at 65 mm apart on each side of the central aperture. The rig was attached to a heavy-duty tripod that was placed directly above the central measurement line at the first of the four measurement distances.

The technique used was to fire the laser at the tangential point on the side of the cylinders that approximated the widest point of the hips and the narrowest point of the waist. The laser point was allowed to skim across the cylinder edge and to fall onto the background from where its position was marked and measured from the centre line. This was done from each of four distances (800mm, 1380mm, 3360mm and 6270mm.) and from each of the three aperture positions. From these data points, the waist-hip ratios of the occluded area in each of the conditions were calculated and are shown in Figure 8.6. Figure 8.7 shows the line of best fit for the central and binocular (equivalent) measurements. Figure 8.8 shows the trigonometric calculations for the ratio of the visual angles of the peanut shape 0.7 object from zero and 65mm of disparity.

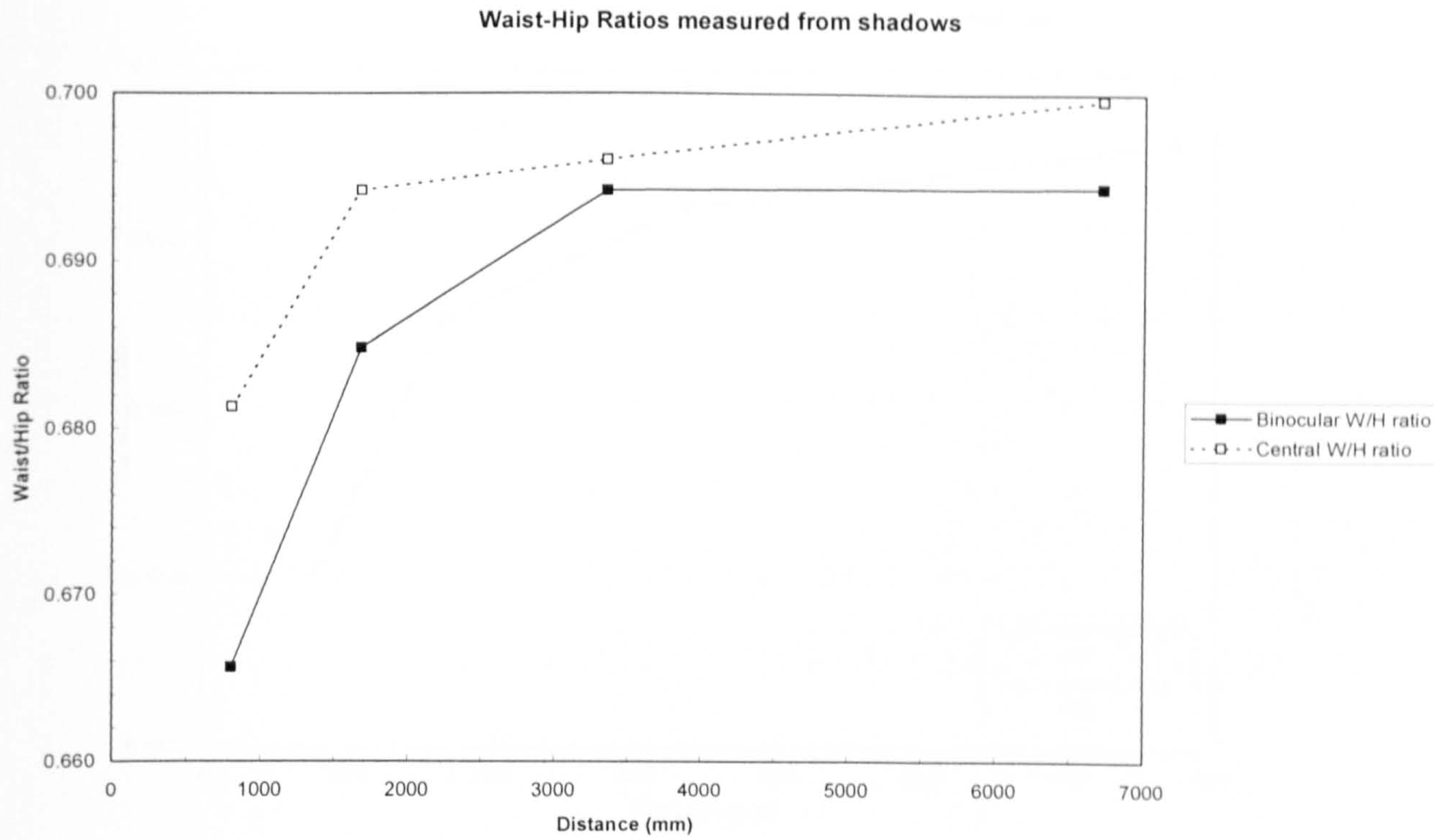


Figure 8.6. Shows how the 0.7 waist-hip ratio model occludes an actual background at four distances (0.8-6.27M) and with zero and 65mm disparity using a laser to trace its outline

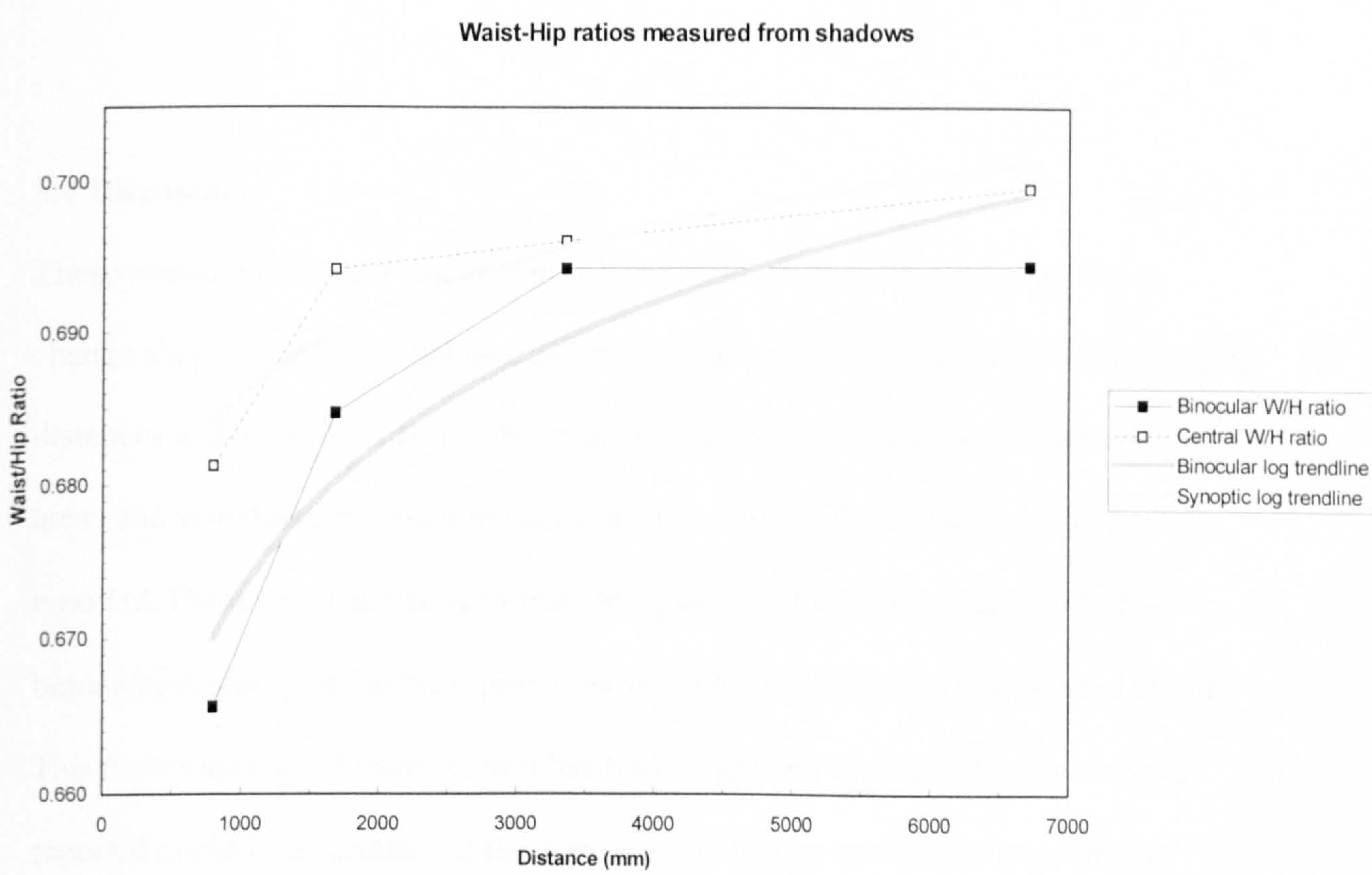


Figure 8.7. The line of best fit for the data in figure 8.6. Lines appear to merge at approximately 10 metres

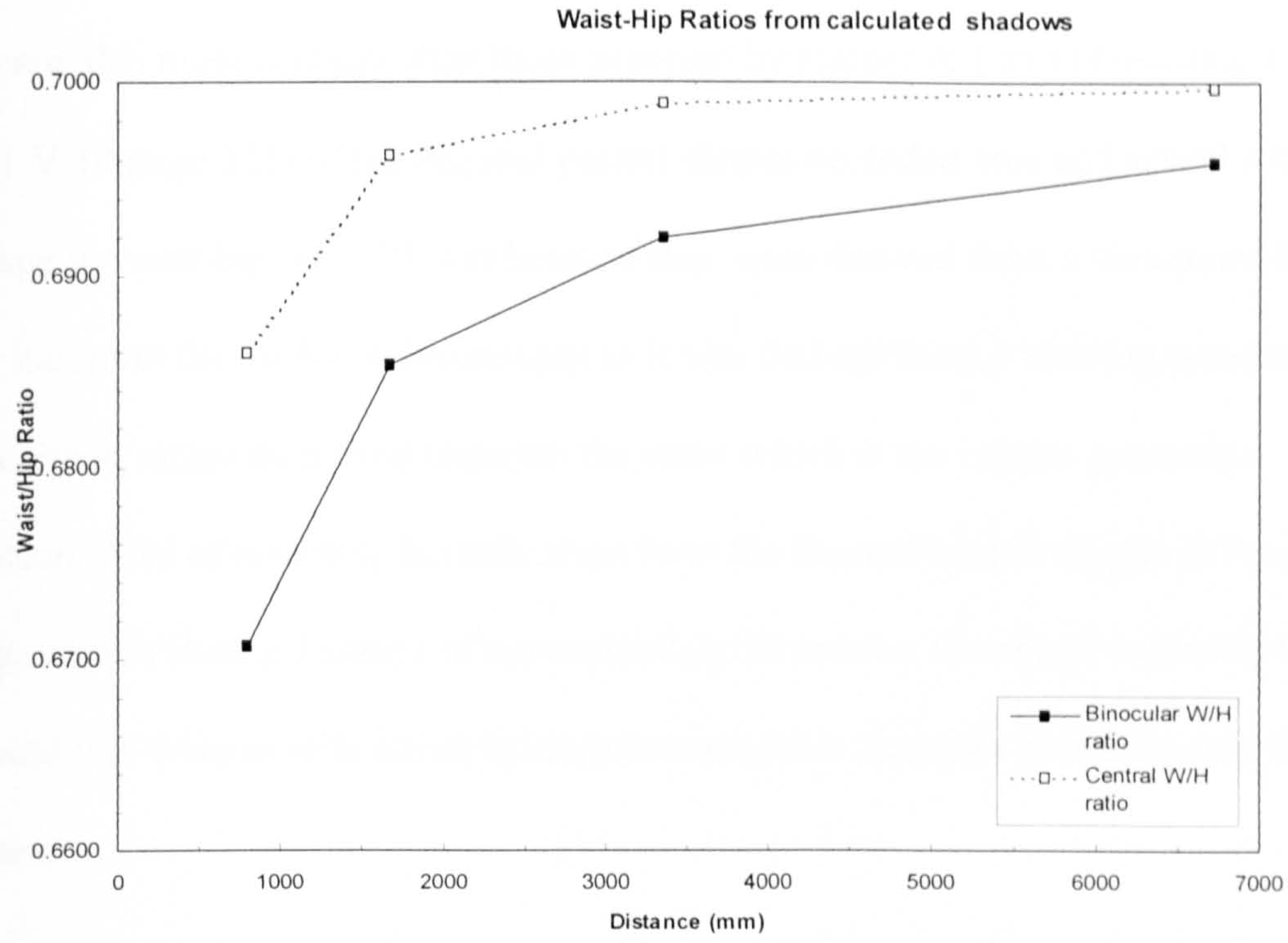


Figure 8.8. Trigonometric calculations of visual angles of the peanut shape 0.7 object from zero and 65mm of disparity.

8.4 Discussion

These measurements and calculations clearly indicate that objects can appear to change shape (in that they subtend different visual angles on the retina) with changing distances and disparities. If the object shape was constant, then the measured occluded areas and visual angles would remain constant at any of the distances of disparities reported. However, it can be seen from the graphs that the slimming effect of binocular disparity and subject proximity is a robust and easily demonstrated effect. This is an important discovery, as it has been suggested that the effect as we have reported could be an artefact of the stereoscopic display used in the experiments reported in chapters 4 and 5. Also of importance were the trigonometric calculations as they confirm that the measurement technique used was valid. These calculations also are also more accurate than those reported by Harper & Latto (Presence, June 2001 V.10 page 321) of the original peanut shapes occluded area and reveal a larger change in waist-hip ratio. This is because they were derived from a viewpoint that was incorrect for the hip measurement as it was derived from a viewing position opposite it, rather than from opposite the waist which is the correct geometric position. Also of note was the indication from the lines of best fit (figure 8.7) which suggests that from a distance of approximately 10 metres, the object is perceived as an 0.7 and that there is little stereo information available from the object beyond that distance.

CHAPTER NINE

Conclusions and Discussion

9.1 Introduction

This thesis has investigated the theory that photography is intrinsically distorting in terms of body image perception. It has produced strong evidence that the belief that the camera "adds 10lbs" to the perceived weight of people can be easily reproduced in experimental conditions that emulate normal photographic practices. The primary questions addressed related to the extent of the fattening effect, to what perceptual mechanisms it can be ascribed and if it can be counteracted in common practice. It has also investigated whether the level of trust people have in its accuracy is appropriate, specifically by examining the reliability of photography in relation to reproducing the face and form of human subjects. A summary of the main findings are:

- **Photography is an innately fattening medium at all camera-to-subject distances greater than 70cm.**
- **Photographs can only provide slimming effects using ultra-wide angle lenses at camera-to-subject distances of less than 70cm.**
- **Orthostereoscopic photography and life-sized orthostereoscopic reproduction are the only techniques that can produce images of people with undistorted size and shape reproduction.**
- **2D photography is a sexually dimorphic medium and produces fatter images of women compared to men photographed in identical conditions.**

9.2 Review of the main findings in Chapter 5:

9.2.1 Experiment 1: Swimsuit Experiment

A one factor (viewing condition) ANOVA showed that the models were rated as significantly slimmer when viewed stereoscopically (F=15.072, df=1, p=0.001). This result confirmed that the apparent change in size seen in side-by-side comparisons between monocular and binocular images of people could be reproduced, even when individual models were seen on only one occasion and participants were unaware that the viewing conditions were randomly changing between 2D and 3D projection.

9.2.2 Experiment 2: Slimming Effect with a non-human object

The results confirmed the prediction that the waisted object was viewed as slimmer or smaller in the stereo presentation ($\chi^2(2, N = 20) = 13.3, p < 0.001$). The stimuli were carefully controlled to ensure that the monocular and binocular stimuli were identically sized and that the only difference between them was in their stereoscopic disparities. Yet three times as many viewers saw the object as slimmer or smaller when viewed binocularly compared to the synoptic image. 95% of all responses gave different size estimates of targets that subtended the same visual angle and only one participant in the study saw the stimuli as an identical size both conditions.

9.2.3 Experiment 3: The First Peanut Experiment

This experiment revealed a previously unseen effect. The 0 Disparity 2D stimuli (card 9) was correctly matched to its projected equivalent (0.7 waist-hip ratio) by over half of the participants. The mean perceived waist-hip ratio of the groups responses was 0.694, which is almost identical to the occluded area as shown in Figure 8.5b of 0.692. However, when the viewers were shown the same shape orthostereoscopically with 65mm of convergence disparity (corresponding to the normal geometry of human stereo vision), a match with a significantly slimmer waist-hip ratio was selected. Conventional parallel-lens axis stereography was simulated with image

(65P) and the mean perceived waist-hip ratio did not differ significantly from the synoptic condition. However, there was much more variation in responses, suggesting that the lack of convergence and vertical disparities in this condition are less natural and more difficult to fuse. There was no correlation between the participants stereo acuity, measured with the TNO test. Neither was there an interaction between the effect of disparity on size judgements and the performance in the stereo acuity test, indicating that the TNO stereoacuity test is not a reliable predictor of performance in orthostereoscopic experiments.

9.2.4 Experiment 4: The Second "Textured" Peanut Experiment

The revised textured images (with accurate vertical disparities) and methodology combined with the original white images confirmed the experiment 3 findings of a slimming effect of binocular disparity. Increasing levels of stereoscopic disparity affected size judgement. With increasing disparity (from 0 to 120mm of interaxial separation), smaller matches were chosen by the participants. Overall, surface texture has no effect on size judgement. However, at 120mm of disparity, smaller matches were chosen in the textured surface condition than in the white surface condition. The textured peanut experiment showed for the first time that projected synoptic stimuli could be seen as larger than life-size. The stereoscopic stimuli showed not only a relative slimming effect of binocular disparity compared to the synoptic target, but also to the 0.7 WHR comparison target. These slimming effects appeared to follow a dose-response curve where the larger the disparity projected, the smaller the size estimations were induced. The participants were also asked to judge how far from the screen the background in the polarized images appeared to be, at each level of disparity. The mean estimated distances at 0, 30mm, 60mm, and 120mm of disparity were 0.24m, 0.77m, 1.44m, and 2.40m, respectively. A correlation between the perceived distance of background and disparity proved significant. As disparity increased, so did the perceived distance of the background in the stereoscopic image. Yet the stereoscopic separations for the background could only be fused by divergence, suggesting that the HVS is surprisingly flexible in it ability to fuse crossed disparities in people with normal stereoscopic vision.

Like the 'TNO test for stereoscopic vision,' the Frisby Stereotest' is designed to measure stereoscopic acuity, on which the size-matching task is based. To find out whether there was any relationship between the scores gained on the separate tests and performance in the size-matching task, a correlation was determined between both the TNO and Frisby tests, and the performance score. The obtained correlation between the TNO test and the performance score was not significant. However, the correlation gained between the Frisby test and the performance score showed that there was a statistically significant relationship between the two. The TNO test is therefore not a good predictor of performance in the size-matching task, wheras the Frisby test can predict performance. No significant interaction was found between disparity and gender, There would seem to be no difference between the sexes in their perception of 3-D stereoscopic information.

9.2.5 Discussion of experiments 1-4

The experiment 1 data supported the subjective impression that people looked slimmer in stereoscopic images. The data from experiment 2 confirmed that the slimming effect seen in experiment 1 was not caused simply by participants acting more favourably to a more realistic human presence and was robust enough to be seen with a non-human target. It also strongly suggested that the computer generated peanut shape used in the first CGI pilot study was seen as slimmer because it was a stereoscopic rendering, and not because the background design drove participants to make slimmer size estimations with stereoscopic projection. Experiment 2 also showed that stereoscopic images could induce smaller size estimations to otherwise identical monoscopic stimuli in 14 out of 20 participants. Yet it is possible that in a more carefully controlled version of this experiment, an even higher number of participants would report that the object was smaller in the stereoscopic condition. The hand-held stereo viewers lenses were designed for resolution and not magnification, so the images produced appeared to be slightly smaller than life size. Also, the stereoscopic photography was of an object that was 1.68 metres away. If it had been smaller and closer, there would have been more stereoscopic information and the image and possibly a larger change between the monocular and binocular conditions.

Experiment 3 used the first of a number of versions of the 07 WHR peanut shape and demonstrated a relative slimming effect of binocular disparity. It also demonstrated that increasing disparities gave smaller size percepts. It did not however show synoptic images as larger than life size. Experiment 4 investigated size perception of the peanut as a function of disparity (interaxial separation) and surface texture in 3-D images projected with corrected vertical disparities. The presence of texture and veridical vertical disparities, (and the resulting greater number of stereo cues present on the object) aided stereo fusion and allowed participants to choose larger than life-size matches with the synoptic stimuli and smaller size matches with widening disparities. An important observation regarding all of the CGI peanut experiments was that the area of zero disparity on all of the peanut shapes was also the area with the smallest geometric differences between the monocular (synoptic) and stereoscopic presentations. The largest areas of differences and widest disparities presented were found on the shoulder, hips and background plane in all of the binocular projections when viewed without polarizing glasses, Yet when alternating between viewing conditions, the waist area was seen to become smaller while the rest of the image seemed to retain its overall shape and size. This suggests that the HVS may operate on a powerful principle of ecological validity, whereby changes in an image that are well understood (such as a waists being made smaller by muscle action) are what we actually perceive, even though the actual changes in the image are outside of the area of attention.

In summary, the results confirm that increasing disparity results in smaller size-matches. Although texture appears to have no overall effect on size perception, at 120mm of disparity there were significant differences between the two conditions, with smaller matches for the textured stimulus. The reasons for this are unclear but may be related to object surface stereo cues aiding fusion of the two image channels. The study found that with increasing disparity the perceived distance of the background retreated further behind the screen, but that estimates of background distance at zero disparity did not correlate with the performance score for the size-matching task. No differences were detected between the error scores gained from the card-ordering task, nor was a correlation established between the combined mean error score and the performance measure.

Stereoacuity measured by the 'TNO test for stereoscopic vision' did not correlate with performance in the size-matching task. However, performance in the size-matching task was found to correlate with stereoacuity measured by 'The Frisby stereotest'. No effects of gender were revealed. The study therefore supports the use of orthostereoscopic techniques for stereo image capture and projection, using interaxial separations close to the inter-ocular distance of the human eyes (60-65mm). Results also point to the use of 'The Frisby stereotest' instead of the 'TNO test for stereoscopic vision' for predicting stereo performance, probably because it uses the perception of actual depth targets rather than the random dot stereograms of the TNO test. This confirms Heron et al (1985, cited in Rogers and Howard, 2002 p.147) who found that the Frisby test had the best test/retest reliability of the available clinical tests.

9.3 Experiment 5: Portraits made at varying camera to subject distances I

9.3.1 Results.

The experiment showed that as camera to subject distance (and focal length) increases, a higher score was given on the Likert scale estimating bodyweight. A one-factor (camera to subject distance) ANOVA found an overall effect of distance on weight judgements ($F(4,76) = 8.858, p < 0.001$). Photographs taken at 0.45 metres distance showed no slimming or fattening effect. However, using two-tailed t-tests, it was shown that the wide-angle, close proximity images (0.32 m) showed underweight estimations ($t(19) = 4.073, p = 0.001$) compared to the 0.45 metre condition. The standard lens image (0.71 m) showed a slight but not significant overweight estimation ($t(19) = 1.097, p = 0.287$). The telephoto distance images (1.32 m and 2.7 m) showed overweight estimations ($t(19) = .2.101$ & $5.101, p = 0.049$ & < 0.001).

9.3.2 Discussion.

Because of the limitations of the photographic location and lenses available, it was not possible to test if extending the range of focal lengths would show a continuing positive relationship between focal length and perceived bodyweight. It was also not possible to derive statistically useful gender based date from the five models (two male, three female). These effects were examined in experiment six, as was the anecdotally reported aging effect of photography.

9.4 Experiment 6: Portraits made at varying camera to subject distances II

9.4.1 Results

This experiment showed that as camera to subject distance (and focal length) increases, a higher rating of weight in kilograms was given. In this respect it repeats the pattern of rising weight estimation using a Likert scale of bodyweight descriptors alone shown in experiment 5. Additionally, it shows that mean weight estimations across all models were under their actual weight with the wide-angle lenses used for photographs taken within one metre distance. Overweight estimations rise with distances beyond 1 metre, though the effect appears to level-off with all images made beyond 3 metres.

Using a two-way mixed (within and between) ANOVA with factors of distance and gender, changes in camera to subject distance were significant ($F(5,50) = 16.727$, $p < 0.001$). Gender was not significant ($F(1,10) = 0.240$, $p = 0.635$) and there was no interaction between them (F 15,50) = 0.665, $p < 0.665$). Using estimated weight minus the actual weight in kilograms, with a two-way mixed (within and between) ANOVA with factors of distance and gender, the effect of distance was again significant ($F(5,50) = 16.727$, $p < 0.001$). However, gender was also significant: ($F(1,10)$ =

13.775, $p < 0.004$). Again, there was no interaction between them (F 15,50) = 0.647, $p < 0.665$). This data supports the observation that there are gender differences in the way that bodyweight is estimated and that women's weight is highly likely to be overestimated in normal photographic conditions. All 36 participants (15 males and 21 females) were asked to to look at the following statements and underline which one agreed with own opinion:

"Photographs are generally a reliable and accurate method of recording the appearance of people and normal scenes."

"Photographs can be an unreliable and inaccurate method of recording the appearance of people and normal scenes."

Twenty seven of the responses (66%) indicated that photography was reliable method of recording people and general scenes. Of this group, the gender split was 14 male and 13 female (52%/48%). Of the 9 respondents who believe that photography is an unreliable medium, only one of the responses was from a male participant (11 % male, 89% female). So from this group, almost nine times as many females as males believe that photography is an inaccurate and unreliable method of recording the appearance of people and normal scenes.

The mean results of the size ordering task (of placing eleven portraits in order of thinnest to fattest) shown in figure 6.10 had the expected reliability of responses. Two randomised sets (male and female) of the 11 portraits shown in figures 6.1 a & b were presented. Without any knowledge of the original order, four of the participants placed one set of photographs in this order (widest angle to longest telephoto) using the criteria of slimmest to fattest images placed left to right. 15 participants ordered one set with one transposition of position. 17 participants made two or more transponsitional errors. The most common transpositional error was to place photograph 7 in position 6 (and vice versa). This may have been caused by the similar perspectives of the 80mm and 100mm lenses used. Judgements may also have been affected by a slight size error in the printing that made image 6 slightly larger than image seven. The same effect was probably responsible for the size ordering error in the female portraits at positions 10 and 11.

When examining the effects of the most commonly used telephoto lenses used in portraiture, movies and television, body weight estimations rose with distance and gender. However, while only 12.5 % of weight estimations were overweight for the average BMI male models, 93% of all responses indicated that average BMI females were over their actual bodyweight in images made using these lenses. However, when apparent age was reported, no effect of changes in camera to subject distance could be found in the responses recorded.

9.4.2 Discussion.

The results of significantly increasing weight estimations with portraits made from increasing camera to subject distances in experiment 5 were confirmed in experiment 6. It was also shown that the combined gender data in figure 6.5 concealed a gender effect of underweight estimations for all males in all camera conditions and overweight estimations for all females in all conditions. The use of Body Mass Index as a method of establishing whether each participant was over or under ideal weight showed that the male group mean was +7 kg overweight. This was mainly due to two males who between them were 31 kg. overweight and only one male who was slightly under weight. The female group had two slightly overweight and three underweight females, so the mean for this group was -2kg. So the result of overweight estimations for female models confirms the prediction made after examining the data from experiment 5. However, the results of underweight estimations for the male models was not predicted. Matched BMI subgroups were derived from the BMI data were approximately 1.5 kgs. overweight and contained only four models. Nevertheless, the gender data was almost significant ($F(1,6)= 5.372$, , $p <0.060$), and showed a trend in the expected direction.

The result of the size ordering task was expected. It had been determined using informal trials with fellow photographers that the task was very difficult if one tried to order the images from widest angle to longest telephoto, but comparatively easy if one ordered them from fattest to thinnest. It is predicted that if there were consistent spacing between lens focal lengths, fewer conditions and better control of print magnification then most participants would find this size ordering task as very easy. However, attempting to do the same task by estimating focal length alone would be very difficult (even for photographers) and a correct order almost impossible for a non-photographer.

9.5 Experiment 7. Changes in apparent size of a Synoptically viewed target.

9.5.1 Results

Nineteen of the participants with good stereovision made sixty size judgements of a model head using direct vision and via a prismatic synoptic viewer. No respondent identified that the 2D image of the target was in fact smaller when viewed through the synopter. 80% of the responses indicated that the object looked larger when viewed through the synopter. Eleven of the participants indicated that the target looked larger in all three conditions. Seventeen of the participants reported that people looked larger or fatter when viewed through synopter and the other three reported no change in size.

9.5.2 Discussion

The responses from this synopter experiment indicate that when it is used to view a 3D target head, the "flattened and fattened" subjective experience it gives is strongly supported by the data collected. It is reasonable to speculate that these 2D presentations of faces accurately simulate the flattening effect in photographic portraits of people using a standard (same size reproduction) lens camera. It also supports the data from Harper & Latto (2001, p. 323) which showed that 2D portraits taken with a standard lens gave overweight estimations when averaged across all subjects and indicated that in most conditions, the portraits gave the appearance of an increase in perceived bodyweight.

9.6 Experiment 8. Changes in apparent size of a Synoptically viewed target II

9.6.1 Results

A mirror synoptic viewer was constructed from engineering plastic and balsa wood and the mirrors were aligned using a laser. The responses from this second "mi2rror" synopter experiment indicate that when it is used to view a 3D target head, the flattened and fattened" subjective experience it gives is strongly supported by the data collected. Although the close-up condition was difficult to fuse on 11 occasions out of 44, the overall magnifying effect of the mirror synopter appears to have been reproduced with the more simple device at viewing distances of 1 metre and beyond. It also supports the data from Experiment 7 regarding the reported fattening effect on randomly selected people. Even though they were in close proximity to the subject of their judgement, most participants had no hesitation in saying that the synopter made the person appear to be fatter than when viewed directly. In the post experiment briefing, most participants expressed great surprise when told that image generated by a synopter gave a smaller image on the retina than with direct viewing.

9.7 Discussion of the synoptic effect.

A possible explanation for the fattening or enlarging effect of synoptically viewed images may be because the device forces the observers eyes into a parallel axis orientation before it can fuse the two images presented by the bi-ocular viewfinder. Normally the HVS uses parallel optical viewing to fuse on objects at the horizon, or targets well beyond the normal range of accommodation and stereoscopic fusion. So the brain may have size constancy mechanisms linked to parallel viewing that make objects subtending a large visual angle on each retina be perceived as large if the eyes do not need to converge to fuse the two images. Perhaps it is simply the case that targets seen through a synopter are "encoded for infinity." If this encoding is combined a view of a near target that subtends a large visual angle on the retina, it may be that the visual system defaults to a larger size percept. It may also be the case that the flat 2D reproduction of a face or scene in a conventional photograph is also "encoded for infinity" because it has none of the disparities that we would expect to see in a close-up normal scene. Instead, the image has the horizontal and vertical disparities of a target viewed at infinity and so again may force the visual system into seeing the target as further away, and therefore larger.

9.8 Jaw-Neck Ratio

The experiments reported here suggest that photographs made without orthostereoscopic information will lead to predictably distorted perception. However, it is also argued that the conveyed neck size in photographs in relation to jaw width

Figure 9.1. These images were scaled so that the inter-ocular distance is the same on each face. At increasing camera-to-subject distances, the relative thickness of the neck in relation to the jaw width also increases.

seems to be an important factor in body-weight perception. Differences in apparent neck width can lead to sexually dimorphic perception with larger estimations of bodyweight made for female faces than male faces under the same photographic conditions. The reason for this seems to be that wide necks can have a sexually dimorphic character in humans. In males, a widened neck could be caused by an excess of muscle mass produced by exercise and the action of testosterone. However, it seems females cannot produce thick neck muscles by exercise alone and lack testosterone to make wide necks and square jaw lines. So for them, any appearance of a thickened neck and a wide jaw line could be a strong indicator of excess body fat, even when the whole of the body is not visible. There may also be a link to a phenomena called Boundary Extension (Gottesman & Intraub 1999) whereby viewers of photographs produce mental models of scenes beyond the confines of the border of the presented image. When viewing the portraits presented in the 2D vari-focal experiments, the participants were asked to estimate a body size that was outside of the picture area from minimal face and neck information. It is open to speculation to whether the two phenomena are linked and if a relationship can be demonstrated experimentally.

9.9 Summary of findings

The experiments undertaken during this period of research have demonstrated a number of effects in the perception of photographic images that have not been previously studied. These novel effects include:

- When comparing monoscopic and orthostereoscopic images, the binocular image is perceived as slimmer than otherwise identical monocular stimuli.
- The slimming effect of binocular disparity was observed in photographs of people, non-human objects and computer generated images.
- The slimming effect of binocular disparity follows a dose-response trend: as larger disparities were presented, smaller object sizes were observed.
- The area of the largest apparent change in the stereoscopic projections (the waist) was the area that was rendered with zero disparity and had the least differences (or best registration) when viewed without the polarizing glasses.
- Zero disparity binocular projected stimuli are seen as larger than life size.
- Orthostereoscopic images can be viewed for almost unlimited periods without apparent eyestrain.
- The TNO test for stereoacuity is not a useful predictor of observer performance when viewing orthostereoscopic images.

- The Frisby Stereotest for stereoacuity is a useful predictor of observer performance when viewing orthostereoscopic images.
- Changes in camera-to-subject distance will lead to changes in perceived bodyweight in portrait photography.
- Wide-angle lenses can have a slimming effect with camera proximities of less than one metre.
- Standard and Telephoto lenses can have a relative fattening effect at distances of one metre and beyond.
- The fattening effect reaches a peak at approximately 2.7 metres and the curve appears to level off at further distances.
- There is a powerful gender effect of photography in that males tend to be recorded as correct weight or underweight in most conditions, whereas females are seen as overweight in almost all normal photographic conditions.
- The slimmest female models (by BMI) produced the largest overweight estimations compared to their actual weight.
- When asked if photography was a reliable method of recording the appearance of people and general scenes there was a marked gender effect: Women were almost nine times more likely to doubt the reliability photography than male respondents.
- Synoptic 2D images are almost always seen as larger than life-size, even when the comparison target is a real object.

9.10 Future experiments.

Since these experiments were conceived, there have been a number of technological changes that could be used in revised experiments. Digital still image cameras have now reached a level of development whereby they are superior to film cameras in many respects. It would now be possible to photograph many more people with much greater control and consistency if these cameras were used. Also, digital images can be presented by computers. So it would be possible to program the image presentations using more conventional experimental psychophysics techniques and derive statistics from automated data-logging etc. Video projectors too have greatly improved in quality and could now replace the 35mm projectors used in these experiments. Large rear-projection screens are now available that can display polarised images. This would allow a viewer to use a chin rest that would place their eyes in the ideal viewpoint to see orthostereoscopic images without occluding the light path of the projectors. These rear-projection screens are also free of collimating elements, so can be viewed at distances of closer than one metre. Real images (or CGI) orthostereoscopic stimuli could be created to be shown life-size at distances of one metre.

These revised images would be usefully within the optimal range of normal stereoscopic vision, and not near the edge (1.68 metres) as was the case in all of the stereoscopic experiments reported here. Stimuli made to be shown at closer distances would have larger disparities and may show even larger slimming effects than could be demonstrated here. This should be the case even for stimuli that are smaller in virtual space, as they could be designed to occupy the same visual angles as the existing peanut stimuli. It should also be possible to design a stereoacuity test that could be projected as part of the experiment, rather than rely on a physical test taken in isolation from the main experimental apparatus and procedures.

The 2D "vari-focal" portraiture experiments could also be revised. Larger groups of matched BMI males and females would give easier to interpret statistics. Also, none of the models used in these experiment were assessed for apparent age by observer groups. If the apparent age in reality could be compared with apparent age in photographs (manipulated by gender and distance), it would be possible to find out if the anecdotal aging effect of photography could be quantified for the first time.

9.11 The Media and the Message

Figure 9.2. This is a half-size reproduction of the controversial photograph of Kate Winslet first shown in GQ magazine in February 2003 . The main image has been heavily re-touched to make the actress appear to be slimmer. However, the mirror image in the background was not given the same treatment and shows her waist, and hips and arms before re-touching.

The photograph in figure 9.2 (first reproduced here as figure 1.5) is almost unique in the field of retouching in that elements of the un-retouched photograph are reproduced in the same image. The re-touching is extensive and there is evidence of the bottom area, thighs, tummy and arms all being made dramatically slimmer. There is a high likelihood that the face and neck have also been re-touched as these are almost always enhanced by re-touching artists. The following text is typical of the media treatment of the GQ article and is reproduced from Hello Magazine, 10 JANUARY 2003 (20):

"RETOUCHING IS 'EXCESSIVE' SAYS SLIMLINE COVERGIRL KATE WINSLET"

A newly statuesque Kate Winslet towers on a pair of surprisingly slimline pins on the cover of next month's GQ magazine. The result of a pre-festive season crash diet? No, more like a little digital manipulation on the part of the magazine's art department.

The Titanic star, who has always made it clear that she doesn't believe a woman has to be slim to be attractive, has come under fire in some quarters for being involved in the project. According to her agent, however, although the actress had approved the original photos she was not consulted about the digital changes.

In fact the star confirmed her attitude in the interview accompanying the photos. "What is sexy?" she says. "All I know from the men I've ever spoken to is that they like girls to have an arse on them. So why is it that women think in order to be adored they have to be thin?"

Meanwhile, the magazine's editor, Dylan Jones, while admitting that the photos had been doctored, claims that Kate has actually lost a great deal of weight recently. "These pictures are not a million miles away from what she really looks like," he says. "Kate is currently thinner than I have ever seen her, petite and sexy."

Ms Winslet herself, however, doesn't seem to share his opinion. "The retouching is excessive. I do not look like that and more importantly I don't desire to look like that," she said. "I actually have a Polaroid that the photographer gave me on the day of the shoot... I can tell you they've reduced the size of my legs by about a third. For my money it looks pretty good the way it was taken."

For this thesis, the important element of the GQ story is the comment of its editor, Dylan Jones. He makes it clear that the job of the re-touching artist in this case was to make the actress look as slim in the photograph as she did when he met her. It seems likely that this type of image manipulation is common in the print media when the aim of the story is to show someone as they really are, or to remove temporary blemishes such as spots which may disappear a few days after the photo was taken but would be seen as a permanent disfigurement due to the archival nature of this type of image. Also, the size of enlargement in many magazines can be above life-size. This can allow the viewer to exert a level of scrutiny that is not possible with the real person and so emphasize flaws that cannot be seen when meeting the model in reality.

The logical conclusion that can be drawn from the GQ article and the work in this thesis is that *all 2D photography of women should be re-touched to counteract the fattening and unflattering effects of conventional 2D photography.* The assumption that photographs are always a reliable method of record of what the mediated viewer would have seen if they were present when the photograph was taken can no longer by sustained by argument or normal professional photographic practices. In the case of the Kate Winslet photographs, the re-touching artist appears to have excessively slimmed the image in a crude and clearly visible way. Yet if the treatment had been more subtle, it is possible that the actress might have not been aware that any intervention had taken place and considered the image to be a good representation of how she appeared at that time. There is a strong case to be made here for the original photographer to take responsibility for all retouching and to ensure that the image meets the required ethical standards in addition to the technical and artistic standards he has traditionally been responsible for. For picture editors and journalists, the converse is also true: Unless you know that an image has been corrected for the fattening effect, you should assume that it is not representative of a person's real body image. And if the image appears to show a thin model or a starving family in a refugee camp, be aware that by the time the effects of starvation are visible in a photograph, the people photographed may be frighteningly slim or even terminally underweight.

9.1.2 The effect of "Talking Movies" on photographic imaging techniques.

The psychologist James E. Cutting of Cornell University contributed a chapter in a collection of articles called *Moving image theory: Ecological considerations* on conventions in cinematography, from which the following quotation is taken:

Perceiving Scenes in Film and in the World

"In photography relative size and density are manipulated through the use of lenses Swedlund, 1981). Perhaps the most familiar example of issues concerning relative size occurs in portrait photography. Here the photographer typically stands back from the subject and uses a long lens. For 35-mm film, the standard lens has a focal length of 50 mm, and a lens with a focal length greater than about 100 mm is considered a long lens, also called a called telephoto lens. With a short focal length lens on the camera, the camera must be placed close to the person being photographed, with the result that the difference between camera-to-nose distance and camera-to-ear distance is great, and the person's nose appears large. With a long lens on the camera, the camera can be placed farther away from the person being photographed. With the camera farther away the difference between camera-to-nose and camera-to-ear distances becomes negligible, so the person's features appear close to their actual sizes. This is also one reason why most shot-reverse-shot sequences in cinematic dialogs are taken with relatively long lenses. They make the actors look better."

J. D. Anderson & B. F. Anderson (Eds.), Moving image theory: Ecological considerations. Carbondale, IL: Southern Illinois University Press.

Cutting appears to argue that the seemingly compressed perspective found with telephoto lenses is more accurate because it reduces the changes in apparent size caused by natural perspective. Conversely, "short focal length lenses" (normally called wide-angle lenses) are dismissed because they increase the differences in apparent size. Yet he does not consider the case of the "standard lens" in photography (see appendix H, section 5) which neither compresses or expands perspective and reproduces the appearance of natural size constancy when its images are viewed at life-size magnification.

A more likely explanation for the use of telephoto lenses in film production is for the purpose of noise suppression in "talking movies." Early camera shutters used in silent films were noisy and without extensive noise reduction strategies they would contaminate the recording of live dialogue. One strategy was to record the scene silently (figure 9.3) and to add the dialogue later in a dubbing studio.

Figure 9.3 Silent movies could use very close camera to subject distances because the noise of the shutter and transport mechanism was not recorded. (The Movie, encyclopaedia p.26 Orbis Press 1982)

This technique is still used today when background noise makes location sound recordings unusable. But most film dialogue is recorded live on a sound stage and an early technique was to hang a microphone above the actors from a movable boom. This had to be close to the actors because early microphones had low sensitivity. To exclude the boom from the shot, telephoto lenses were used to crop the image to just

Figure 9.4 A sound proof booth (with double or triple gazing) used in early talking pictures made increases in camera to subject distances inevitable, and close-up photography very difficult as this simulated scene shows. (The Movie, encyclopaedia p.22 Orbis Press 1982)

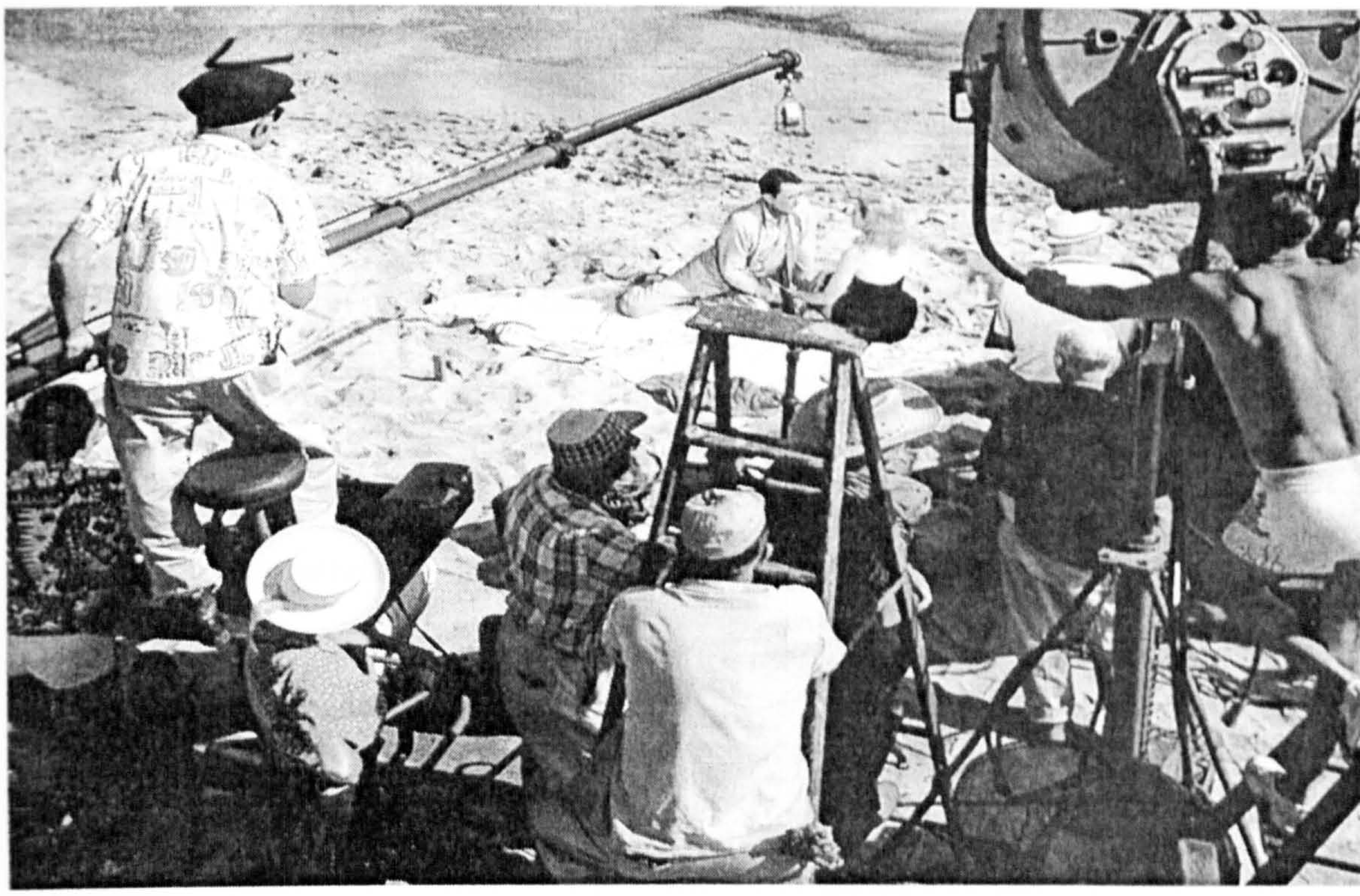


Figure 9.5 A very long boom microphone is used to record dialogue close to actors and away from camera noise (The Movie, encyclopaedia p.1 Orbis Press 1982).

above the actors' heads. This also had the advantage of allowing the camera to be positioned further away and placed in a sound-proof booth (figure 9.4). Later cameras were smaller, quieter and when fitted with sound deadening kits could be used at closer distances. They could now be moved around on tracks or cranes, but were rarely used in close proximity to an actor when sound was being recorded. Small and discrete radio microphones have made boom microphones less necessary in recent years. And modern film and video cameras are now so quiet that close-up photography is now possible with live dialogue. The recent technique of steady-cam photography (used on many location shoots in confined spaces) requires the use of wide-angle lenses, as telephoto lenses tend to amplify any imperfectly damped motion of the camera. Yet it is rare that a modern camera is used with actors at distances closer than two metres.

The significance of the convention of using "large" camera-to-subject distances is that all people imaged at these distances will be perceived as larger or fatter than they are in reality. The experiments reported in chapter six show not only that perceived bodyweights rise with increasing distances, but that the effect reaches its maximum between 2-4 metres. *This is precisely the range of distances used in moving and still photography by amateur and professional photographers.* The sexual dimorphism of photographic portraiture shows that photographs made at the most common distances are more fattening for women than they are for men. For professional models and actresses, this is highly likely to affect casting decisions that would favour unusually slim models: they would be cast to counter-act the fattening effects of the camera and this could be the explanation why models and actresses are much slimmer than women of the same age in the general population (BMA report 2000).

9.1.3 The implications for understanding the origins of BID in anorectics

The four experiments reported in chapter five in this thesis are the only ones conducted to date using size estimation of life-sized monoscopic and stereoscopic images of people with varying body sizes. They were first published in Harper & Latto (2001) and yet have not been cited by subsequent authors in the fields of eating disorders or body image distortion reviewed in Chapter 3. Had these authors been aware that changing between monocular and binocular viewing conditions had produced slimmer size estimations in these experiments, their general conclusion that "body percept" was not a significant issue in the understanding of Body Image Distortion (see, for eample, Williamson, et al, 2005) might be a more difficult assertion to make. The data reported in this thesis and in Harper & Latto (2001) also concurs with the original Brodie and Slade (1994) observations:

- "Individuals suffering from anorexia or bulimia tend to overestimate their physical body size and especially, bodily widths." Participants in the Harper and Latto studies also appeared to overestimate the widths of human and inanimate bodies only in monocular conditions

- "A tendency to overestimate body size is therefore not unique to anorexia or bulimia patients, so cannot be used solely as diagnostic criteria." Harper and Latto (2001) also observed in their experiments that the tendency to overestimate body size was not unique to a subgroup of participants, but common to all viewers of monocular stimuli when compared to otherwise identical stereoscopic/binocular versions of the same stimuli.

The conclusion made by Harper and Latto was that all people viewing monocular 2D images of bodies are therefore likely to significantly overestimate its size in relation to an otherwise identical stereoscopic image. An implication of this finding is that any person viewing a real body (their own, in a mirror perhaps) would have a tendency to overestimate its size if the percept was monocular. Clinical studies in underweight children (Dowdeswell, 1995) have pointed to relationships between reduced neurodevelopment and visual functions such as stereoacuity, colour vision, and contrast sensitivity. Singh (2005) also reported that malnutrition will reduce eye-hand coordination and stereoacuity. It has yet to be studied, but a testable hypothesis would that if stereo acuity failure where to happen without the conscious awareness of a person suffering from malnutrition, they may not be aware that the monocular nature of their vision would give them larger size percepts of their own image in a mirror. Thus, reported BID in some anorexia patients may be a direct result of the effects of malnutrition on their vision and have no "body concept" component to its appearance.

9.1.4 Future technology

As section 9.10 indicated, we need imaging systems that can show reality more like it appears to people present when the images are made. Imaging technology is developing very rapidly, with dramatic improvements in miniaturisation, convenience and resolution. However, there has been a much slower development in the technology required to produce images that avoid the perceptual distortions currently inherent in 2D images generated through the use of long focus lenses, large camera-to subject distances and the lack of orthostereoscopic information. The work that there has been on developing stereo imaging in the film industry has been driven by its potential entertainment value rather than its increased perceptual accuracy and in any case has often ignored the need to match the technology to the way the human visual system works. A wider appreciation of the perceptual distortions identified in this thesis may lead to a greater awareness of this need and lead to the production of imaging systems that are more life-like and more accurate than have been seen to date.

Bibliography & World Wide Web References

Allebeck, P., Hallberg, D., & Espmark, S. (1976). Body image: An apparatus for measuring disturbances in estimation of size and shape. *Journal of Psychosomatic Research*, 20, 583–589.

Ames, A. (1952) The Ames Demonstrations in Perception, *New York, Hafner Publishing*

Askevold, F. (1975). Measuring body image: Preliminary report on a new method. *Psychotherapy and Psychosomatics*, 26, 71–77.

Banton, T., & Levi, D.M. (1991). Binocular summation in vernier acuity. *Journal of the Optical Society of America A*, 8(4), 376-380.

Blake, R., Sloane, M. & Fox R. (1981). Further developments in binocular summation. *Perception & Psychophysics*, 30(3), 266-276.

Blakemore & M. Weston-Smith (Eds.), Images and Understanding (pp. 94-107). Cambridge, U.K.: *Cambridge University Press.*

Bowden, P. K., Touyz, S. W., Rodriguez, P. J., Hensley, R., & Beaumont, P. J. V. (1989). Distorting patient or distorting instrument? *British Journal of Psychiatry,* 155, 196–201.

Bradshaw M.F., Glennerster A, Rogers BJ (1996) The effect of display size on disparity scaling from differential perspective and vergence cues. *Vision Research* 36 1255 - 1265

Brodie, D. A., & Slade, P. D. (1988). The relationship between body-image and body-fat in adult women. *Psychological Medicine*, 18, 623–631

Brinded, P. M. J., Bushnell, J. A., McKenzie, J. M., & Wells, J. E. (1990). Body-image distortion revisited: Temporal instability of body image distortion in anorexia nervosa. *International Journal of Eating Disorders,* 9, 695–701.

Bruch, H. (1962). Perceptual and conceptual disturbances in anorexia nervosa. *Psychosomatic Medicine*, 24, 187-194.

Cash, T. F., & Brown, T. A. (1987). Body image in anorexia nervosa and bulimia nervosa: A review of theliterature. *Behavior Modification*, 11, 487–521.

Collins, J. K., Beumont, P. J. V., Touyz, S. W., Krass, J., Thompson, P., & Philips, T. (1987b). Variability in body shape perception in anorexic, bulimic, obese, and control subjects. *International Journal of Eating Disorders,* 6, 633–638.

Cutting, J.E. (2004). Perceiving scenes in film and in the world. In J.D. Anderson & B.F. Anderson (Eds.) Moving image theory: Ecological considerations.Carbondale, IL: *University of Southern Illinois Press*

Duke, P.A. & Howard I.P. (2004) Vertical-disparity gradients are processed independently in different depth planes. *Vision Research* 45 2025–2035

Dumbreck, A.A. (1995) 3DTV Production Grammar, a report for the Independent Television Commission, AEA Technology Report TS/DOC/829.

Ee, R. van, Banks M.S., Backus B.T.: Perceived visual direction near an occluder. Vision Research 39, 4085-4097, 1999

Erkelens, C.J., van Ee R.: The role of the cyclopean eye in vision: sometimes inappropriate, always irrelevant. *Vision Research* 42, 1157-1163, 2002.

Erkelens, C. J., & van de Grind, W. A. (1994). Binocular visual direction. *Vision Research*, 34, 2963–2969

Evans, R M. (1959). Eye, Film, and Camera in Color Photography. Huntington, N.Y., U.S.A.: Krieger.

Evans, R.M. (1974). The Perception of Color. New York, N.Y., U.S.A.: Wiley.

Evans, R. M. (1976) Using photography to preserve evidence. Kodak Publication M-2

Field, A. (2005) Discovering Statistics using SPSS. 2nd edition. *Sage*: London.

Fisher, B & Poggio, G.F. (1979) Depth sensitivity of binocular cortical neurons of behaving monkeys. *Proc. R. Soc. Lond.* B. 204, 409-414

Franzen, U., Florin, I., Schneider, S., & Meier, M. (1988). Distorted body image in bulimic women. *Journal of Psychosomatic Research*, 32, 445–450.

Freeman, J., Avons, S.E., Meddis, R., & Pearson, D.E. (2000). Behavioural realism: Using postural responses to estimate presence. *Presence: Teleoperators and Virtual Environments*, 9(2), 149-164.

Freeman, J., Avons, S.E., Pearson, D., & IJsselsteijn, W. (1999). Effects of sensory information and prior experience on direct subjective ratings of presence. *Presence: Teleoperators and Virtual Environments*, 8(1), 1-13.

Freeman, R. J., Thomas, C. D., Solyom, L., & Hunter, M. A. (1984). A modified video camera for measuring body image distortion: Technical description and reliability. *Psychological Medicine*, 14, 411–416.

Frisch, R. E. (1987). Body fat, menarche, fitness and fertility. *Human Reproduction,* Vol. 2, No. 6, pp. 521-533,

Furnham, A., & Alibhai, N. (1983). Cross-cultural differences in the perception of female body-shapes. Psychological Medicine, 13, 829-837.

Gardner, R. M. & Bokenkamp, E. D.(1996) The role of sensory and non-sensory factors in body size estimations of eating disorder subjects. *Journal of Clinical Psychology* 52, 3-15.

Garfinkel, P. E., Moldofsky, H., Garner, D. M., Stancer, H. C., & Coscina, D. V. (1978). Body awareness inanorexia nervosa: Disturbances in ''body image'' and ''satiety.'' *Psychosomatic Medicine*, 40, 487–498.

Garner, D. M., Garfinkel, P. E., Stancer, H. C., & Moldofsky, H. (1976). Body image disturbances in anorexia nervosa and obesity. Psychosomatic Medicine, 38, 327–336.

Garner, D. M., Olmsted, M. P., & Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders*, 2, 15–34.

Gardner, R.M. & Moncrieff, C. (1988). Body image distortion in anorexics as a non-sensory phenomenon: a signal detection approach. *Journal of Clinical Psychology*, 44, 101-107.

Gluckman, M. L., & Hirsch, J. (1969). The response of obese patients to weight reduction: III. The perception of body size. Psychosomatic Medicine, 31, 1–7.

Gottesman, C.V., & Intraub, H. (1999). Wide-angle memories of close-up scenes: A demonstration of boundary extension. *Behavior Research Methods, Instruments & Computers*, 31, 86-93.

Gregory, R. L. (1963) "Distortion of Visual Space as Inappropriate Constancy Scaling," *Nature*, 119, 678

Gunby, L. (2000, September 14). Notes & Queries. The Guardian, G2, p.16

Harper, B. & Latto, R. (2001). Cyclopean vision, size estimation and presence in orthostereoscopic images. *Presence: Teleoperators and Virtual Environments*, 10(3), 312-330.

Hendrix, C., & Barfield, W. (1996). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators and Virtual Environments*, 5(3), 274-289.

Hagen, M. A. (1976). Influence of picture surface and station point on the ability to compensate for oblique view in pictorial perception. *Developmental Psychology*, 12(1), 57-63.

Hagen, M. A. (editor). (1980). The Perception of Pictures: Volume 1, Alberti's Window: The projective model of pictorial information. New York, U.S.A.: Academic Press.

Hagen, M. A. (editor). (1980). The Perception of Pictures: Volume 2, Dürer's Devices: Beyond the projective model of pictures. New York, U.S.A. Academic Press

Hering E (1861) Beitrage zur Physiologie, Heft 1. Englemann, Leipzig

Hering, E. (1879/1942). Spatial sense and movements of the eye (translated by Raddle, A.), Baltimore, MD: *American Academy of Optometry.*

Holway, A. H., & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *American Journal of Psychology*, 54, 21-37.

Howard, I.P. (2002). Seeing in Depth. Volume 1 Basic Mechanisms. Thornhill, Ontario, Canada: I.Porteus.

Howard, I.P., & Rogers B.J. (2002). Seeing in Depth. Volume 2 Depth Perception. Thornhill, Ontario, Canada: I.Porteus.

Huon, G. F., & Brown, L. B. (1986). Body images in anorexia nervosa and bulimia nervosa. *International Journalof Eating Disorders*, 5, 421–439.

IJsselsteijn, W., de Ridder, H., Hamberg, R., Bouwhuis, D., & Freeman, J. (1998). Perceived depth and the feeling of presence in 3DTV. *Displays*, 18(4), 207-214.

Intraub, H. (2004) Anticipatory spatial representation of 3D regions explored by sighted observers and a deaf and blind observer. *Cognition*, 94, 19-37.

Intraub, H., & Richardson, M. (1989). Wide-angle memories of close-up scenes. Journal of Experimental Psychology*: Learning, Memory, & Cognition*, 15, 179-187

Intraub, H., Gottesman C.V., Willey E.V., Zuk I.J., 1996. Boundary extension for briefly glimpsed photographs: Do common perceptual processes result in unexpected memory distortions? *Journal of Memory & Language* 35, 118-134

Jones, R.K., & Lee, D.N. (1981). Why two eyes are better than one. Journal of Experimental Psychology: *Human Perception & Performance*, 7(1), 30-40.

Julesz, B.(1971). Foundations of Cyclopean Perception. Chicago, Ohio U.S.A.: University of Chicago Press.

Kelly, L., (1998, September 20). Q and A. Sunday Mirror Personal Magazine, p.3.

Künnapas, T. M. (1957a). Interocular differences in the vertical-horizontal illusion. *Acta psychologica*, 13, 253-259.

Künnapas, T. M. (1957b). The vertical-horizontal illusion and the visual field. *Journal of Experimental Psychology*, 53, 405-407.

Künnapas, T. M. (1957c). Vertical-horizontal illusion and the surrounding field. *Acta Psychologica*, 13, 35-42

Lipton, L. (1982). Foundations of Stereoscopic Cinema, A Study in Depth, VAN NOSTR AND REINHOLD Company, New York

Khokhotva, M., Ono, H., & Mapp, A. P. (2005). The cyclopean eye is relevant for predicting visual direction. *Vision Research*, 45, 2339–2345.

Koenderink, J.J. (1998). Pictorial relief. Philosophical *Transactions of The Royal Society of London*, 356(1740), 1071-1086.

Koenderink, J.J., van Doorn.A.J. & Kappers, A.M.L.(1994). On so–called paradoxical monocular stereoscopy. *Perception*, 23(5), 583-594.

Langford, M. (1989). Advanced Photography. London, U.K.: Focal Press.

Maisey, D.M., Vale, E.L.E, Cornellison, P.L., & Tovée, M.J. (1999). Characteristics of male attractiveness for women. *The Lancet*, 353, 1500.

Marks, D.L., Stack, R.A., Brady, D.J., Munson Jr., D.C., & Brady, R.B. (1999). Visible cone-beam tomography with a lensless interferometric camera. *Science* 284, 2164-2166.

Maunsell, J. & Van Essen, D. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neuroscience*, 3(12), 2563-2586.

Muggeridge, M. (1977) Christ and the Media. Sevenoaks, Kent, U.K.:Hodder & Stoughton.p.30

Nakayama, K. and Shimojo, S. (1990) DaVinci stereopsis: depth and subjective occluding contours from unpaired image points. *Vision Research*, 30, 1811-1825,

Norris, D. L. (1984). The effects of mirror confrontation on self-estimation of body dimensions in anorexia nervosa, bulimia, and two control groups. *Psychological Medicine*, 14, 835–842

Ono, H., Mapp, A. P., & Howard, I. P. (2002). The cyclopean eye in vision: the new and old data continue to hit you right between the eyes. *Vision Research*, 42, 1307–1324

Perrett, D., Harries. M., Mistlin, A.J. & Chitty, A.J. (1990). Three stages in the classification of body movements by visual neurons. In H. Barlow, C. Blakemore & M. Weston-Smith (Eds.), *Images and Understanding* (pp. 94-107). Cambridge, U.K.: Cambridge University Press.

Prinzmetal, W., & Gettleman, L. (1993). Vertical-horizontal illusion: One eye is better than two. Perception & Psychophysics, 53(1), 81-88.

Probst, M., Braet, C., Vandereycken, W., De Vos, P., Van Coppenolle, H., & Verhofstadt-Dene`ve, L. (1995). Body size estimation in obese early adolescents: A controlled study with the video distortion method. *InternationalJournal of Obesity*, 19, 820–824

Proctor, L., & Morley, S. (1986). ''Demand characteristics'' in body-size estimation in anorexia nervosa. BritishJournal of Psychiatry, 149, 113–118.

Reitman, E. E., & Cleveland, S. E. (1964). Changes in body image following sensory deprivation in schizophrenic control groups. Journal of Abnormal & Social Psychology, 68, 168-176.

Richards, W. (1971) Anomalous stereoscopic depth perception. J.opt. Soc. Am. 61, 410-414

Robinson, T., Killen, J., Litt, I., Hammer, L., Wilson, D., Haydel, F., Hayward, C., & Taylor, C. (1996). Ethnicity and Body Dissatisfaction: Are Hispanic and Asian Girls at Increased Risk for Eating Disorders? Journal of Adolescent Health, 19,(6), 384-393.

Rogers, B. J., & Bradshaw, M. F. (1995). Disparity scaling and the perception of frontoparallel surfaces. Perception, 24(2), 155-179.

Rosenblum, N.(1984). A World History of Photography. New York: *Abbeville Press*,

Ruff, G. A., & Barrios, B. A. (1986). Realistic assessment of body image. *Behavioral Assessment*, 8, 237–251.

Schilder P (1935).: Image and Appearance of Human Body. London, Kagan, Paul, Trench, Trubner & Company,.

Secord, P. F., & Jourard, S. (1953). The appraisal of body cathexsis: Body cathexsis and the self. *Journal of Consulting and Clinical Psychology*, 17, 343–347.

Skrzypek S, Wehmeier PM, Remschmidt H. (2001) Body image assessment using body size estimation in recent studies on anorexia nervosa. A brief review. European Child Adolescent Psychiatry, 10 (4), 215-21. PMID 11794546.

Slade P. D., (1988) Body image in anorexia nervosa British Journal of *Psychiatry* Supplement, Jul, (2): 20-2

Slade, P. D., & Russell, G. F. (1973b). Experimental investigations of bodily perception in anorexia nervosa and obesity. *Psychotherapy and Psychosomatics*, 22, 359–363.

Slade, P,D. & Brodie, P.A. (1994) Body image distortion and eating disorder: A reconceptualization based on the recent literature. European Eating Disorders Review. *2*, 32–44

Singh, D. (1993) Adaptive significance of female physical attractiveness: Role of waist

- to-hip ratio. *Journal of Personality and Social Psychology*, 65, 293-307.

Spottiswoode, R., Spottiswoode, N.L., & Smith, C. (1952, October). Basic principles of the three-dimensional film. *Journal of the Society of Motion Picture & Television Engineers*, 59, 249-285.

Swami, V., & Tovée, M.J. (2005a). Female physical attractiveness in Britain and Malaysia: A cross-cultural study. *Body Image*, 2, 115-128.

Swami, V., & Tovée, M.J. (2005b). Male physical attractiveness in Britain and Malaysia: A cross-cultural study. *Body Image*, 2, 383-393.

Taylor, D.W & Boring, E.G. (1942). Apparent visual size as a function of distance for monocular observers, *The American Journal of Psychology* , 55, 102-105

Thompson, J. K. (1990). Body image disturbance: Assessment and treatment. New York: Pergamon.

Thompson, J. K., & Spana, R. E. (1988). The adjustable light beam method for the assessment of size estimation accuracy: Description, psychometric and normative data. *International Journal of Eating Disorders*, 7, 521-526

TNO Test for Stereoscopic Vision (10th ed.) (1972). Veenendaal, The Netherlands: Laméris Ootech B.V..

Tovée, M.J., Hancock, P., Mahmoodi, S., Singleton B.R.R., & Cornellison, P.L. (2002). Human female attractivenes: Waveform analysis of body shape. *Proceedings of the royal society of London* (B), 269, 2205-2213.

Traub, A. C., & Orbach, J. (1964). Psychophysical studies of body image: I. The adjustable body-distorting mirror. *Archives of General Psychiatry*, 11, 53–66.

Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience*, 8, 1401–1410.

Warner, M. (1995, Summer). Stealing souls and catching shadows. tate: the art magazine, 6, 40-46.

Williamson, D.A., White, M.A., York-Crowe, E., Stewart, T.M. (2004). Cognitive-behavioral theories of eating disorders. *Behavior Modification,* 28(6), 711-738.

Yamanoue, H. (1997, April). The relationship between size distortion and shooting conditions for stereoscopic images. *Journal of the Society of Motion Picture and Television Engineers*, 106, 225-232.

Yamanoue, H. Okui, M., & Yuyama, I. (2000). A study on the relationship between shooting conditions and cardboard effect of stereoscopic images. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(3), 411-416.

World Wide Web References

(All web references were retrieved between March and July 2006)

(1) http://www.hrc.utexas.edu/exhibitions/permanent/wfp/4.html

(2) As if you were there; Matching Machine Vision to Human Vision Richard Jay Solomon, April 2002. http://www.hybridvigor.net/publications.pl?s=human

(3) http://www.usatoday.com/money/industries/technology/patent-chart.htm

(4) http://www.kodak.com/US/en/corp/researchDevelopment/achievements/

(5) The Right White Light. Avoid metameric pressure and make colors match. Brian P. Lawler. April 27, 2000 http://www.creativepro.com/story/feature/5475.html

(6) http://www.microsoft.com/whdc/archive/convrg1.mspx

(7) http://www.guardian.co.uk/notesandqueries/query/0,,-12228,00.html

(8) http://en.wikipedia.org/wiki/Telescope

(9) http://www.guardian.co.uk/women/story/0,,257823,00.html

(10) http://www.guardian.co.uk/Archive/Article/0,4273,4031595,00.html

(11) http://www.prdomain.com/feature/feature_details.asp?id=187&area=J&typ=A

(12) http://www.eating-disorders.org.uk/docs/media.doc

(13) http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/HENKEL/research/stereo/Cyclops/

(14) http://www.ient.rwth-aachen.de/engl/team/ohm/publi/ei99.pdf.

(15) http://www.stereoscopy.com/isu/glossary-a.html

(16) http://www.stereoscopic.org/library/foundationsbook3a.gif

(17) http://www.alleydog.com/glossary/definition.cfm?term=Ecological%20Validity

(18) http://news.bbc.co.uk/1/hi/entertainment/2643777.stm

(19) http://www.frisbystereotest.co.uk/FST%20Page1.htm

(20) http://www.hellomagazine.com/film/2003/01/10/katewinslet/

(21) http://www.visualexpert.com/Resources/photoevidence.html

APPENDIX

A Paper 1: *Cyclopean Vision, Size Estimation and Presence in Orthostereoscopic Images by Bernard Harper and Richard Latto,* Presence: Teleoperators and Virtual Environments, MIT Press Vol. 10, No. 3, June 2001

B Paper 2. *Virtually There*? A Vision on Presence Research by Wijnand Ijsselsteijn & Bernard Harper, published by the Presence Research Working and Information Society and Technolgy, IST 2000-31014 December 2001

C Paper 3: *The Non-Realistic Nature of Photography: Some More Reasons Why Turner Was Wrong by* Richard Latto and Bernard Harper. Leonardo Journal 2007 (in press).

D Two articles from the Guardian News paper; Thin, thinner, thinnest by
Anita Chaudhuri, Monday October 18, 1999. Thin end of the wedge by Hadley Freeman, Wednesday June 21, 2000

E Programming rules for Vertical Disparity correction in Orthosterescopic Images by Philip Berridge

F Script for Experiment Six proceedure

G An appreciation of synoptic viewing using The Cyclopter, by Terry Pope

H Pilot studies into perceptually distorted images and the

limitation of Photographic Reality

APPENDIX A

BEST COPY

AVAILABLE

Variable print quality

Bernard Harper
bernard.harper@liverpool.ac.uk

Richard Latto
Department of Psychology
University of Liverpool
Liverpool, L69 7ZA
United Kingdom

Cyclopean Vision, Size Estimation, and Presence in Orthostereoscopic Images

Abstract

Stereo scene capture and generation is an important facet of presence research in that stereoscopic images have been linked to naturalness as a component of reported presence. Three-dimensional images can be captured and presented in many ways, but it is rare that the most simple and "natural" method is used: full orthostereoscopic image capture and projection. This technique mimics as closely as possible the geometry of the human visual system and uses convergent axis stereography with the cameras separated by the human interocular distance. It simulates human viewing angles, magnification, and convergences so that the point of zero disparity in the captured scene is reproduced without disparity in the display. In a series of experiments, we have used this technique to investigate body image distortion in photographic images. Three psychophysical experiments compared size, weight, or shape estimations (perceived waist-hip ratio) in 2-D and 3-D images for the human form and real or virtual abstract shapes. In all cases, there was a relative slimming effect of binocular disparity. A well-known photographic distortion is the perspective flattening effect of telephoto lenses. A fourth psychophysical experiment using photographic portraits taken at different distances found a fattening effect with telephoto lenses and a slimming effect with wide-angle lenses. We conclude that, where possible, photographic inputs to the visual system should allow it to generate the cyclopean point of view by which we normally see the world. This is best achieved by viewing images made with full orthostereoscopic capture and display geometry. The technique can result in more-accurate estimations of object shape or size and control of ocular suppression. These are assets that have particular utility in the generation of realistic virtual environments.

1 Introduction

Photographers are sometimes aware that the scenes they see with their normal direct vision will differ significantly from the 2-D representations produced when the scenes are imaged and transferred to photographic paper or a projection screen. Almost everything about the originally captured scene is conveyed in a modified or degraded form. The descriptions of classical image aberrations (for example, Langford (1989, ch. 2)) cover the effects of simple uncorrected lenses on only the shape or color of the imaged scene. However, there are many other changes in the transition from

Presence, Vol. 10, No. 3, June 2001, 312–330

the reality to the image. One of the best known (and most disconcerting to the subject) is the fattening effect of photography.[1]

The most obvious loss in conventional imaging is presence derived from stereo depth information (Freeman, Avons, Meddis, & Pearson, 2000; Freeman, Avons, Pearson, & IJsselsteijn, 1999; Hendrix & Barfield, 1996; IJsselsteijn, de Ridder, Hamberg, Bouwhuis, & Freeman, 1998). However, there are other more subtle effects of which we are often unaware that are worthy of note. Peripheral vision objects and scaling cues are usually excluded from photographic images. Photography almost always fails to reproduce scenes at their original (same-size) magnification. Even when this is achieved, photography cannot reproduce the detail that can be seen with normal vision from the original viewpoint while maintaining the angle of view. Natural brightness ranges are immensely difficult to reproduce because each image generation can add contrast and lose shadow or highlight detail. Accurate color reproduction is also almost impossible with conventional imaging, and subject color failure can be found in most types of image capture. These "fidelity failures" are often corrected for by trial and error or custom and practice techniques derived from professional knowledge (Langford, 1989, ch. 8).

The only thing that appears to be unchanged in a photograph is the point of view. However, the single-point perspective that makes a photographic image appear to be an accurate representation of the original scene can also convey inaccurate object information. Humans, too, perceive the world from a single-point perspective. By the process of cyclopean vision (Julesz, 1971), we see the world through a "cyclopean eye" that generates a single artificial viewpoint from a location midway between each real eye. In normal human vision, the processes of foveal convergence, accommodation, and stereo fusion allow the brain to construct a new perspective that differs from those seen by either eye individually. This cyclopean point of view appears to be similar to a 2-D photographic perspective. However, a single-lens system cannot reproduce the way in which we can focus/fuse on an object with two eyes and see both diverging and converging optical paths (figure 1) from the same position. With close-up objects, we have the ability to see the normal photographic perspective and also have "look-around" vision from a single head position. The result is that close-up objects viewed stereoscopically occlude less of the background than do their 2-D photographic equivalents. This paper investigates the possibility that the failure to reproduce this geometry in a display is a major cause of the fattening effects associated with conventional photographic images. A previous study (Yamanoue, 1997) found evidence of changes in size estimations in stereoscopic conditions. His experiments linked widening camera lens interaxial separations to smaller size perception and the "puppet theater" effect. He used direct observation of a mannequin and compared it with a same-size, parallel-imaged stereo video reproduction. In a later paper, Yamanoue, Okui, and Yuyama (2000) supported the use of lens separations and magnifications similar to those of the human visual system in order to reduce the appearance of an image artifact known as the "cardboard effect." In the stereo experiments reported here, only photographic images were viewed and only the stereoscopic disparity and convergences were changed.

In psychophysical experiments, monocular vision has consistently been linked to lower performance when compared to binocular vision, with the exception of the horizontal-vertical illusion (Prinzmetal & Gettleman, 1993). Tasks such as luminance increment detection, contrast sensitivity with sine wave gratings, color discrimination, vernier acuity, letter identification, and visual search (Banton & Levi, 1991; Blake, Sloane, & Fox, 1981; Jones & Lee, 1981) all show improved performance in the binocular condition. It is argued here that, whenever the visual system is presented with im-

1. It is commonly said in the fields of photography, film, and television that the camera "can put ten pounds on you." Yet we can find no academic reference for this effect, despite researching this phenomena with a number of institutions such as the British Journal of Photography, the Independent Television Commission, the Moving Image Society (BKSTS), the Royal Television Society, members of the American Society of Cinematographers, and more-conventional scientific resources. Although distortions are regularly mentioned anecdotally (Gunby, 2000; Kelly, 1998; Warner, 1995), until the present study, it appears that no one has examined the fattening effect of photography in a systematic way.

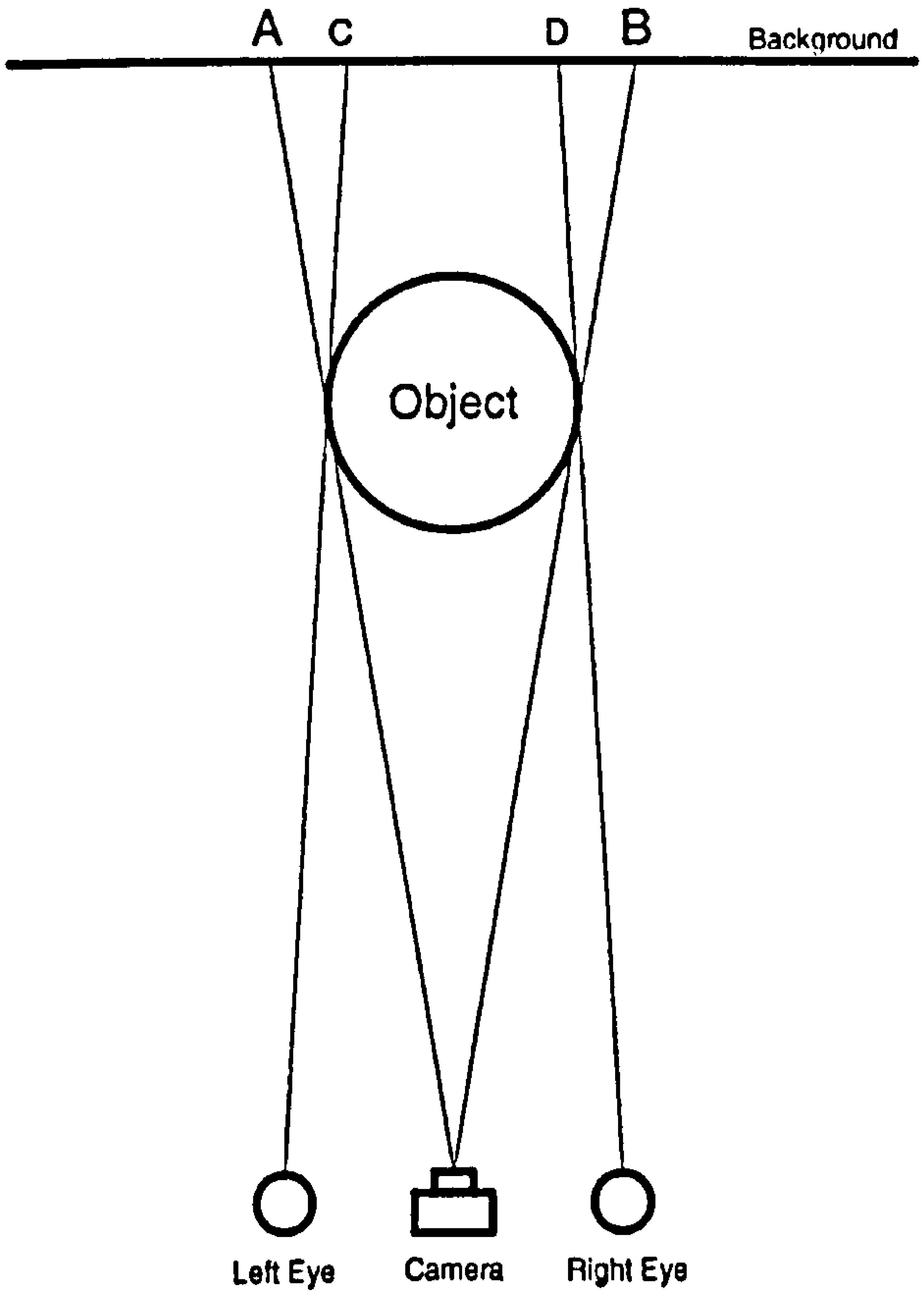


Figure 1. *The difference between a camera point of view and human stereo vision from the same position. The viewed object occludes more of the background in a 2-D photograph (AB) than in stereo vision (CD).*

ages that do not allow it to form a normal cyclopean view, predictable perceptual disturbances will occur: the display medium will be flawed in its ability to convey objects and people in their original proportions, size, and background occlusion characteristics. We propose that only a full orthostereoscopic capture and display system (Spottiswoode, Spottiswoode, & Smith, 1952)[2] can reproduce natural viewing geometries and provide a more lifelike visual experience.

2 General Method

The experiments reported here use orthostereoscopic imaging to investigate the distorting effects of photographic images. Two-dimensional images are less able to convey volumetric, contour, or shading information and can generate monocular optical illusions that fail with direct stereo vision (such as an Ames room). The hypothesis is that 2-D images distort because they do not present object information in the same way as a real object would under direct human observation. To minimize possible photographic distortions, the experiments use stereo image-capture geometry that is as close as possible to that of the human visual system. Conventional 3-D photography, which we are grouping under the term *parallel stereography*,[3] is inadequate because most stereo camera/display arrangements are not designed to match the geometry of human stereo vision.[4]

It was considered that viewing comfort should have a high priority in the presentations. There are limits (Panum's fusional area) to how far out of horizontal or vertical alignment binocular stimuli can be before there is loss of fusion and diplopia or suppression of one image (Howard & Rogers, 1995). We decided that the point

2. While recognizing the theoretical advantages of orthostereoscopic imaging and that this technique was "the condition of perfect image reproduction," Spottiswoode et al. (1952, p. 263) argued that this would constrain the artistic freedom of directors and cinematographers. So, their pragmatic solution was to reject these constraints for more-flexible and practical combinations of magnification, lens interaxial separations, and alignments. This often meant that images were captured using long telephoto lenses, wider-than-normal lens interaxials, "narrower than natural" convergences, and that the stereo window of reproduction was often placed behind the plane of focus/screen plane. They also considered that the primary orthostereoscopic conditions were 65 mm interaxial separation and same-size magnification.

3. Parallel stereography in this paper refers to stereo image-capture geometries that do not converge the lens axes on the center of focus and interest at the object plane and generate a double image at the plane of reproduction.

4. Almost all stereography uses different combinations of lens interaxial separations, magnifications, and convergences from those the human visual system would use when viewing the original scene. For instance, the average human interocular distance is approximately 65 mm, but stereo camera separations are often much wider than this. Also, they usually fail to reproduce the point of zero disparity from the original scene with zero disparity in the display. This means that they show a single point from the captured scene as two points on the screen and the viewers are required to "force fuse" these points to form a single stereo image.

of focus for each camera should coincide with the convergence point for each lens axis, and that this must be reproduced as a point of zero disparity in the display. This alignment was most likely to give comfortable viewing because, when the points of each camera's focus are horizontally aligned in the display, the center of interest (a face, for instance) appears as a single image. Zero separation in the display (no double image at the center of interest) means that relatively flat objects can be viewed successfully without polarizing spectacles. Typically, this condition has a high degree of 2-D compatibility as only the out-of-focus areas are not aligned at the screen. Polarizing spectacles allow the viewer to separate these areas into discrete channels by which they can then perceive the original scene depth. The principle that underlies all of the stereo experiments reported here is that orthostereoscopic images are presented to the participants for comparison with 2-D images from the same viewpoint and camera to subject distance. In practice, this means that, when participants are making size or shape judgments under experimental conditions, they are presented with images in which the only differences are of disparity.

2.1 The Stereo Camera

In experiment 1 and 2, a stereo camera was constructed using two Olympus OM1 cameras mounted vertically on a common baseplate and tripod mount. Standard 50 mm, f1.8 lenses were used to closely approximate the human eye's angle of view and magnification. The lens separation was 64 mm and the optical axis of each lens was converged on the point of focus 1.68 m away. The framing was for adults of normal height; the horizontal crop lines falling above the knees to just above head height (figure 2). Each shutter was triggered by a dual cable release staged to fire the flash lighting on the opening of the second curtain to ensure correct synchronization. This method allowed for bright, even illumination of the subject and for consistent exposures using small apertures (f16). It also ensured that the maximum depth of field and apparent sharpness would be recorded.

The images were recorded onto high-resolution Fuji

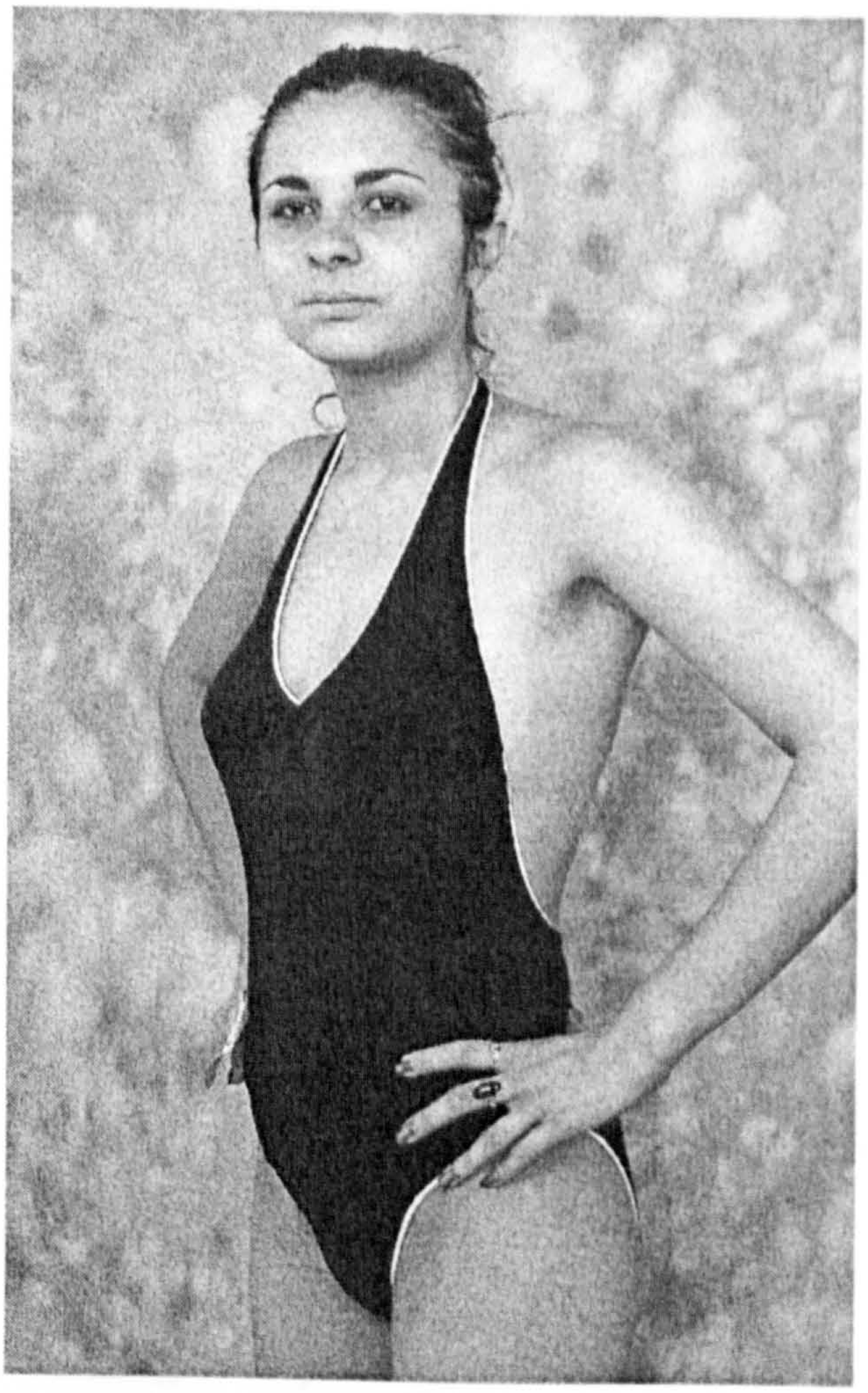


Figure 2. *Experiment 1: typical swimsuit image.*

50 ASA transparency film. The transparencies were processed, selected for technical quality, and mounted into annotated 35 mm registration mounts. The left camera images were also copied to same-size magnification, and two color-matched copies were produced for synoptic[5] presentation. The exposures were carefully controlled, because the stereo images were intended for two-channel projection using cross-polarized filters and viewing

5. Following Koenderink, van Doorn, and Kappers (1994), we are using the term *synoptic* to describe the situation in which both eyes see exactly the same image with no binocular disparity, as in viewing a photograph, television screen, or a landscape at infinity.

through standard polarizing spectacles. This technique allows high-quality, full-color stereo images to be seen, but it causes a 50% loss of image brightness. Some of this brightness loss can be recovered because the technique requires the use of a polarization-maintaining (metalisized) projection screen. These are often used simply as high-brightness screens and, together with illumination by two projectors, this ensures a projected image of adequate brightness.

2.2 Stereo Projection

The transparencies were projected onto the metalisized screen using two carousel-type (Kodak Ektar) projectors with matched Kodak f2.8, 85 mm lenses. Because of their large size, the projectors could not be mounted side by side for correct orthostereoscopic projection, so a surface-silvered mirror was used to establish the correct optical path (figure 3). The right projector images were loaded normally, but the left projector images were laterally reversed to compensate for the mirror reversal in its optical path. Calibration images were then projected to same-size scale so that the projected model's interocular distance and height measured on the screen closely matched the measurements taken from the real person.

Side-by-side projection like this allows for the stereo window in which objects and scenes are reproduced to be easily moved towards or away from the viewer. For instance, it is possible by cross-converging the projectors (that is, moving one image horizontally) to place the background plane onto the projection screen and have the object appear to be reproduced in virtual space at the original camera-to-object distance. The projectors can also be diverged so as to move the object/stereo window behind the plane of reproduction. However, both of these alignments would require the images on the screen to be presented out of registration (figure 4). We speculated that this could cause the viewer to see objects as slimmer than they really are, because it might affect their perception of the true object boundary as it occludes the background. Incorrect vertical or rotational registration also might cause shape misperception (figure 5) for the same reason. So all of the images in

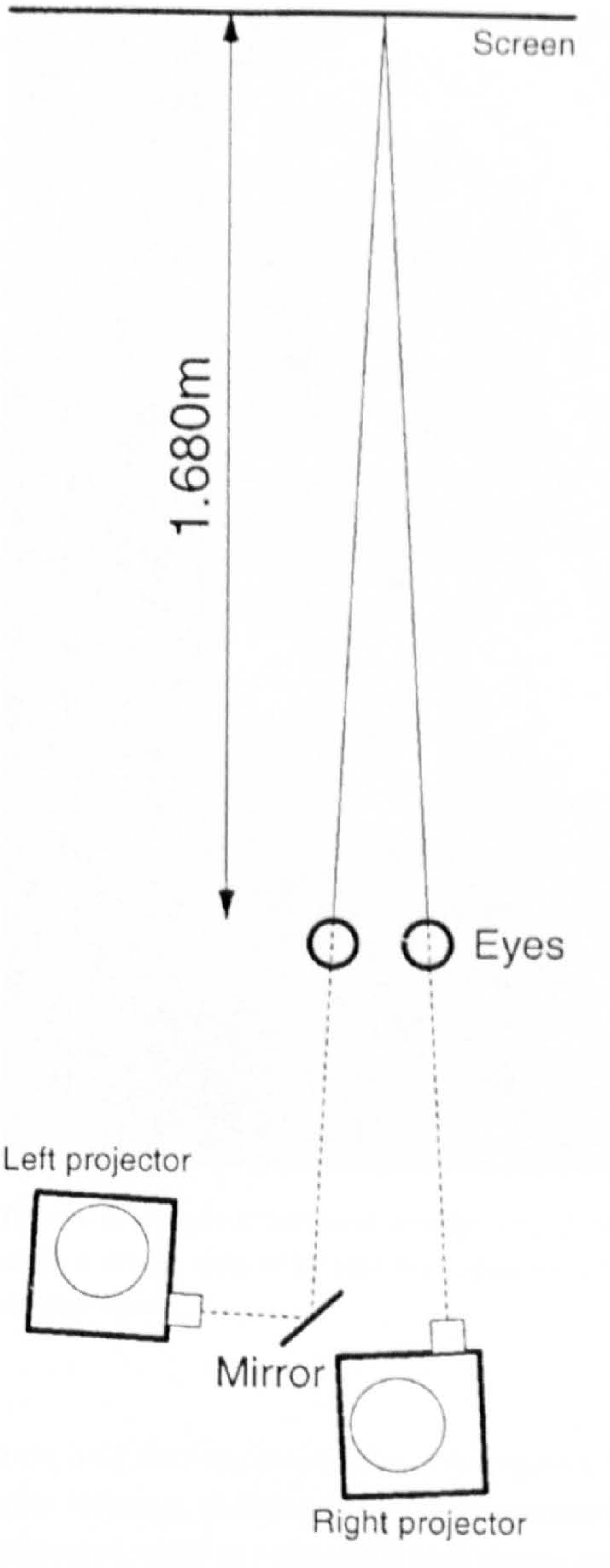


Figure 3. *Plan view of the projector alignment and viewing position for experiments 1 and 3. The viewers were positioned below the projectors' lenses to avoid occluding the image.*

these experiments were presented so that the vertical and horizontal registration of the point of interest/focus were of zero disparity at the plane of reproduction. Successful stereo projection also requires that image cross-talk (whereby one image channel can "leak" into another) be kept to a minimum. This can be achieved

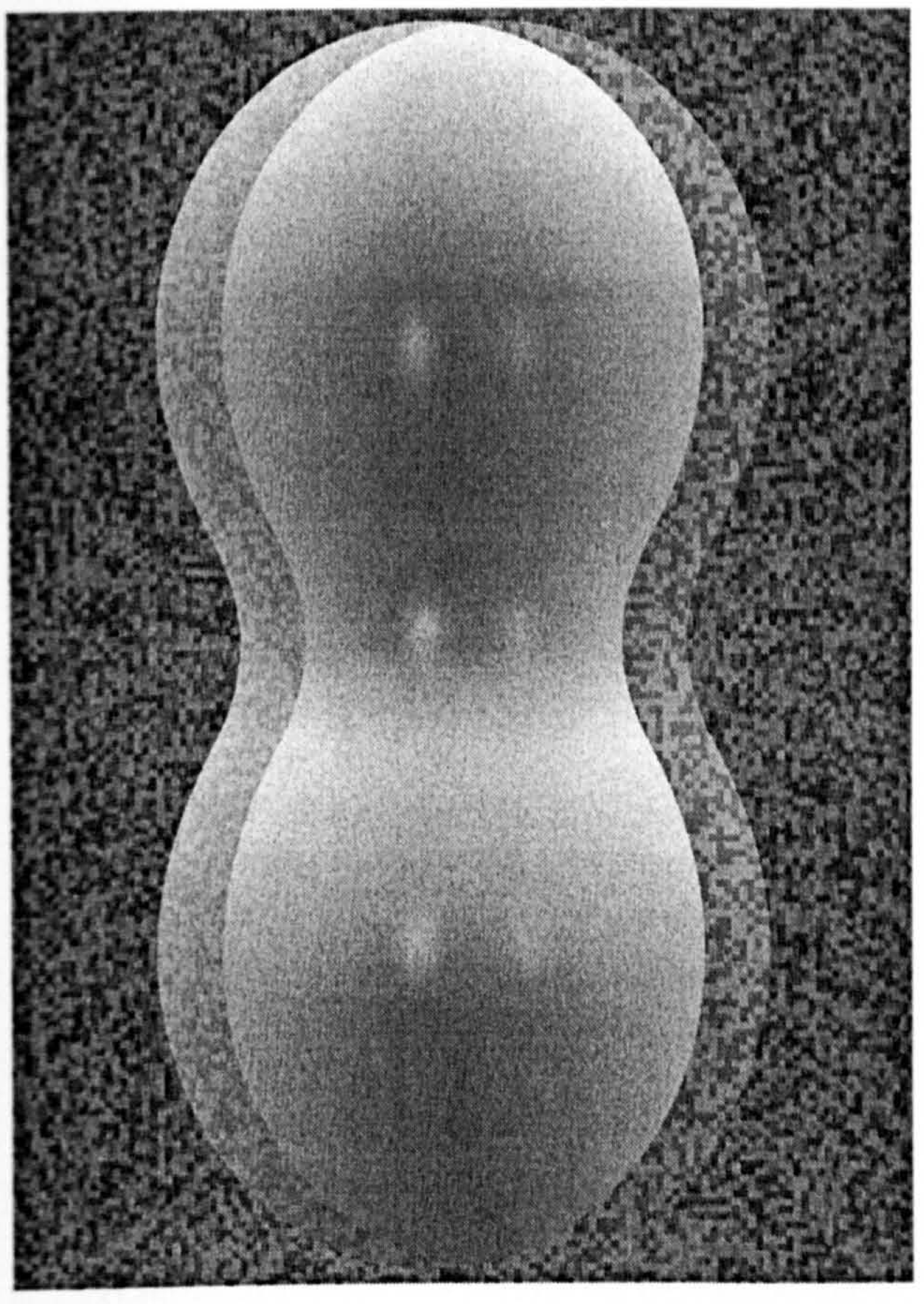


Figure 4. *Horizontal misalignment of the stereo window could cause a slimming effect by confusing viewers as to the true object boundary.*

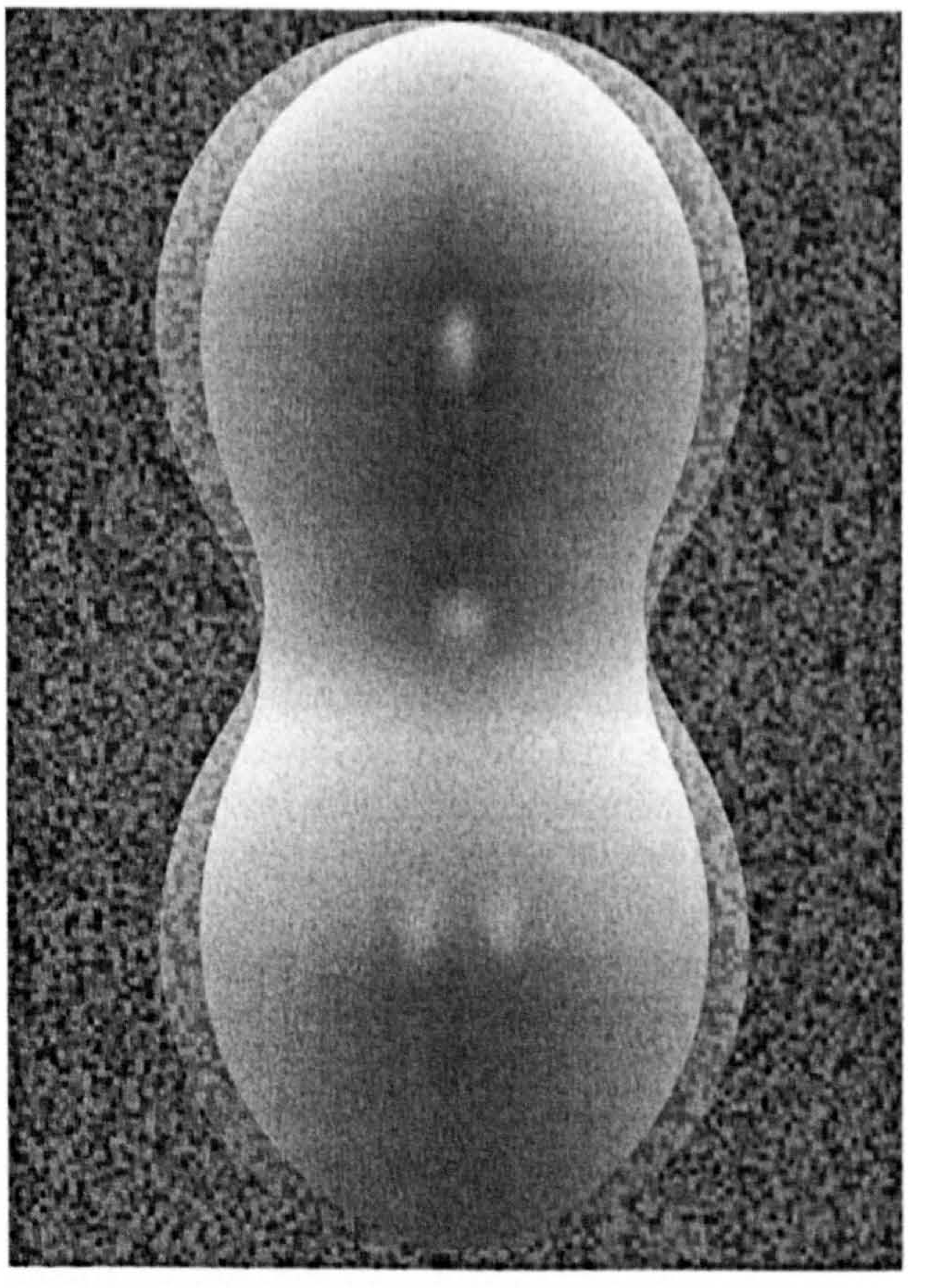


Figure 5. *Vertical or rotational misalignment of the projectors could cause a smaller waist to be seen in comparison with the hips and shoulder areas.*

by using professional-quality polarizing filters over each projector lens. These must be correctly aligned to 45 deg. (left and right) from the vertical to match the polarization angles of conventional 3-D movie spectacles. Image depolarization and cross-talk can still occur with these filters if the screen surface is not designed to maintain the polarization of the reflected image. In these experiments, image cross-talk was kept below 5% in each channel.

To test for the possibility that the slimming effect might be an artifact of projected stereo images, two Wheatstone viewers were used to present the transparencies in experiment 2. The advantage with this type of viewer (Pinsharp Viewer) is that it offers near same-size magnification, very high central resolution, zero cross-talk, and user control of the convergence for comfortable viewing. It also permits the presentation of a pair of conventionally mounted 35 mm stereo transparencies in one viewer and synoptic 2-D same-size copies in the other. When stereo pairs were shown to the participants, they could be asked to make comparisons between the 3-D and synoptic image while ensuring that the only difference between the conditions were the disparities presented.

2.3 The Virtual Stimulus

For experiment 3, a virtual "peanut-like" object was designed with the same imaging geometry as experiments 1 and 2 (see figures 11 and 12) using an architectural computer-aided design program (StrataVision 3D 4.0 from Strata, Inc.), with sophisticated rendering and lighting capabilities. The real-world image quality available with StrataVision is unlikely to generate the variable pixellation that could occur with simpler 3-D programs. It could also incorporate a random-dot background that was derived directly from stock Adobe PhotoShop files. When rendering stereo disparities using a computer-aided design package, it is important that the model is very accurately described, because small changes in topography or brightness due to aliasing can alter the stereoscopic detail within the image. The overriding design priority was that the virtual experiment could be repeated with a real object using stereo photography. It is therefore possible, should it be desired, for the virtual object and its background to be constructed and the camera/lighting simulation to be accurately reproduced.

3 The Fattening Effect of Zero-Disparity Images

A series of studies was performed to test the hypothesis that the absence of stereo depth information in 2-D images causes size and shape misperception of people and objects.

3.1 Experiment 1: Images of Female Models

3.1.1 Method. The stimuli, participants, and procedure of experiment 1 are as follows.

3.1.1.1 Stimuli. Ten female volunteers were photographed in stereo using the stereo camera described in section 2.1. The stereo photographs were taken with the models at three-quarter profile (figure 2). After being weighed and accurately measured, each model wore a dark swimsuit and was positioned in front of a flat photographic background over a floor mark. The left stereo photograph was copied to make a synoptic 2-D pair for the presentation.

3.1.1.2 Participants. Twenty-eight Liverpool University undergraduates were tested individually.

3.1.1.3 Procedure. Participants began by taking the TNO stereo acuity test (TNO, 1972) and viewing a series of projected 3-D slides to accustom them to stereo viewing. They were then shown life-size projected images of the ten models in alternating stereo and synoptic 2-D images, so that each model was never shown to the same participant in both 2-D and 3-D. Half the participants saw models 1 through 5 in stereo and models 6 through 10 in synoptic 2-D, and half saw 1 through 5 in synoptic 2-D and 6 through 10 in stereo. Trials were self-paced, and, during each presentation, participants rated the bodyweights of each model on a seven-point Likert scale: Very overweight, Overweight, Slightly overweight, Correct, Slightly underweight, Underweight, Very underweight.

3.1.2 Results. The mean perceived weight estimates of the ten models viewed either stereoscopically or synoptically are shown in figure 6. As the means and the very small standard errors indicate, there was a strong centralizing tendency in the participants' judgments, partly because the range of bodyweight in the models was not high but partly also probably because of a reluctance on the part of the participants to make negative judgments on the models. Nevertheless, a one-factor (viewing condition) ANOVA showed that the models were rated as significantly slimmer when viewed stereoscopically ($F(1,26) = 15.072, p = 0.001$).

3.1.3 Discussion. Although there was a significant slimming effect of stereoscopic presentation, it was possible that this was an indirect effect of evoking increased presence in 3-D presentations. Informal reports from several participants suggested that they sometimes felt they were in the presence of real people. Perhaps increased presence may have led the participants to give judgments that were less harsh to models that they felt were more present in the laboratory. Although this seems unlikely, particularly as most viewers were unaware that the presentation mixed 2-D and 3-D images, it was decided in experiment 2 to test this finding using an inanimate object. The generalizability of the initial finding was also tested further by using

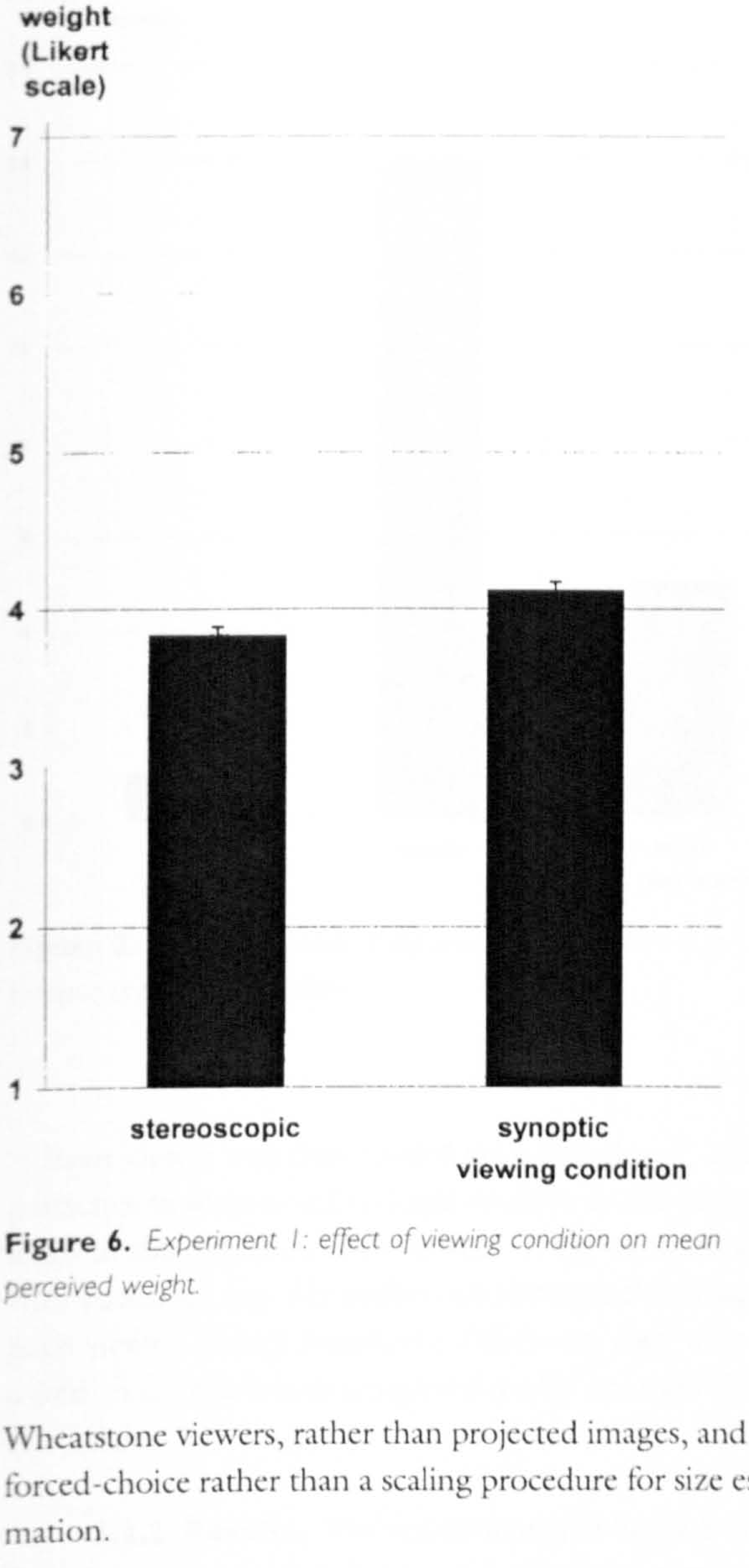


Figure 6. *Experiment 1: effect of viewing condition on mean perceived weight.*

Figure 7. *The stimulus used in experiment 2.*

Wheatstone viewers, rather than projected images, and a forced-choice rather than a scaling procedure for size estimation.

3.2 Experiment 2: Images without Human Presence

3.2.1 Method. The stimuli, participants, and procedure of experiment 2 are as follows.

3.2.1.1 Stimuli. Two large flower pots were arranged to form a waisted object (figure 7) which was then photographed using the same camera and image-capture geometry as used for the stimuli in experiment 1. The stereo transparencies were made using the method described in section 2, but the object was daylight-illuminated with the background plane imaged at infinity. The transparencies were mounted in a Wheatstone-type handheld stereo viewer. The horizontal/vertical field of view was 40 deg., and the viewer had user-variable vergence control. A second viewer held two same-size copies of one of the stereo transparencies, forming a synoptic pair.

3.2.1.2 Participants. Twenty Liverpool University undergraduate participants were tested individually.

3.2.1.3 Procedure. While viewing a series of pretest stereo images, each participant was shown how to use the two Wheatstone viewers.

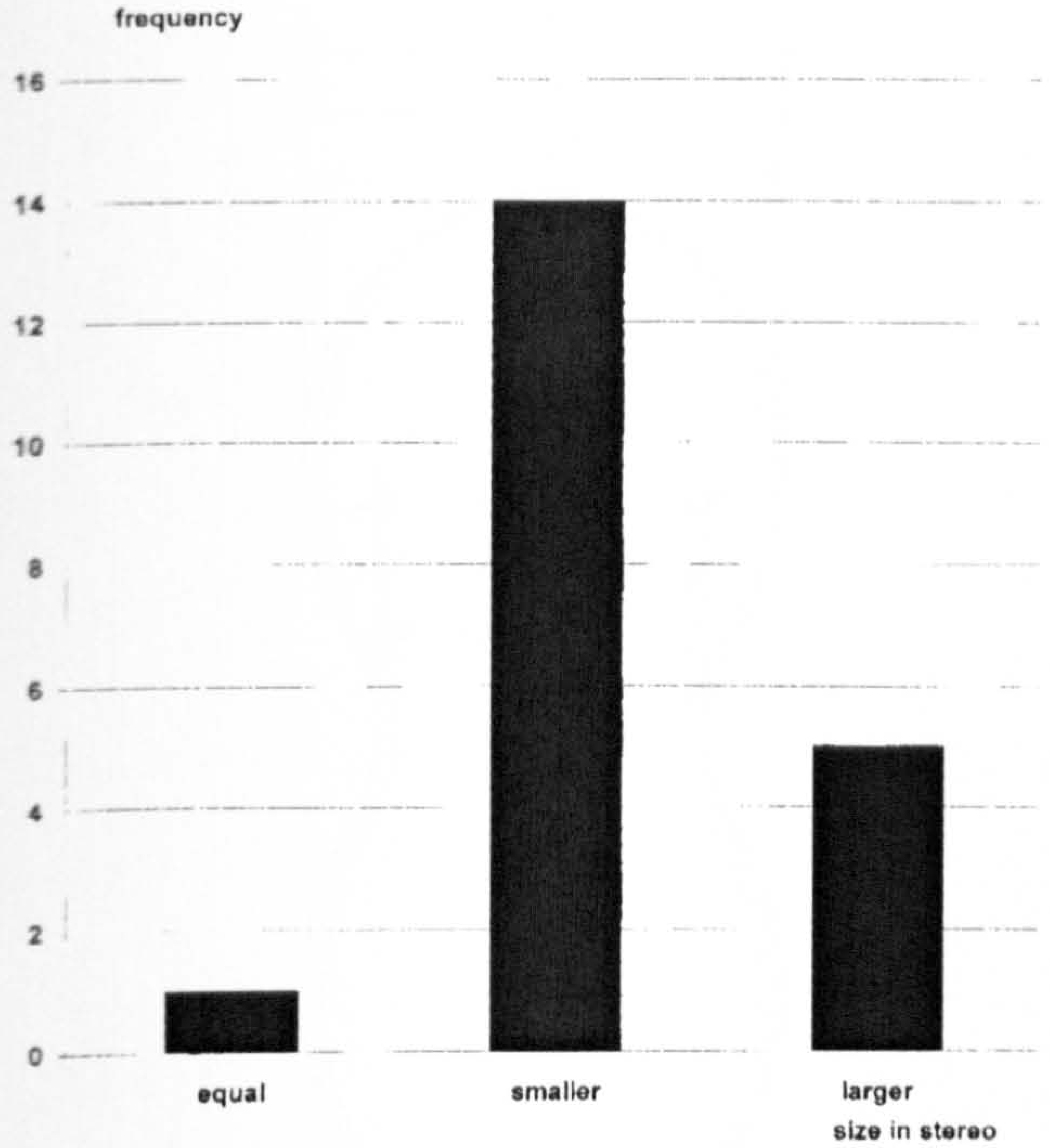


Figure 8. *Size comparisons of the stimuli in experiment 2 in synoptic and stereo conditions.*

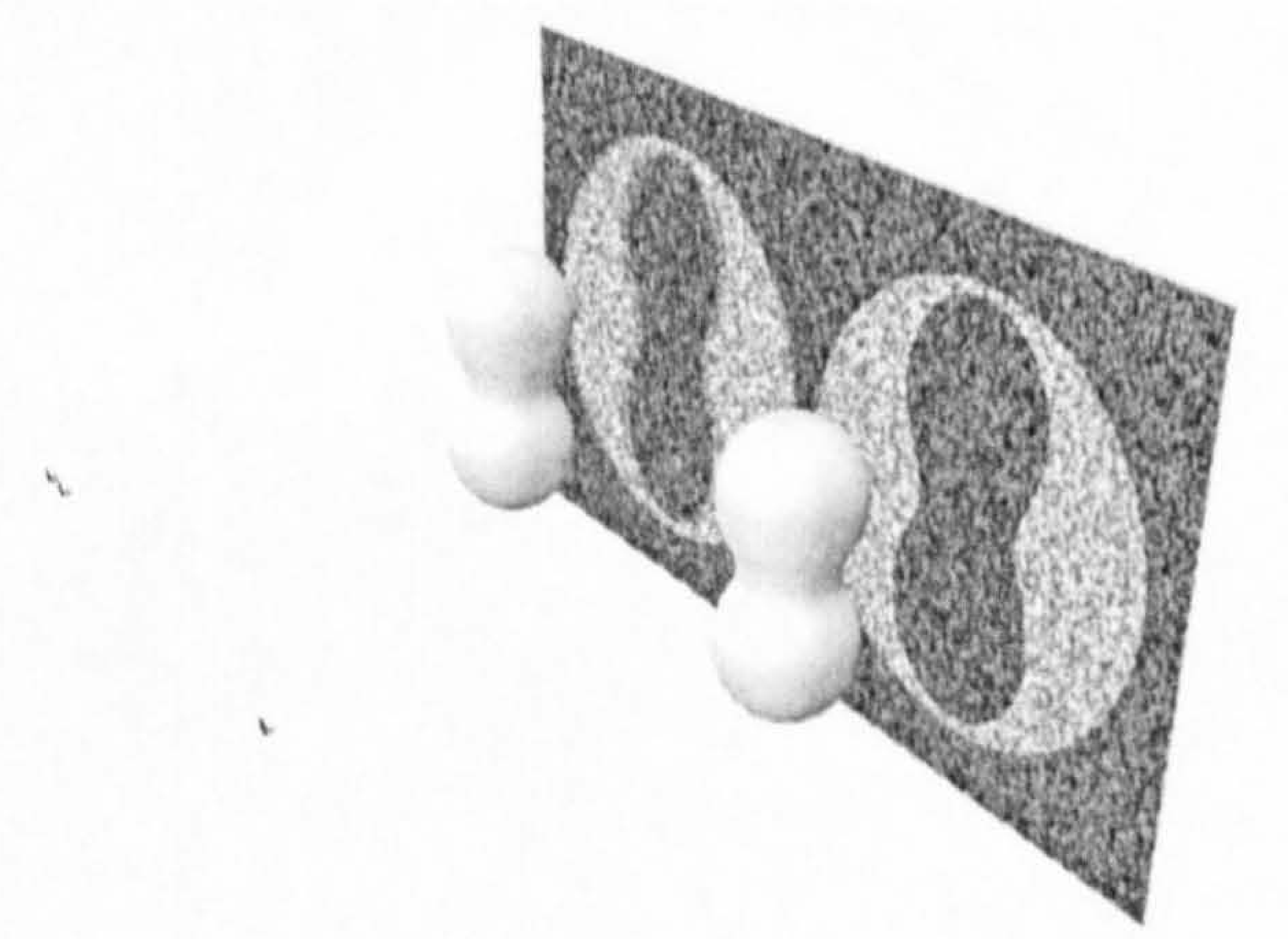


Figure 9. *The size and shape of the occluded area behind the object. The occluded area not only becomes smaller with disparity (left image), but the waist-hip ratio also changes; the wider the disparity, the lower this ratio becomes. (See also figure 10.)*

Each viewer was then loaded with the stimuli, and the participants were asked to look carefully at the dimensions of the object in both viewers. They were asked if they could see any size difference between the images in each viewer. If they reported a difference, they were asked to choose which image was wider or larger than the other.

3.2.2 Results. The results shown in figure 8 confirm the prediction that the waisted object was viewed as slimmer or smaller in the stereo presentation ($\chi^2(2, N = 20) = 13.3, p < 0.001$). Almost three times as many viewers saw the object as slimmer or smaller when viewed binocularly compared to the synoptic image.

3.3 Experiment 3: Digital Variable-Waist Images

When directly comparing the synoptic and stereo images of female models in experiment 1, it seemed that not only were the models appearing to be slimmer but also that their proportions were subtly altered. Both necks and waists appeared to be disproportionately slimmer than their associated jaw and hip widths. The flowerpot stimuli used in experiment 2 also seemed to support this view, and simple trigonometry confirmed that this was possible (figure 9 and 10). A new shape-matching experiment was designed to test whether perceived waist-hip and jaw-neck ratios could be affected by changing between 2-D and stereo image presentation. Two additional conditions were also introduced. Two different disparities in the binocular condition were used to look at the relationship between the degree of size distortion and the magnitude of the disparity. A parallel-axis stereogram was also included to allow the direct comparison of the distortions in parallel and convergent stereo. All of the participants were also tested for stereo acuity using the TNO test to establish if this was a reliable predictor of performance.

3.3.1 Method. The stimuli, participants, and procedure of experiment 3 are as follows.

3.3.1.1 Stimuli. A peanut-shaped 3-D model (see section 2.3) was designed. The widest part of the stimu-

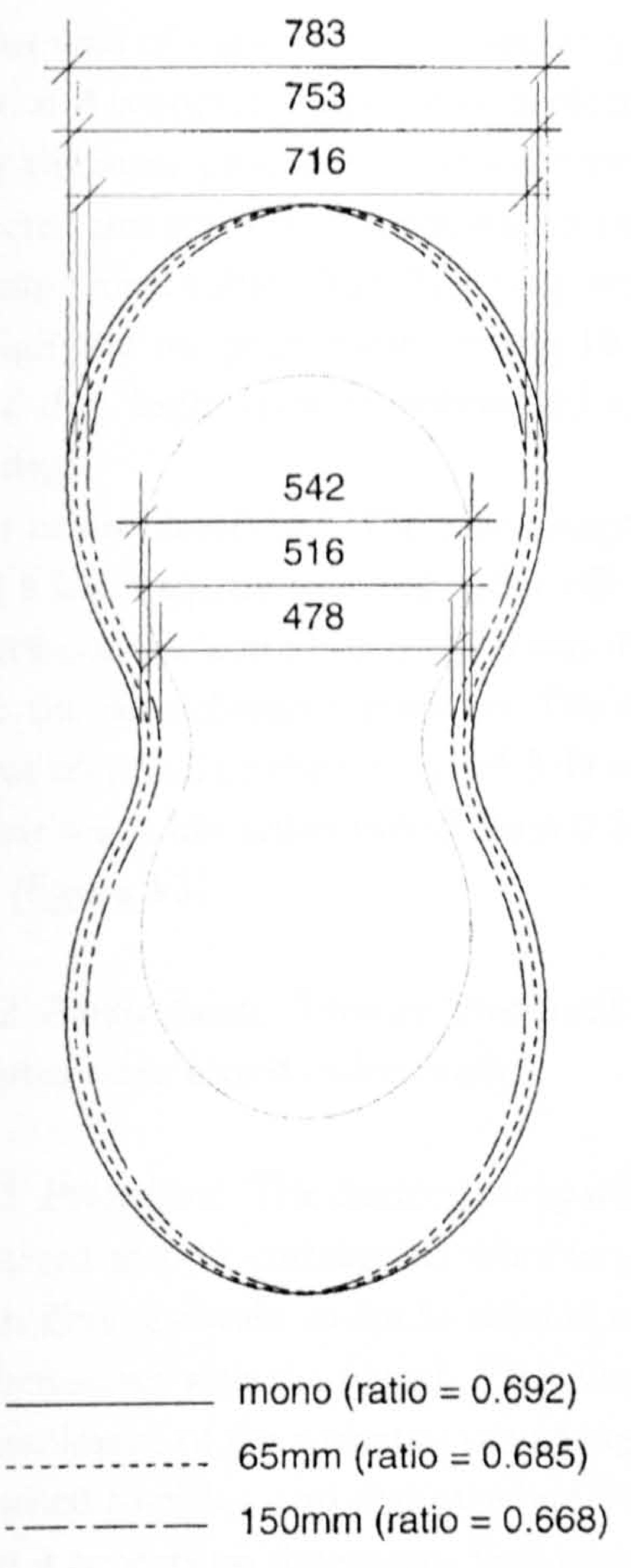


Figure 10. *Diagrammatic representation of the size and shape of the occluded area behind the object shown in figure 9 quantifying the way in which the occluded area becomes smaller with disparity and its waist-hip ratio lowers with increasing disparity. (It should be noted that the occluded area from the monocular position does not equal the 0.7 waist-hip ratio of the foreground object, because it was not imaged from a camera at infinity.)*

lus is described in these experiments as the "hips." The narrowest is the "waist." The waist circumference in figure 11 is 70% of the size of the hips, and this is described as a 0.7 waist-hip ratio. All of the stereo and synoptic images of the stimuli in experiment 3 are of this 0.7 ratio. Its surface was rendered without texture so that the only stereo information available to the viewer was from lighting-derived contour and shading and the trapezoidal distortion (perspective keystoning) of the background.

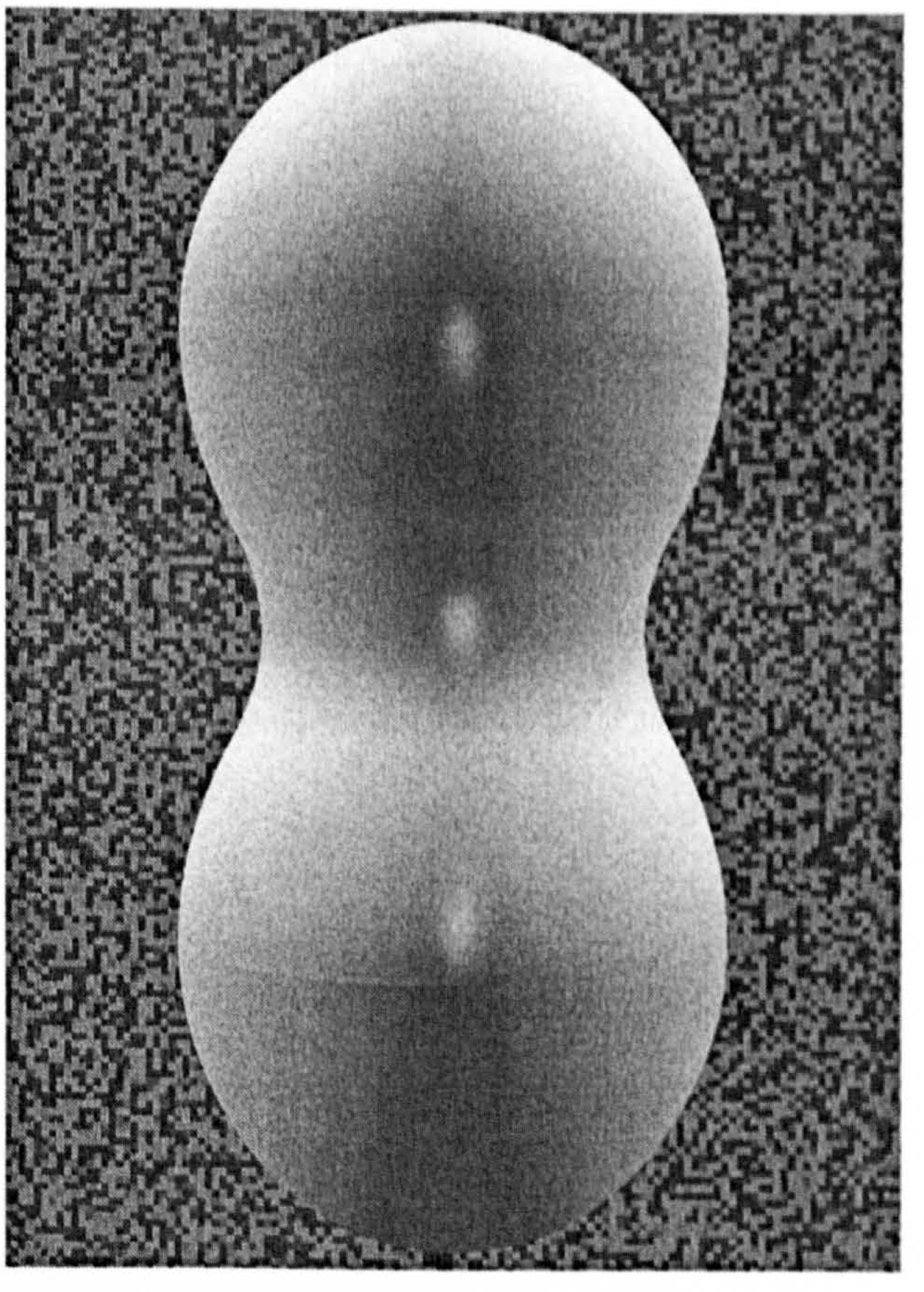


Figure 11. *An example of the peanut shape used in experiment 3 with a waist-hip ratio of 0.7. (This image is cropped for reproduction so the background is smaller than in the test stimulus.)*

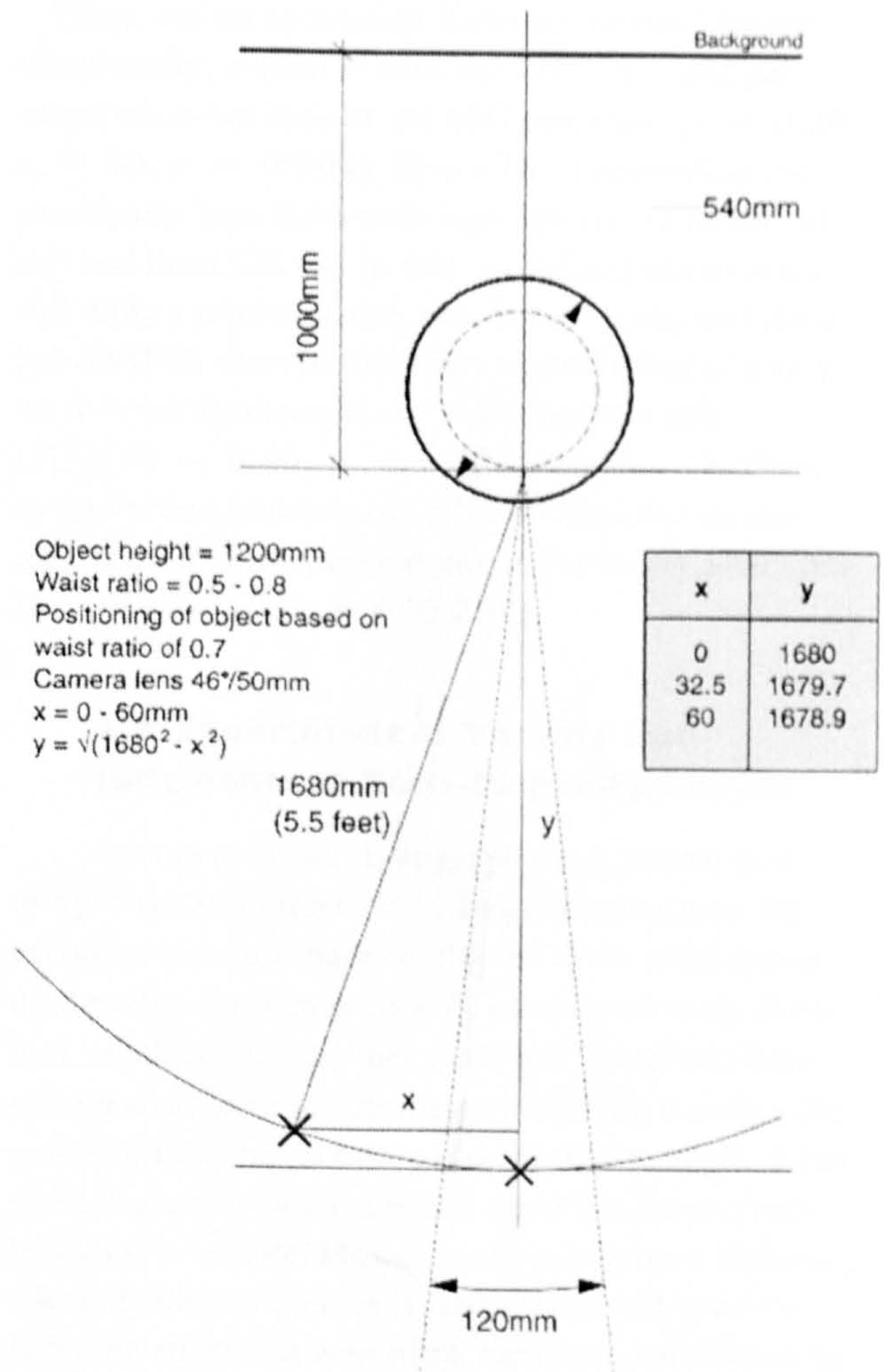
Four computer-generated stereogram pairs of the 0.7 peanut model were rendered for polarized projection to individual participants. These images were made in a series of widening disparities with 00 (synoptic, 2-D), 65P (65 mm, parallel axis), 65C (65 mm, convergent axis) and 120C (120 mm, convergent axis) interaxial equivalent separations. The peanut was constructed to approximate the "ideal" 0.7 waist-hip ratio of a healthy adult female (Singh, 1993) but with rotational symmetry in order to have the same shape from any horizontal angle. In the plan view (figure 12), the peanut and its relationship to the virtual cameras and the background are shown. These were designed to be identical to the arrangement used in experiment 1. The background was a random-dot wall of light-gray and dark-gray pixels. Stereoscopic and synoptic images were projected onto a screen using the same procedure as in experiment 1. These projected images were the equivalent of life-size, with the background subtending 31.6 deg. wide by 47.0 deg. high and the peanut subtending 18.6 deg. wide by 39.3 deg. high. Its waist subtended a visual angle of 13.0 deg.

The order of presentation of the four images was rotated round a Latin square to avoid order effects. A set of thirteen A4 comparison photographs was made of the peanut from the zero-disparity position. The image on each card was identical to the projected 3-D images except that their waist-hip ratios varied from 0.5 to 0.8 in 0.025 steps (figure 13).

3.3.1.2 Participants. Twenty Liverpool University undergraduates were tested individually.

3.3.1.3 Procedure. The thirteen comparison cards were randomized and the participants asked to place them in order from slimmest waist to fattest waist in order to familiarize themselves with the stimuli. They were then shown the first image of the sequence of varying-disparity images and asked to pick a card that matched the shape of the peanut as it appears on the screen. This was repeated with the remaining three images.

3.3.2 Results. Figure 14 shows the frequency distributions of participants' matches for the four different disparities. Figure 15 shows the overall group means for these choices. A one-factor (disparity) ANOVA found an overall effect of disparity on size judgement ($F(2,38) = 7.628$, $p = 0.002$). Post-hoc paired comparisons showed that the only significant differences were between the 0 deg. (synoptic) and the 65 deg. (stereo) ($t(19) = 3.367$, $p = 0.003$, two-tailed) and between the 0 deg. (synoptic) and the 120 deg. (stereo) $t(19) = 3.286$, $p = 0.004$, two-tailed).



x	y
0	1680
32.5	1679.7
60	1678.9

Figure 12. *Plan view of the dimensions of the peanut shape used in experiment 3, its relationship to the virtual camera positions, and the plane of the background. The virtual cameras generated views at each disparity in an arc to ensure that the magnification was constant in every image. The right-hand bold X shows the position of the camera when it was in the straight-ahead position (zero disparity). The left-hand bold X shows the position of the left-hand camera at a distance x from the straight-ahead position. The disparity this generates is defined as 2X mm.*

3.3.3 Discussion. It can be seen in figure 14 that the image-capture geometries (or disparities) used in this experiment reveal a previously unseen effect. The zero-disparity 2-D stimuli (card 9) was correctly matched to its projected equivalent (0.7 waist-hip ratio) by more than half of the participants (figure 14a). The average perceived waist-hip ratio of the group was 0.694 (figure 15). This is almost identical to the oc-

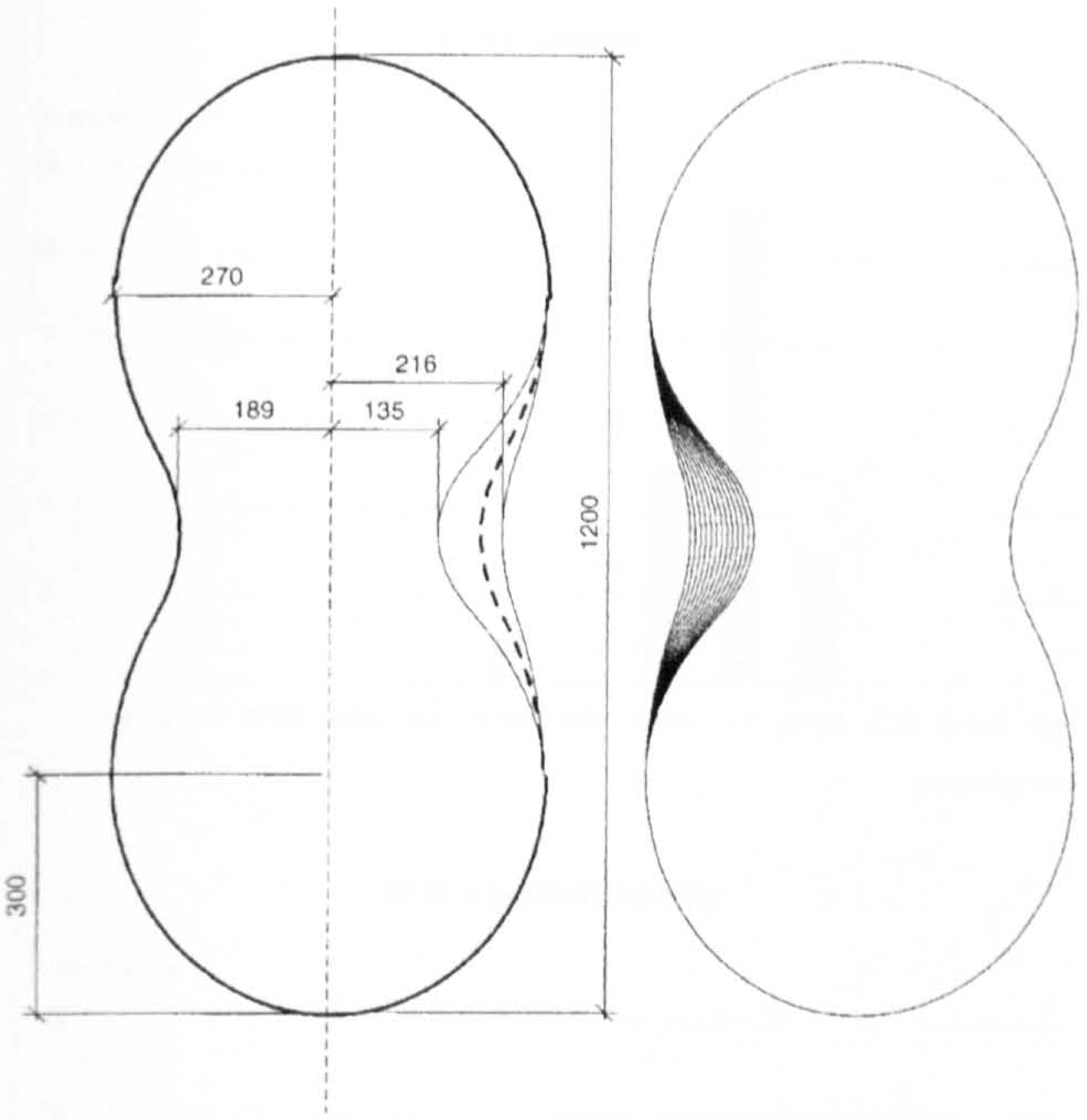


Figure 13. *The waist-hip ratios of the thirteen comparison stimuli used in experiment 3. Each stimulus was printed onto an A4 card (with a random-dot background, as in figure 11). Card 1 had the slimmest waist-hip ratio of 0.5. Each of the subsequent cards had a ratio that increased in 0.025 graduations. Card 9 (see also figure 11) was the same 0.7 ratio as the stereo and synoptic images. Card 13 was at a ratio of 0.8. The left diagram shows the largest and smallest physical dimensions of the varying waist sizes. The diagram on the right shows all of the intermediate ratios. The card images were scaled so that they were approximately the same size as the projected image when held at arms length.*

cluded area as shown in figure 10 of 0.692. However, when the viewers were shown the same shape but in stereo with 65 mm of convergence disparity (corresponding to the normal geometry of human stereo vision), a match with a significantly slimmer waist-hip ratio was selected. Conventional stereo cameras do not capture images with convergent lens axes but use parallel capture geometry. When this condition was simulated with a test image (65P), the mean perceived waist-hip ratio did not differ significantly from the synoptic condition. It can also be seen in figure 14c that there is much more variation in responses in this condition.

There was no correlation between the participants' stereo acuity, measured with the TNO test, and perceived waist-hip ratio in the 65C condition ($r = 0.29$, $n = 20$, $p = 0.904$) (figure 16). Subdividing the participants into those with high (15 sec. to 60 sec. of arc) and low (120 sec. to 480 sec. of arc) stereo acuity and using a mixed-design, two-factor (acuity and disparity) ANOVA showed that there was no effect of acuity on their performance in the size-judgment task ($F(1,18) = 0.46$, $p = 0.506$). Neither was there an interaction between the effect of disparity on size judgments and the performance in the stereo acuity test ($F(3,54) = 1.49$, $p = 0.228$).

3.4 Experiment 4: Varying Size Judgments in Zero-Disparity Images

In conventional photography, it is known that using lenses of different focal lengths can change the perceived size and shape of objects. Wide-angle lenses used in close proximity to scale models can make them look much larger than they really are. Telephoto lens compression can trick the viewer into misperceiving the spatial relationship between objects. For example, it can make the moon look oversized when it is framed with buildings or people. However, the perspective-flattening effect of telephoto lenses is rarely associated with the fattening effect that is so often mentioned in relation to photographic portraits, film, and television. Experiment 4 was designed to test the hypothesis that bodyweight appears higher in telephoto images and lower in wide-angle images. Of particular interest was the effect of different focal lengths of lens on the perceived width of the model's neck relative to the width of the jaw. Figure 17 shows how varying lens-to-subject distances can change the measured waist-hip ratio of the occluded area (as well as the expected size change) behind the peanut shape. It should be noted that quoting focal lengths in millimetres can be misleading. Lens calibrations can offer different image magnifications depending on the camera used. For instance, a 50 mm lens on a 35 mm SLR is considered to be a standard lens. On a 6 × 6 camera, it is a wide-angle lens. On a video camera, it would be a telephoto lens. For experiment 4, the inde-

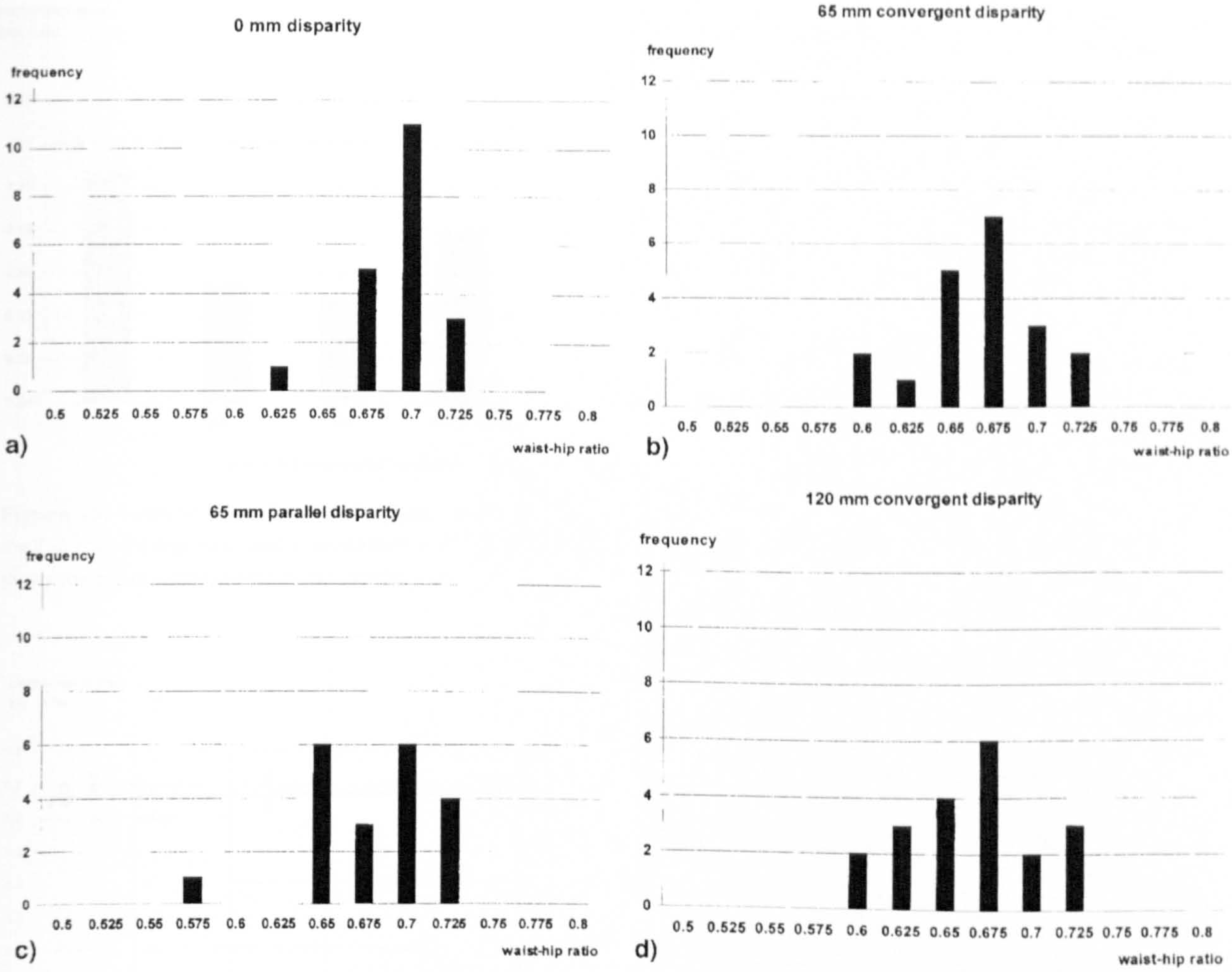


Figure 14. *Experiment 3. The matches that the participants made when shown the shape with a waist-hip ratio of 0.7: a) synoptically (0 mm disparity); b) stereoscopically with 65 mm, convergent disparity; c) stereoscopically with 65 mm, parallel disparity; d) stereoscopically with 120 mm, convergent disparity. Increasing the convergent stereo disparity to 120 mm results in a lower perceived waist-hip ratio.*

pendent variable reported is therefore camera-to-subject distance while maintaining a same-size image, because this is repeatable regardless of the camera system or lens design used.

3.4.1 Method. The stimuli, participants, and procedure of experiment 4 are as follows.

3.4.1.1 Stimuli. Two men and three women were photographed in identical poses using zoom lenses in a series of five focal lengths from wide-angle to telephoto. Using guides in the viewfinder, the lenses were zoomed very accurately for each of five camera-to-subject distances. This method made it possible to record the facial features of each model to the same magnification at the film plane from distances of 0.32 m, 0.45 m, 0.71 m, 1.32 m, and 2.70 m. Prints were made from the portraits and five sets were assembled, each containing one photograph of each model at one of the five focal lengths. Examples from the range are shown in figure 18.

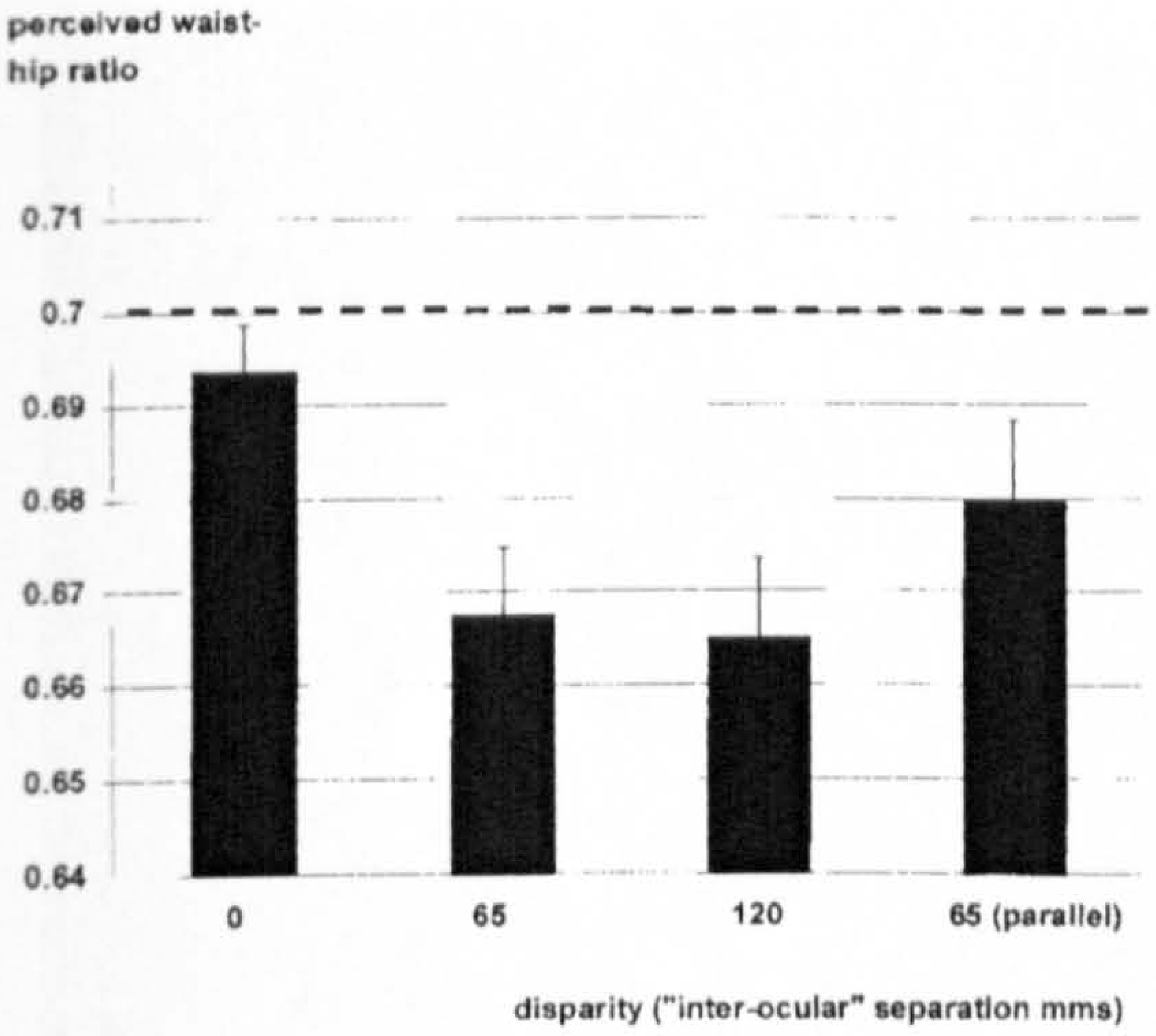


Figure 15. *Perceived waist size when the object was projected in the four different disparities used in experiment 3. The dashed line shows the actual waist-hip ratio of the stimulus.*

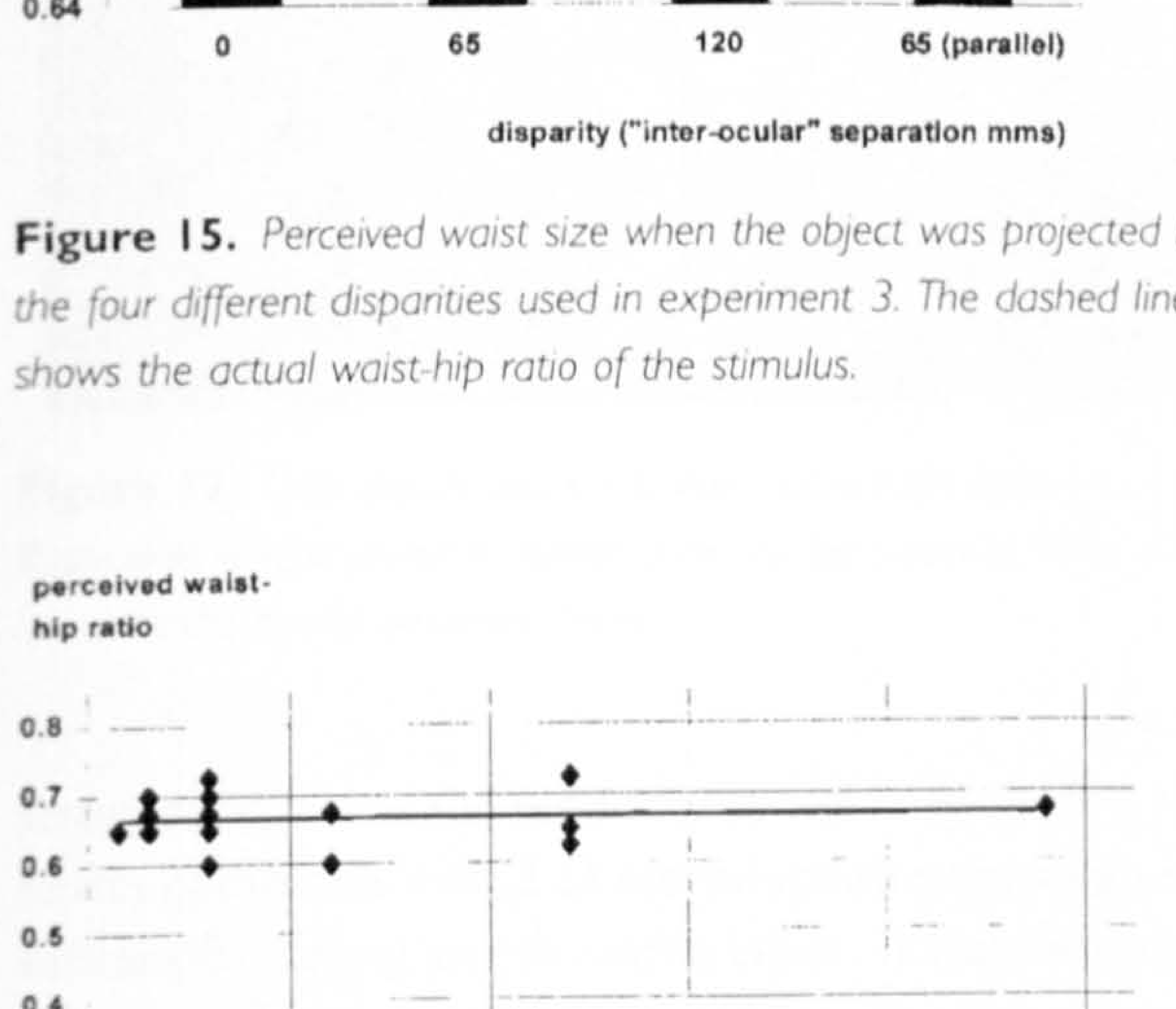


Figure 16. *The relationship between perceived waist-hip ratio in the 65 mm, convergent disparity condition, and the stereo acuity of the individual participants in experiment 3.*

3.4.1.2 Participants. Twenty Liverpool University undergraduates were tested individually.

3.4.1.3 Procedure. One set of photographs was shown to each of four groups of five participants. They were never shown the same model photographed at more than one focal length (unlike the examples in figure 18). Each participant was asked to place the five different model portraits in a rising order of apparent bodyweight using the same seven-point Likert scale as in experiment 1 (see section 3.1.1.3), and to apply a number from 1 to 7 to each image. A number greater than 4 was given to people who appeared to be overweight and numbers less than 4 to people who appeared to be underweight. The most overweight would be given a score of 7, the most underweight a score of 1.

3.4.2 Results. Figure 19 shows that, as camera-to-subject distance (and focal length) increases, a higher score was given on the Likert scale ($r = 0.824$, $N = 5$, $p < 0.05$, one-tailed). A one-factor (camera-to-subject distance) ANOVA found an overall effect of distance on size judgment ($F(4,76) = 8.858$, $p < 0.001$). Planned comparisons using two-tailed t-tests, showed that the wide-angle, close-proximity images (0.32 m) showed underweight estimations ($t(19) = 4.073$, $p = 0.001$). The standard lens image (0.71 m) showed a slight but not significant overweight estimation ($t(19) = 1.097$, $p = 0.287$). The telephoto distance images (1.32 m and 2.7 m) showed overweight estimations ($t(19) = 2.101$ and 5.101, $p = 0.049$ and < 0.001).

3.4.3 Discussion. Because of the limitations of the photographic location and lenses available, it was not possible to test if extending the range of focal lengths would show a continuing positive relationship between focal length and perceived bodyweight. It is likely, however, that the focal lengths used in this experiment cover the range in which the strongest effects could be demonstrated. Extreme wide-angle distortions at one end of the scale and proportionally smaller changes in the depth-compression effect of telephoto lenses at the other would probably act to curtail the effect.

4 General Discussion and Conclusions

These experiments support the theory that conventional imaging methods can convey misleading object information. Images of people seem to carry the

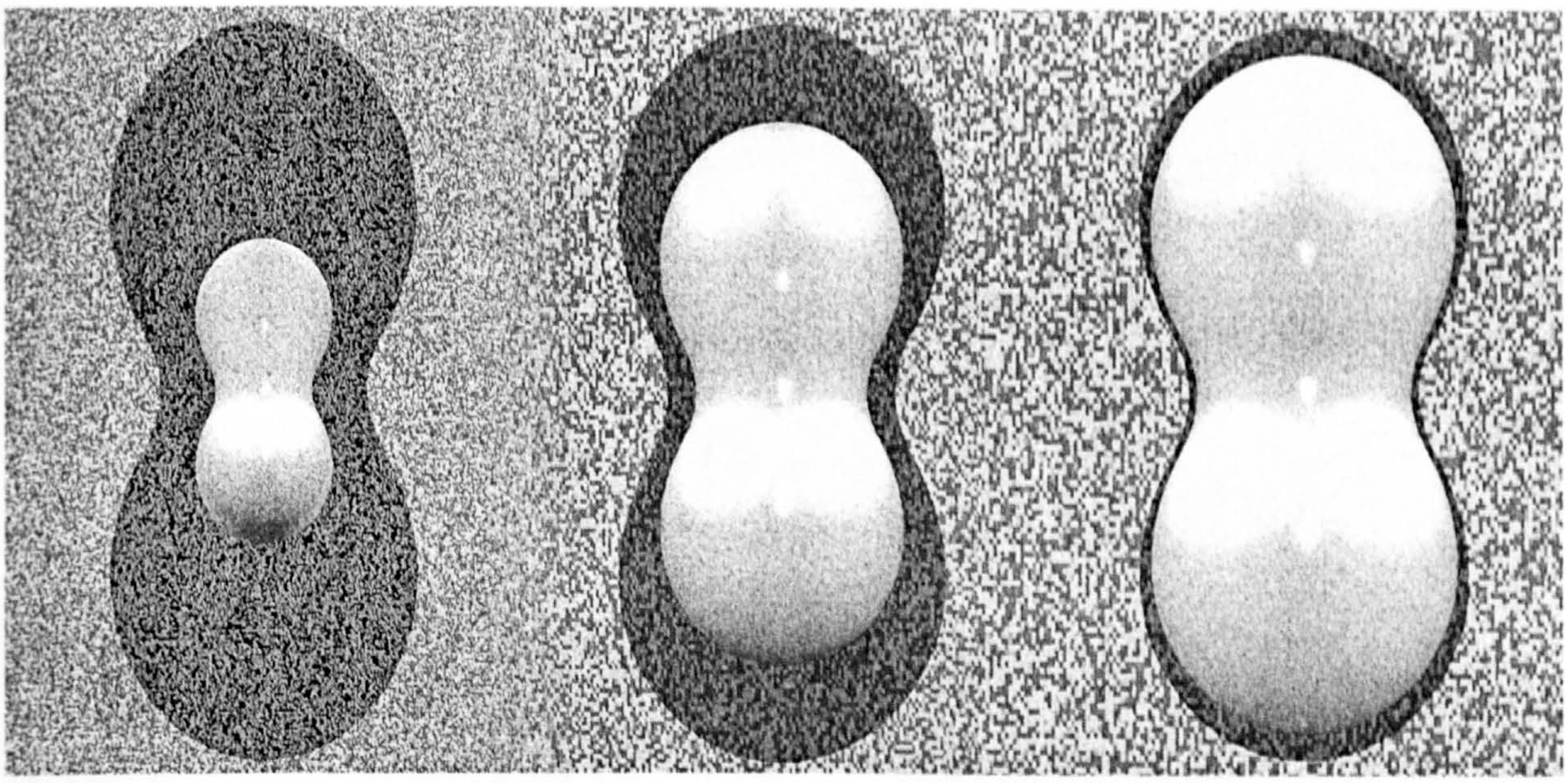


Figure 17. *Only objects viewed or illuminated from optical infinity can generate an occluded area that is the same size as the object. In this illustration, a light source is moved closer to the object in three stages (from right to left). The waist-hip ratio of the occluded area becomes lower as the source becomes closer.*

strongest effect as the tendency to use long-focal-length lenses combined with 2-D reproduction produces a significant flattening and fattening effect. It may be that we have specific mechanisms for shape recognition of the human body (Perrett, Harries, Mistlin, & Chitty, 1990) which are particularly sensitive to interference by different methods of imaging. Experiment 1 supported the theory that people look slimmer when viewed stereoscopically. Experiments 2 and 3 showed that the slimming effect of binocular disparity is seen with inanimate objects as well as human participants. Experiment 4 indicated that 2-D photography, which is usually considered to be a veridical method of record, can cause inaccurate size judgments under certain common conditions. In portraiture, it is likely that a model's directly seen jaw-neck ratio will be perceived as slimmer than in a conventional 2-D photographic image taken from the same viewpoint. The body-image distortion described here could be reduced by comparatively simple changes in 2-D imaging techniques. Some correction of the most common fattening effects can be achieved by using wide-angle lenses with carefully controlled subject proximity. However, only a well-designed stereoscopic or volumetric display can properly solve all of these problems.

We have also demonstrated that orthostereoscopic images can affect object ratio judgments in shape perception. In experiment 3, the circumference of the waist of the peanut shape was seen as 5.4% slimmer (when averaged across all participants) in the 65C condition than the waist of the synoptically viewed object. It should be noted, however, that the participant's view of the orthostereoscopic display geometry used in experiments 1 and 3 was not as well corrected as it could have been. Firstly, although the stereo transparency pairs were converged at the object's waist, it was not possible to actively adjust the vergence angles so that all of the other gaze points on the stimuli were seen as having zero disparity as they were viewed. In this respect, the display could not perfectly simulate direct viewing of a real object as a small amount of nonveridical vertical disparity was fused as the observers moved their gaze away from the center of interest. However, it was at the waist (the area of zero disparity) that the object ap-

Figure 18. *Three images (from a sequence of five) of two of the five photographic models used in experiment 4. The images on the left are extreme wide-angle photographs with a camera-to-subject distance of 0.32 m. The central images use a standard lens at 0.71 m. The images on the right were taken with a telephoto lens and a camera-to-subject distance of 2.7 m.*

peared to change shape. The background was perceived as flat throughout, even though the vertical disparity increased towards the image periphery. Incorrect vertical disparities generate pincushion (concave) or barrel (convex) distortion that would affect the perceived flatness of the background plane. As the background in

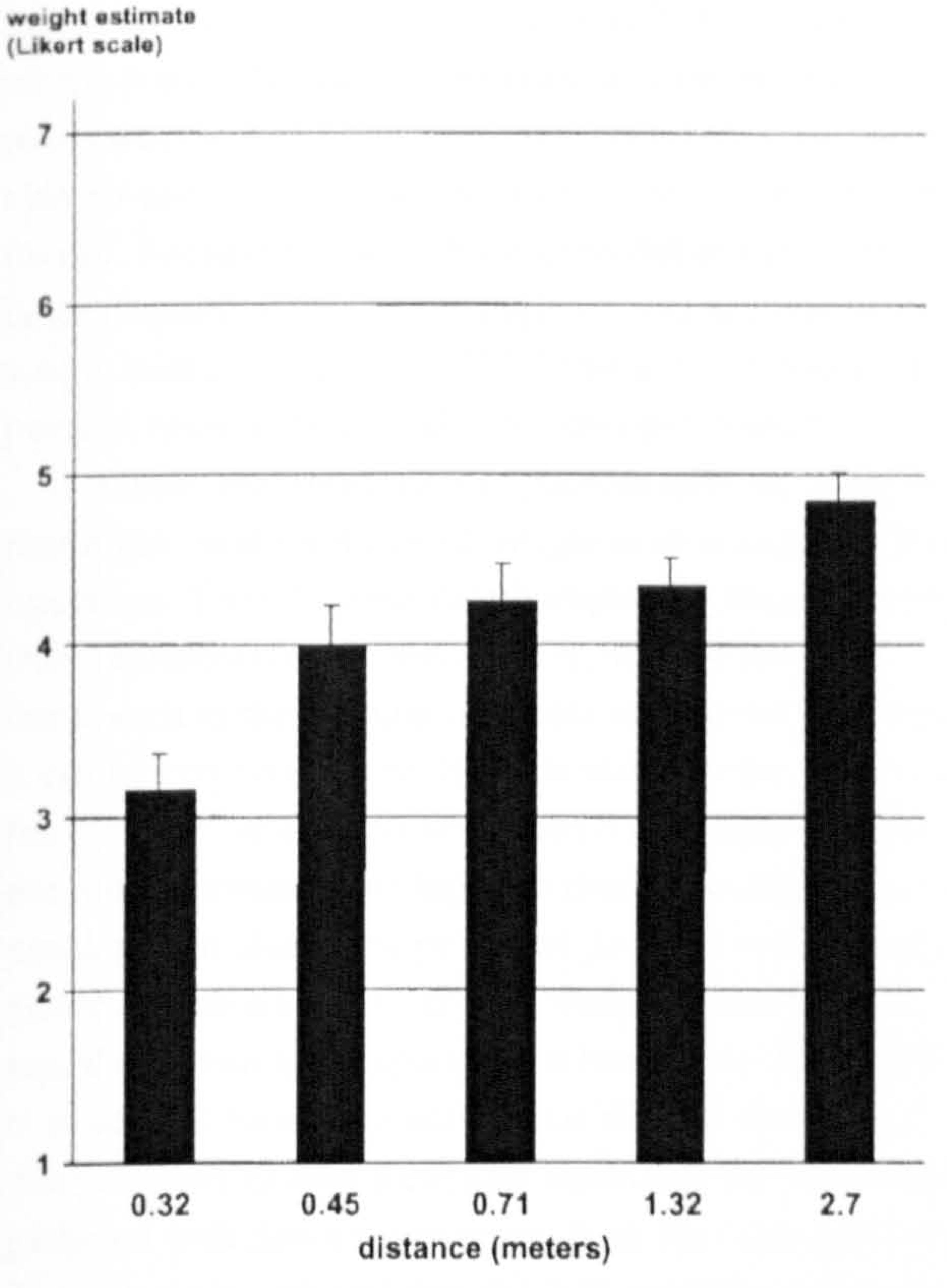


Figure 19. *The mean perceived bodyweight for the five different camera-to-subject distances (in meters), and therefore five different lens focal lengths, used to photograph the models in experiment 4.*

experiment 3 was perceived as flat, it can be inferred that the nonveridical vertical disparities did not generate obvious image artifacts. Secondly, any projected/reflected image system is likely to be compromised by the fact that the ideal viewing position (Koenderink, 1998) will occlude the projectors' optical paths. In these experiments, this was partially addressed by placing the viewing position between, but slightly below, each projector lens. The Wheatstone viewer used in experiment 2 resolves the occluded projection problem (and provides high-brightness images with zero cross-talk) but introduces others. The simple optics in this viewer are likely to induce slight curvature of field and resolution fall-off towards the edge of the image. A back-projected stereo display could, in theory, solve these problems and give a very high-brightness image. However, back projection tends to depolarize light, and, as yet, the materials required to manufacture a low cross-talk screen are currently not available. Despite these limitations, the results reported indicate that the orthostereoscopic technique used in these experiments appears to offer some advantages in veridical perception over 2-D representations of the same scene. Two-dimensional compatibility is another useful feature demonstrated by the orthostereoscopic display used in these experiments. Aligning the convergence to the point of zero disparity allows a viewer to see a single image at the center of interest in a scene without the need for polarizing glasses. This is especially true of scenes captured with low disparities.

We had expected, based on previous experience of stereoscopic displays, that some participants or experimenters would experience a degree of viewing discomfort during our experiments. However, in debriefing, no participant reported viewing discomfort in any of the experiments reported here and no experimenter experienced viewing discomfort despite very long exposure to the images. We therefore speculate that the polarized orthostereoscopic image could probably be viewed continuously for extended periods. Orthostereoscopic imaging may allow the muscles of the eyes to converge each optical axis in a natural and unstrained way. This is difficult with conventional stereography where the image separations at the screen plane require the eyes to "force fuse" two images, as if an object is at a closer position than would be the case with direct vision. Also, it can be seen in the "peanut experiment" (experiment 3, section 3.3) that shape perception may be more difficult in 65 mm parallel stereo image and causes more variation in shape matching than was found with the convergent orthostereoscopic images (figures 14b and c).

The analysis of the TNO stereo-acuity data also supports the view that the convergent images were easier to fuse for all the participants than were conventional parallel stereo images. The TNO stereo-acuity test uses a parallel stereo image-capture technique for its random-dot anaglyph plates. These anaglyph disparities are rendered to indicate the limit of a subject's ability to fuse red-green "double images." We had predicted that those participants who had above-average measured stereo acuity would perform consistently better in the size-matching task than would those with below-average stereo acuity. No such correlation was found. Participants who scored poorly on the TNO stereo test were able to easily fuse the stereo stimuli used in our experiments. Because the stimuli we used did not contain large disparities, this result suggests that the stereoscopic stimuli used in the TNO test differ in some important respects from orthostereoscopic images.

It is likely that most users of photography are unaware that it can produce distorted images in its normal modes of operation. Two-dimensional photography purports to be a truly representational medium. Yet, in common conditions, such as the imaging of people and close-up objects, it can be very misleading. It is reasonable to speculate that the "peanut" stimuli in experiment 3 correlates not only to the human female waist-hip ratio that it was designed to simulate, but also to the perceived jaw-neck ratio of both genders. This is because its waist design is similar to the way the human neck separates the head from the shoulders in males and females. It seems clear that a 2-D image of this geometry cannot accurately reproduce the information gathered with direct stereo vision from the same position. Thus, it can be inferred that the 2-D condition is almost always likely to distort when compared with an otherwise identical stereoscopic image. As parallel stereoscopic imaging seems to convey object information that causes more variation in the size-matching task, it appears that only an orthostereoscopic image can convey truly lifelike information (and therefore the presence) of objects and people.

Acknowledgments

The stimuli in Experiment 3 and all of the technical illustrations were designed and rendered by Philip Berridge of Liverpool University Department of Architecture and Building Engineering.

References

Banton, T., & Levi, D. M. (1991). Binocular summation in vernier acuity. *Journal of the Optical Society of America A, 8*(4), 376–380.

Blake, R., Sloane, M., & Fox, R. (1981). Further developments in binocular summation. *Perception & Psychophysics, 30*(3), 266–276.

Freeman, J., Avons, S. E., Meddis, R., & Pearson, D. E. (2000). Behavioral realism: Using postural responses to estimate presence. *Presence: Teleoperators and Virtual Environments, 9*(2), 149–164.

Freeman, J., Avons, S. E., Pearson, D., & IJsselsteijn, W. (1999). Effects of sensory information and prior experience on direct subjective ratings of presence. *Presence: Teleoperators and Virtual Environments, 8*(1), 1–13.

Gunby, L. (2000, September 14). Notes & queries. *The Guardian, G2*, p. 16.

Hendrix, C., & Barfield, W. (1996). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators and Virtual Environments, 5*(3), 274–289.

Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis.* Oxford: Oxford University Press.

IJsselsteijn, W., de Ridder, H., Hamberg, R., Bouwhuis, D., & Freeman, J. (1998). Perceived depth and the feeling of presence in 3DTV. *Displays, 18*(4), 207–214.

Jones, R. K., & Lee, D. N. (1981). Why two eyes are better than one. *Journal of Experimental Psychology: Human Perception & Performance, 7*(1), 30–40.

Julesz, B. (1971). *Foundations of cyclopean perception.* Chicago: University of Chicago Press.

Kelly, L. (1998, September 20). Q and A. *Sunday Mirror Personal Magazine*, p. 3.

Koenderink, J. J. (1998). Pictorial relief. *Philosophical Transactions of The Royal Society of London, 356*(1740), 1071–1086.

Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (1994). On so-called paradoxical monocular stereoscopy. *Perception, 23*(5), 583–594.

Langford, M. (1989). *Advanced photography.* London: Focal Press.

Perrett, D., Harries, M., Mistlin, A. J., & Chitty, A. J. (1990). Three stages in the classification of body movements by visual neurons. In H. Barlow, C. Blakemore, & M. Weston-Smith (Eds.), *Images and understanding* (pp. 94–107). Cambridge, U.K.: Cambridge University Press.

Prinzmetal, W., & Gettleman, L. (1993). Vertical-horizontal illusion: One eye is better than two. *Perception & Psychophysics, 53*(1), 81–88.

Singh, D. (1993). Bodyshape and womens attractiveness: The critical role of waist-hip ratio. *Human Nature, 1*(3), 297–321.

Spottiswoode, R., Spottiswoode, N. L., & Smith, C. (1952, October). Basic principles of the three-dimensional film. *Journal of the Society of Motion Picture & Television Engineers, 59,* 249–285.

TNO Test for Stereoscopic Vision (10th ed.) (1972). Veenendaal: Laméris Ootech B.V.

Warner, M. (1995, Summer). Stealing souls and catching shadows. *tate: the art magazine, 6,* 40–46.

Yamanoue, H. (1997, April). The relationship between size distortion and shooting conditions for stereoscopic images. *Journal of the Society of Motion Picture and Television Engineers, 106,* 225–232.

Yamanoue, H., Okui, M., & Yuyama, I. (2000). A study on the relationship between shooting conditions and cardboard effect of stereoscopic images. *IEEE Transactions on Circuits and Systems for Video Technology, 10*(3), 411–416.

APPENDIX B

Virtually There? A Vision on Presence Research

Wijnand IJsselsteijn,

(Eindhoven University of Technology, The Netherlands)

Bernard Harper

(Liverpool University, UK)

and the

Presence Research Working Group

PRESENCE - IST 2000-31014
EC Public Deliverable (D2)
December 2001
(Updated May 2002)

Contents

A Scenario from the Future...

It is 7pm on Wednesday May 28th 2031. You've been looking forward to this evening for the last two months, ever since you finally got around to upgrading your home Immerse-U-There© Technologies communication system as a birthday treat at the start of March.

You've noticed some real benefits since your system upgrade. The early days of teleworking using a home PC and a modem seem so primitive now, just 20 years after it became the modus operandi for the majority of the world's working population. Gone are the days of hoping that your voice communicated the right degree of warmth down the ether to your colleagues (the rudimentary video communications of the 00's didn't allow you to feel like you were really having a face to face meeting with your clients and colleagues).

Today you had a group meeting on the global communication strategy of the natural drinks company you work for. At the meeting was Jaap, the company's head communications strategist, joining direct from his small village just outside Eindhoven. Shigoro, head of marketing for Asia, has just moved to a self-sustaining village just outside Tokyo but that didn't stop him being at the meeting. Judy and Lorraine (heads of marketing communications in North and South America respectively) were in the middle of a trek through the Borneo jungle but they were able to join in too using their portable Immerse-U-There© systems. And finally, Maxine was there. Maxine is your old-timer boss, now in her late 90s, who was resolutely against 'new technologies' changing her work practices until she realised that her new Immerse-U-There© system could keep notes on everything that everyone said in the meeting, generate automatic contact reports and enable her to more or less take up residence at her local golf course, just north of Edinburgh. Constructive meeting all round. Got a bit heated when the ad agency tried to pretend their Immerse-U-There© system was broken as an excuse for joining the meeting late. You've heard that excuse from the agency at least 12 times this year - and you know its been a lie each time. Anyway, the agency came up with the goods just in time. The meeting went on for three hours - lucky really because that is the safe continuous use limit set down by the Global Standards Commission. At 2pm (World Standard Time) everyone tuned out after a hard day's work - your children still don't believe you when you tell them that only 10 to 15 years ago people were working 50 plus hours a week, and that there were different time zones around the world!

The reason you're excited though is that in about half an hour you'll be in Rio de Janeiro for the 2031 World Club Championship Final. Tickets for seats in the ground sold out over ten years ago, but with your new Immerse-U-There© system you literally feel as though you are there. You've got the best seat in the stadium, the smells of Brazil, you can feel the atmosphere (its hot and humid there tonight), you can hear every little whistle or quip that you'd hear if you were really there (and even be buzzed - but not bitten - by the local mosquitoes). And best of all, you can sit and watch with your remote friends (Jim in Hong Kong, Toby in St. Lucia, Patricia in Dubai, and Lucy in Moscow). Your kids might join you for parts of the game, but they spend most of their free time these days watching earth from the Global Space Station (through their personal Immerse-U-There© systems). Your ticket for the game was quite expensive, but this is one of the biggest sports events of the year. And you have to hand it to the system developers, they knew what they were doing. You really do feel like you're in Rio, with your friends. And it really does feel worth it. And the best thing about it all? You've hated flying all your life, and as a result of the Immerse-U-There© systems becoming the world communication standard (in 2022) you haven't had to fly anywhere for seven years. You and your partner still take holidays abroad, but there isn't the pernicious rush to get it over and done with in a fortnight, so you tend to travel by sail boat - after all you can stay in touch and be anywhere at the flick of a switch...

1. Introduction: What is Presence?

"The biggest challenge to developing telepresence is achieving that sense of "being there." Can telepresence be a true substitute for the real thing? Will we be able to couple our artificial devices naturally and comfortably to work together with the sensory mechanisms of human organisms?"
- Marvin Minsky, 1980

"What do people do at work? They go to meetings. How do we deal with meetings? What is it about sitting face to face that we need to capture? We need software that makes it possible to hold a meeting with distributed participants -- a meeting with interactivity and feeling, such that, in the future, people will prefer being telepresent"
- Bill Gates, 1999

With the rapid developments in the area of immersive, multisensory displays and human-machine interfaces, as well as the increased availability of transmission bandwidth, computing power and digital resources, we are able to create and experience reproductions and simulations of reality with an unprecedented sensory quality. However advanced such media systems and services may become from a technological point of view, the social acceptance and uptake (and consequent commercial success) will depend to a large extent on users' perceptions, experiences and responses towards them. Indeed a number of technological advances have become instructive failures because of their inability to connect to real user needs, or to provide true added value in terms of user experiences. A user experience that is of particular relevance in this regard is that of *presence*. Several studies have shown that users report a compelling perception of 'being there' when interacting with virtual environments, advanced broadcast and cinematic displays, teleoperation systems, and advanced telecommunication applications.

Presence is the experience of projecting one's mind through media to other places, people and designed environments. Appropriate presence technologies combine to create an *illusion of non-mediation* (Lombard & Ditton, 1997) - the closest possible approximation to "being there" without physical travel. The experience of 'being there' in a mediated environment is a subjective experience that has proven to be relevant for the design and evaluation of various interactive and non-interactive media systems. In particular through immersive, interactive and perceptually realistic media, user experiences and responses are provoked that are similar to those in nonmediated environments (see e.g. Lombard, 1995; Freeman, Avons, Meddis, Pearson & IJsselsteijn, 2000; IJsselsteijn, de Ridder, Freeman, Avons & Bouwhuis, 2001).

There is consensus that the experience of presence is a complex, multidimensional perception, formed through an interplay of raw (multi)sensory data, perceptual-motor activity and various cognitive and emotional processes – an experience in which attentional factors play a crucial role as well (Draper, Kaber & Usher, 1998; Witmer & Singer, 1998). Figure 1 depicts a diagram of media experience in general, and presence in particular.

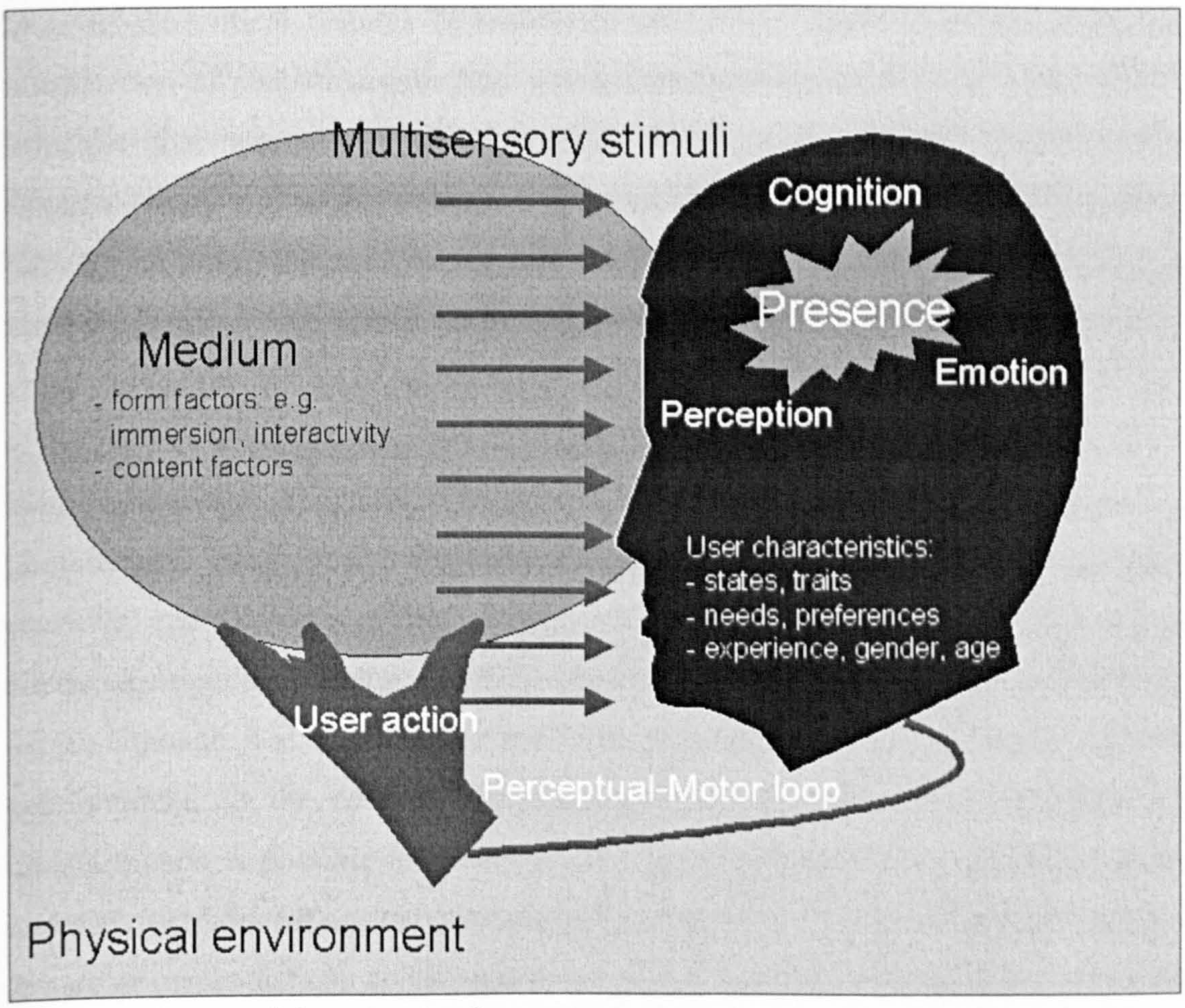


Figure 1. A general framework of presence

Factor analytic studies are starting to shed light on the multidimensional structure of presence. In particular, studies by Schubert, Friedman, and Regenbrecht (1999) and Lessiter, Freeman, Keogh, and Davidoff (2000) reveal very similar factor structures. Schubert et al. (1999) arrived at a 3-factor solution for the presence construct, which they termed 'spatial presence', 'involvement', and 'realness'. Similarly, Lessiter et al. (2000) reported a 4-factor solution for presence, with three factors almost identical to the ones identified by Schubert et al.: 'physical space', 'engagement', and 'naturalness', and a fourth attenuating factor they termed 'negative effects'.

As we move towards increasingly realistic media, each new development in the history of visual media can be viewed as a gradual buildup of perceptual cues that simulate natural perception and enhance the experience of presence (Biocca & Delaney, 1995). Early perspective paintings, dating back to the middle ages, only included static monocular depth information which violates most of the critical features of real-world perception. The end of the 18th century saw the introduction of panorama paintings, which stimulated large portions of the visual periphery, a principle that was also applied to great effect in the cinema of the 1950s (Cinerama, CinemaScope), and in more recent large film formats (e.g. IMAX), projection-based VR (e.g. CAVE), and head-mounted displays. The stereoscope of the 19th century allowed each eye to view the same scene from a slightly different perspective (i.e. stereoscopically), contributing greatly to the perception of egocentric distance and exocentric depth within an image.

With the introduction of cinema in the late 19th century, motion has been added to high-resolution photorealistic imagery as a fundamental perceptual cue. The visual system is highly motion-sensitive, and the onset of motion cannot be ignored - it demands attention and automatically elicits an orienting response. Certain camera movements provide motion parallax as a cue to depth, although it is important to note that observer movement does not transform the image appropriately. In the case of head-mounted virtual environments this viewpoint-dependent transformation is possible in real-time, although with the current state of technology real-time interactivity trades off against photorealism. Importantly, it is not clear at present how much each feature or perceptual cue contributes to the perceived realism of media, or to eliciting a sense of presence for the participant. This is an empirical question which will need to be addressed through further research.

As we are approaching the stage at which it becomes possible to develop media that can simulate the majority of critical features of natural perception (including real-time interactivity appropriately based on observer movement), the media experience is transformed from looking at pictures, to experiencing a space, to eventually visiting it as a place and being able to act within it. Despite the rapid developments it is a considerable challenge to design a realistic multisensory media environment with which we can interact intuitively and where participants can not only look, but also hear, touch, smell, taste, walk, run, pickup objects, etc.

The design goal of achieving optimal presence and making auditory, visual and haptic communications believable requires an interdisciplinary approach, integrating knowledge and ideas from disciplines such as neuroscience, multisensory perception, cognition, artificial intelligence, multimedia development, video compression or telecom engineering. In order to build future systems which can efficiently transmit remote presence, it will be necessary to incorporate and integrate ongoing insights from these fields into next-generation research for advanced, wideband multisensory services and novel telecommunications architectures. Such research would include enhancements that could amplify human perception, as well as ameliorate mediating distortions. It will allow for vastly improved mediated experiences that would bring many benefits in a number of surprisingly different ways.

2. Why study presence ?

Historically, research interest in presence has mainly been motivated by work in three related domains: teleoperation, simulation and telecommunication. Teleoperation has been driven to a large extent by the *ideal* of telepresence, i.e. sensing sufficient information about the teleoperator device and remote task environment, and communicating this to the human operator in a sufficiently natural way, that the operator feels physically present at the remote site (also known as distal attribution).

Advanced simulation systems, such as those used for flight and combat training, have long been known to engender a sense of presence in their participants, and although the relation of presence to task performance within the simulated setting has been a matter of some debate, there is little doubt that training transfer to real-world settings benefits from a realistic simulation. More recently, computing power and rendering possibilities have dramatically increased and have made it possible to create virtual environments of ever increasing realism, at a fraction of the cost of professional simulation systems. Virtual environments can be made so convincing that they can lead to a feeling of presence in the created environment (see e.g Slater & Wilbur, 1997).

From a telecommunication perspective, social presence, or the feeling of being together, has received considerable attention for several decades already (see e.g. Short, Williams & Christie, 1976). The emergence and proliferation of email, mobile communication devices, internet chatrooms, shared virtual environments, advanced tele-conferencing platforms and other telecommunication systems underlines the importance of investigating the basic human need of communication from a multidisciplinary perspective that integrates media design and engineering, multisensory perception, and social psychology. Add to this the increasingly social nature of interfaces and the increase in mediated communications with non-human entities (avatars, embodied agents), it becomes abundantly clear that we need to develop a deeper understanding, both in theory and in practice, of how people interact with each other and virtual others through communication media. The experience of social presence within different contexts and through different applications thus becomes a concept of central importance.

A common theme linking the above examples is that presence research offers the possibility to engineer a better user experience, to optimise the effectivity, efficiency and pleasurability of the different applications. Moreover, the user-centred perspective of presence research will diminish the likelihood of expensive technology failures, such as the introduction of AT&T's PicturePhone in the 1960s and 1970s (Noll, 1992) or .....(see Loretta's doc)

From an application viewpoint, presence research will spur the development of numerous tele-applications in home and professional environments. Such tele-applications have the potential to support a great number of activities, including communication, work, education, shopping, banking, health monitoring, and entertainment. Inasmuch as these tele-applications will partially make redundant the necessity for travelling physically, they will contribute to a more sustainable use of limited resources and will put less strain on the environment in terms of pollution (see figure 2). In addition, they will enable individuals that are physically unable to travel to actively participate in activities previously beyond their reach.

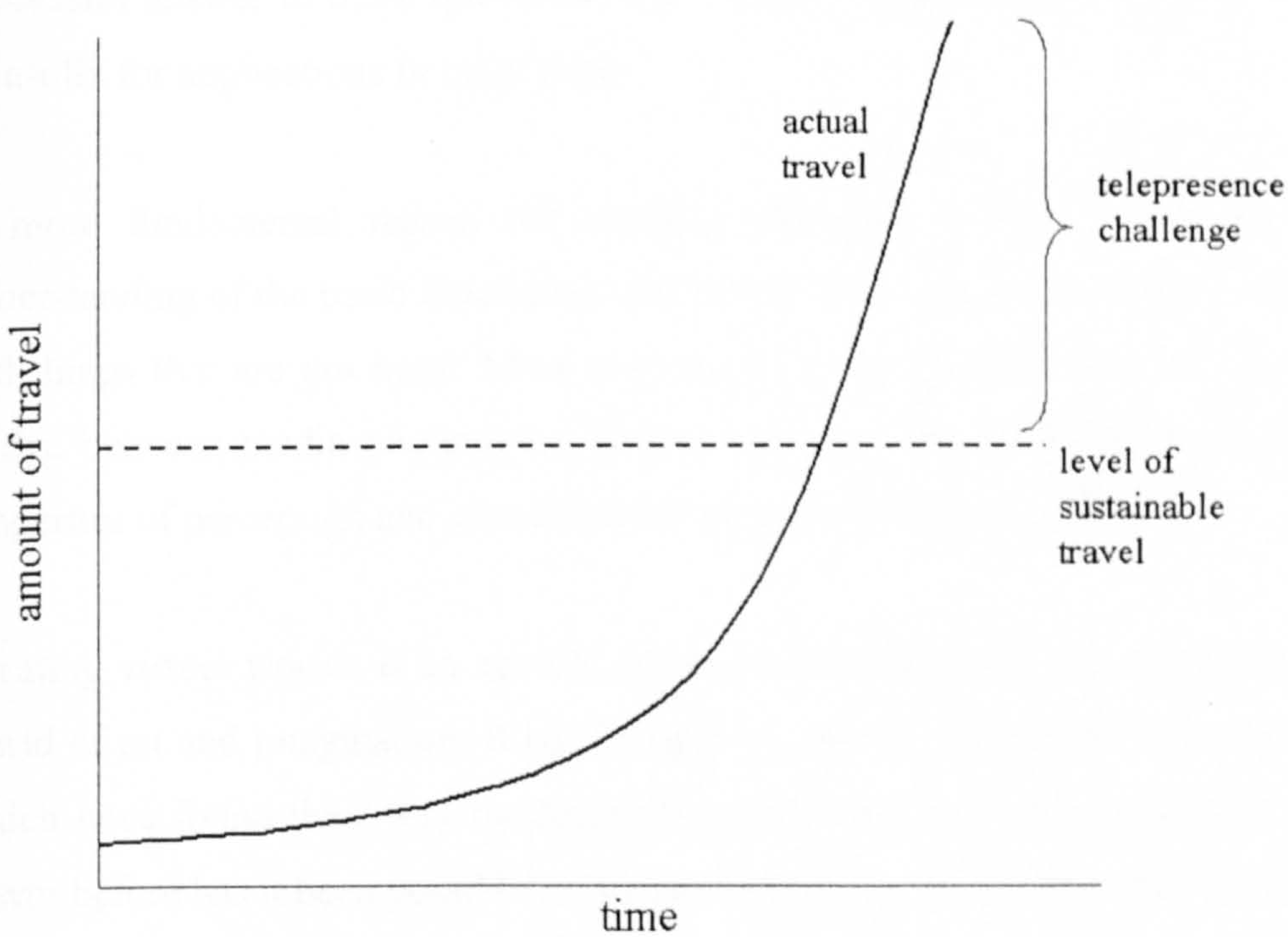


Figure 2. Telepresence technologies may potentially provide a solution for the challenges posed by the increase in physical travel beyond the level of environmental sustainability.

Another recent area of application with the potential of being extremely beneficial to a large number of people is the use of VE technology in the service of psychotherapy, in particular for the treatment of phobias (e.g. fear of heights, fear of spiders, fear of flying, etc.). As VEs allow full control over the feared stimulus (in contrast to in vivo exposure) at a predictable, relatively low cost (one major expenditure to equip the consulting office can be offset against the use of the equipment for many different situations, types of environment and patients), this has stimulated the use of VEs in phobia treatment programmes. However, for a virtual representation to provoke the same responses as the real situation we need to know what kind of VE is needed – is a photograph enough, is a movie enough? Also the answers to these questions might vary during the course of a treatment. Putting this another way: what is necessary in order to enhance sufficient *presence* in the VE that is representing the feared situation? What are the fundamental components in a representation that is necessary in order to achieve this sense of presence? How much like a real spider does the virtual spider have to look and behave? Presence, therefore, is of paramount importance in the successful exploitation of VEs in psychotherapy, and, of course, a successful answer to these questions in the context of psychotherapy may well have significant spin-offs for applications in other areas.

A more fundamental reason for studying presence is that it will further our theoretical understanding of the basic function of mediation: How do media convey a sense of places, beings and things that are not here? More importantly even, presence research provides the necessary bridge between media research on the one hand and the massive interdisciplinary program on properties of perception and consciousness on the other (Biocca, 2001).

Creating virtual worlds is an age-old desire of mankind, that traditionally was addressed by the world of art and imagination. It has brought the world of entertainment, the theatre, the movies, which have found their way by distribution and networks into the privacy of one's own home. Never before has it been possible to interact with the world of imagination in such a realistic way, that it renders the impression of being there, and creates the feeling of doing. In this sense the basic purpose of presence research is to increase the quality of life.

3. Current Issues and Research Challenges

In 1978, the philosopher Daniel Dennett presented us with an interesting and entertaining thought experiment entitled 'Where am I?' as the last chapter of his *Brainstorms* book (Dennett, 1978). Dennett recounts the story of a 'curious episode' in his life where his brain got surgically separated from his body, with each connection between them restored by placing two 'microminiaturized radio transceivers' between each input and output pathway. After the operation he, or rather his body, goes to visit his brain which was placed, in keeping with the best philosophical traditions, in a life-support vat. While looking with his own eyes at his own brain he starts to wonder:

> "*Being a philosopher of firm physicalist conviction, I believed unswervingly that the tokening of my thoughts was occurring somewhere in my brain: yet when I thought "Here I am," where the thought occurred to me was here, outside the vat, where I, Dennett, was standing staring at my brain.*" (p.312)

Dennett reasons that the location of the 'I' he was referring to in the question 'Where am I?' may be related, though not identical, to his *point of view*. He states:

> "*Point of view clearly had something to do with personal location, but it was itself an unclear notion. It was obvious that the content of one's point of view was not the same as or determined by the content of one's beliefs or thoughts. For example, what should we say about the point of view of the Cinerama viewer who shrieks and twists in his seat as the roller-coaster footage overcomes his psychic distancing? Has he forgotten that he is safely seated in the theater? Here I was inclined to say that the person is experiencing an illusory shift in point of view. In other cases, my inclination to call such shifts illusory was less strong. The workers in the laboratories and plants who handle dangerous materials by operating feedback-controlled mechanical arms and hands undergo a shift in point of view that is crisper and more pronounced than anything Cinerama can provoke.*" (p. 314-315)

What Dennett calls 'an illusory shift in point of view' nicely conceptualises the central idea of presence. His examples of this illusory shift are well-chosen. Cinerama, which debuted at the Broadway Theatre, New York in 1952, was one of Hollywood's answers to the growing popularity of television in the early fifties and the resulting decline of sales at the box-office. Cinerama used three 35mm projections on a curved screen to create a 146 degree wide panorama. In addition to the impressive visuals, Cinerama also included a 7-channel directional sound system that added considerably to its psychological impact. The ads for '*This is Cinerama*', the first Cinerama film, containing the famous scene of the vertigo-inducing rollercoaster ride, promised: "You won't be gazing at a movie screen – you'll find yourself swept right into the picture, surrounded by sight and sound." The film's program booklet proclaimed: "Everything that is happening on the curved Cinerama screen is happening to you. And without moving from your seat, you share, personally, in the most remarkable new kind of emotional experience ever brought to the theater" (Belton, 1992). Phrases such as these were used quite often to promote the immersive and multisensory cinema formats of the 1950's, including 3D cinema (e.g. "It happens to YOU in three dimensions"), Todd-AO ("Suddenly you're there..."), CinemaScope, and other formats (Lodge, 2000). Although such statements were sales pitches of the films' marketing people, they do illustrate the fact that the aim of the cinema experience was to enhance the film's psychological impact and entertainment value by making the viewer feel part of the movie. The once passive viewer became an 'active' participant. As Slater and Wilbur (1997) put it, with reference to virtual environments, the "discontinuity between the place of our current reality and the reality showing through the display" seemed to be collapsing.

Although non-interactive media can become highly immersive and realistic, engendering a compelling sense of presence for the participant, they still assume a rather passive role for the observer. Perception, of course, is more than mere 'sensing' of the environment through our various sense organs and subsequently matching these sensations against passive representations or templates of stored information. Rather, perception is a highly activity-dependent and context-dependent process (both embodied and environmentally embedded) that integrates multimodal sensory data, ongoing actions and intentions, and memory processes. Perception serves the individual's need to control behavior (action) within an environment. By allowing real-time action at a distance (through teleoperation) or in virtual space (through interactive computer-generated environments) the participant is able to control certain aspects of the mediated environment, and, as a consequence, his/her perception of the environment. In this way, the participant will become aware that he or she is an *actor* within the environment and it is likely that this experience of willful control or *feeling of doing* will greatly enhance the feeling of actually being there within the mediated environment – the sense of presence (see e.g. Welch, Blackmon, Liu, Mellers & Stark, 1996).

Dennett (1978) too states that interactive teleoperation engenders a "shift in point of view that is crisper and more pronounced than anything Cinerama can provoke." Remote-controlled manipulators (e.g. robot arms) and vehicles are being employed to enable human work in hazardous or challenging environments such as space exploration, undersea operations, minimally invasive surgery, or hazardous waste clean-up. The design goal of smooth and intuitive teleoperation triggered a considerable research effort in the area of human factors, which lies at the root of today's presence research (Johnson & Corliss, 1971; Sheridan, 1992). In fact, the term *telepresence* was first used in the context of teleoperation by Marvin Minsky (suggested to him by his friend Pat Gunkel) in his classic 1980 paper on the topic (Minsky, 1980).

Minsky's paper was essentially a manifesto to encourage the development of the science and technology necessary for a remote-controlled economy that would allow for the elimination of many hazardous, difficult or unpleasant human tasks, and would support beneficial developments such as the creation of new medical and surgical techniques, space exploration, and tele-working. The questions Minsky posed are still valid today. Although the remote-controlled economy didn't arrive in the way he envisioned, the development of telepresence technologies has significantly progressed in the various areas he identified. In addition, the arrival and widespread use of the internet brings us remote access to thousands of homes, offices, street corners, and other locations where webcameras have been set up (Campanella, 2000). In some cases, because of the two-way nature of the internet, users can log on to control a variety of telerobots and manipulate real-world objects.

It seems fair to say that the concept of presence has today become common currency in areas such as virtual environments, advanced broadcast and cinematic displays, teleoperation systems, and advanced telecommunication applications. Since the early 1990's, a growing community of multidisciplinary researchers has turned its attention to presence, looking at what causes it, how the experience may be measured, and what effects it has on the media user. Despite considerable progress, presence research is still very much in its infancy, with a great number of unresolved issues and challenges that set the research agenda of the presence community. The most significant issues are listed below:

- The *structure of presence* is still largely unclear. What are the phenomenological properties? For example, is it an all-or-none or a graded experience? Are there different forms? What are its psychological dimensions? As discussed previously, recent factor analytic studies are starting to shed light on this latter issue. Although a number of conceptual models of presence have been proposed to date, there has not been a generally accepted *explanatory model of presence* that relates current theoretical and empirical insights in perception, emotion, cognition and action to mediated experience, and has neural plausibility. A theory of presence needs to be established that has clear links not only to psychology and engineering but also to ontology, epistomology, consciousness studies and neuroscience.
- *Measuring presence* in a way that is reliable, valid and robust is still a major challenge, although this issue is currently receiving considerable attention from various research labs on both sides of the Atlantic. A good presence measure will allow engineers and media developers to identify the factors (and trade-offs between them) needed to optimise the level of presence for the media user. Moreover, a good presence measure will allow the research community to further develop its understanding and systematic investigation of the construct, which will in turn enable further refinement of measurement methodologies, and so on.
- A number of potential *determinants of presence* have been identified, and some have been experimentally validated. Studies have until now mainly focussed on medium-related form and content factors such as user-initiated control of the simulation (Welch et al., 1996), stereoscopic presentation (Hendrix & Barfield, 1996a; Freeman, Avons, Pearson, & IJsselsteijn, 1999; IJsselsteijn, de Ridder, Hamberg, Bouwhuis, & Freeman, 1998), head tracking (Hendrix & Barfield, 1996a), field of view (Hatada, Sakata, & Kusaka, 1980; Prothero & Hoffman, 1995), spatialized audio (Hendrix & Barfield, 1996b), response latency (Ellis, Dorighi, Menges, Adelstein, & Jacoby, 1997), and meaningfulness of the simulation (Hoffman, Prothero, Wells, & Groen, 1998). The relative contributions and interactions of these and other determinants are still largely unclear however. The resolvement of this issue

depends to a large extent upon the development of suitable measurement methodologies and a proper theoretical framework.

- The *effects of presence* are to a large extent still unclear. For example, under what circumstances does an enhanced sense of presence aid task performance, or learning and memory?
- What are relevant *individual differences* with respect to presence, and what impact do they have? Slater and colleagues (Slater & Usoh, 1993; Slater, Usoh, & Steed, 1994) have performed some initial studies on user characteristics, such as the extent to which a person's field dominance (e.g., visual, auditory) affects their sense of presence. Further study is required however to investigate the potential impact of individual characteristics (traits, states, needs, preferences, experience, gender, etc.) on presence.
- What will be the *social consequences* of the introduction of certain high-presence technologies at work or in the home? How will social relationships be affected by the long-term deployment of high-presence communication applications? Are there any significant *health and safety issues* that need to be taken into account? In which *contexts of use* will presence be of most value? How does presence relate to *basic human needs* such as privacy, control, and social contact?

In fact, the defining characteristics of the concept itself are still being discussed, as is evidenced by recent debates at the Presence Workshops, the European Presence Research Conference, and on presence-related newsgroups. Presence is used in different ways by different scholars, each looking at the concept from their own perspective, applying their own emphasis or using their own specific definition. Lombard and Ditton (1997) reviewed a broad body of literature related to presence and identified six different conceptualisations of presence: realism, immersion, transportation, social richness, social actor within medium, and medium as social actor. Based on the commonalities between these different conceptualisations, they provide a unifying definition of presence as the 'perceptual illusion of non-mediation', i.e. the extent to which a person fails to perceive or acknowledge the existence of a medium during a technologically mediated experience.

The different conceptualisations of presence identified by Lombard and Ditton (1997) can roughly be divided into two broad categories – physical and social. The physical category refers to the sense of being physically located in mediated space, whereas the social category refers to the feeling of being *together*, of social interaction with a virtual or remotely located communication partner. At the intersection of these two categories, we can then identify 'co-presence' or a sense of being together in a shared space, combining significant characteristics of both physical and social presence. Figure 3 illustrates their relationship with a number of media examples that support the different types of presence to a varying extent. For example, while a painting may not necessarily support physical presence to any great extent (although trompe l'oeil and panorama paintings may be examples to the contrary), interactive virtual reality (VR) technology has the potential to engender a high sense of physical presence.

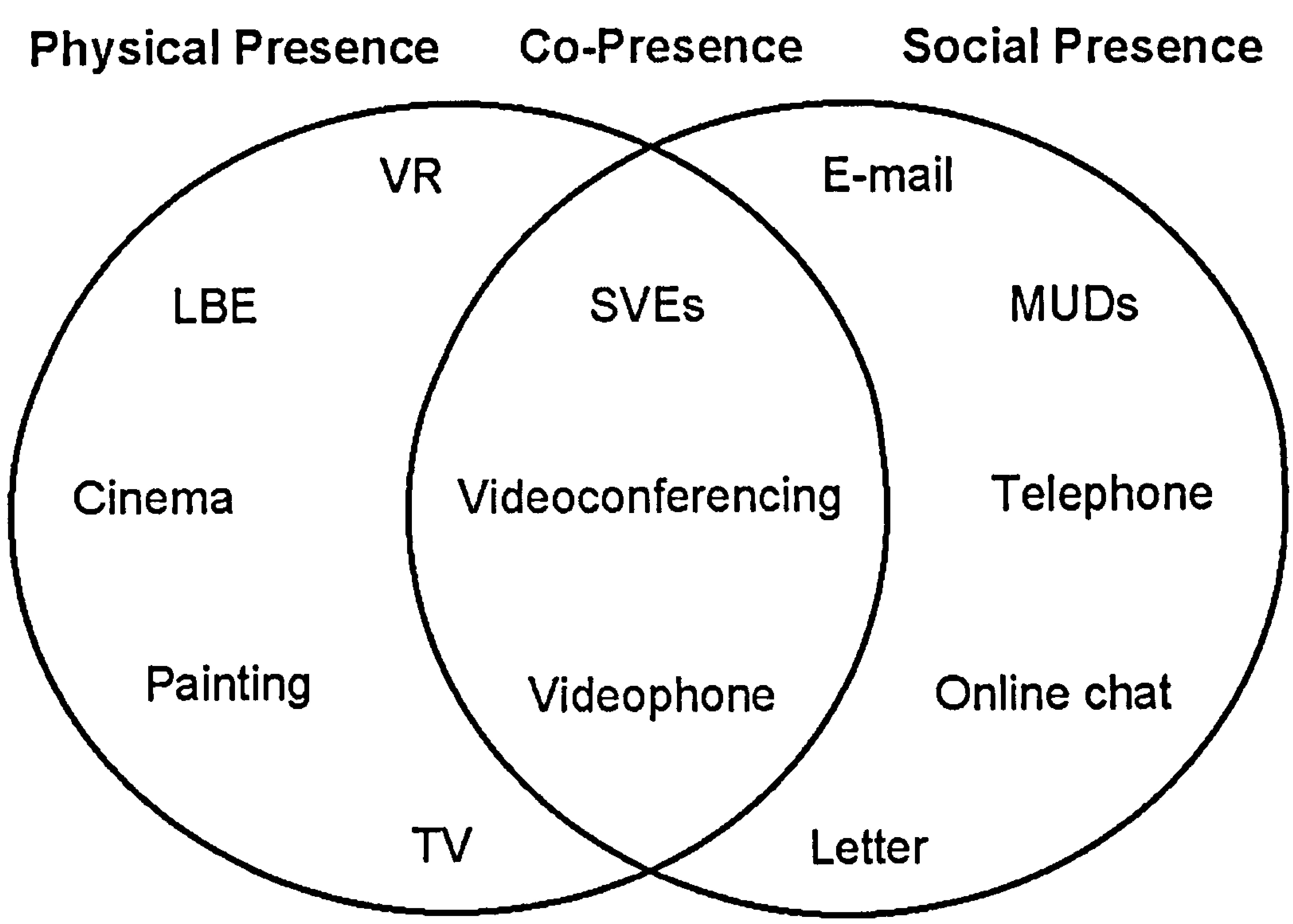


Figure 3. A graphical illustration of the relationship between physical presence, social presence and co-presence, with various media examples. Abbreviations: VR = Virtual Reality; LBE = Location-Based Entertainment; SVEs = Shared Virtual Environments; MUDs = Multi-User Dungeons.

4. Presence and Perception: Media Problems and Research Solutions

There has been an unchallenged assumption amongst many observers that the technologies underlying the current generation of audio and visual media are soundly based and have been thoroughly developed. However, recent psychophysical and perceptual discoveries strongly suggest that existing photography, film, television, audio recording/reproduction and digital transmission technologies are inadequate in important areas. Not least is the problem that some sensory modalities like olfaction, haptic, vestibular and "somatic" sensations are rarely recorded or replayed at all. There is now evidence these "information transmission" inadequacies can adversely affect very many users. This in turn can be demonstrated to cause inefficiencies, misrepresentations, misinterpretation, and even widespread public health issues.

Our basic understanding of human perception has improved significantly in recent years. But many of the standard psychophysical tests used by the audio and visual industries to design and evaluate their products were produced over 50 years ago and have led to very significant failings in current design and performance. Ask yourself a simple question: "Have I ever seen an image or heard a sound that I instinctively believed to be "real", when it was in fact produced by a machine?" Surprisingly, it is not only the general public would say "no" to this. It is almost unheard of for someone to walk up to an object and try to touch it, only to discover that it is an image on a screen. And so far as can be discerned, no one has recorded a natural sound and replayed it so that the reproduction cannot be discriminated from the original source. The reason for this is that the information we sense in our immediate environment is currently recorded and replayed in ways that are not fully compatible with the human perceptual systems. The qualitative, quantitative and temporal events that we experience in the real world are only crudely captured and replayed by existing media technologies. If such information is correctly recorded and replayed, then it is likely that people will no longer be able to discriminate between real and virtual stimuli, and they would report that they are in the *presence* of a "real" event or scene.

Primary sensory stimulation happens when a person experiences stimuli in reality, rather than through a mediated technology. In addition to the five main senses, our other senses (involving the vestibular/somatic systems) can detect body orientation, acceleration, impacts, temperature, air pressure, water immersion, electrical activity and many others. New sensory capabilities are being discovered with surprising regularity (e.g. Todd & Cody, 2000). All of these sensory inputs need to be combined correctly in the brain to give us the sense that we are present in a real environment and allow us to experience stimuli as if we were really there.

Artificial sensory simulation is the process by which current audio/visual technologies re-present captured environmental stimuli to our senses. There are many areas where inadequate environmental information is captured and/or corrupted so that the medium distorts the message in ways that impair its reality. This corruption of environmental information is why we never normally perceive sound and visual images so that they can be easily confused or compared with reality. Our brain's natural ability for pattern recognition means that when current audio/visual technologies provide only a poor approximation of the original stimulus, it can still recognise the source. But it will almost always categorise this source as an obviously artificial reproduction. If mediated artificial stimuli can activate *primary sensory stimulation* (in that it is perceived in the same way as a real event) then our abilities to discriminate the real event from an artificial sensory stimulus would be diminished to the point at which presence could be achieved through the mere perceptual realism of the stimulus. In the next paragraphs we describe two examples of how presence research may bridge the gap between fundamental insights in the mechanisms of perception and the way media represent reality in order to overcome inadequacies and distortions intrinsically present in current media technologies.

Why we need presence research: The case of colour perception and reproduction

Colour discrimination is an important part of human existence. Whether it is choosing between unripe or ready fruits, deadly or tasty mushrooms, live or neutral wires, stop or go lights, toxic or safe substances, colour perception is a fundamental aspect of most peoples everyday lives. Yet the RGB colour space that we currently use to artificially represent natural colours can only show approximately ten percent of the over one million brightness and colour combinations that most people can discriminate (Hurlbert, 2001). A simple but surprising example of this inadequacy can be found on almost any packet of flower seeds. People often buy seeds using the colour photograph on the packaging as a guide. They expect and even rely on the fact that the photograph will show an accurate colour reproduction of the adult plant. For obvious commercial reasons, it is vital that this image is a good match. Colour photography however is very poor at providing adequate representation of subtle colourations and natural hues. Photographers call this phenomenon "subject colour failure" and accept it as a fundamental limitation of their medium. Seed growers overcome the problem by using hand colouring over a monochrome rendition of the adult plant. They compare each colour on the adult plant against a physical colour chart, often from the Pantone printing ink colour space. These colour values are conveyed to the printers who then ensure that the correct dyes are printed over the monochrome photograph to ensure a good match. Commercial colour printers rarely use the RGB colour space, as they have understood its limitations for decades. Most photomechanical reproduction systems are based on the CMYK colour space, yet can have up to sixteen colour layers in the final print. Even then, it is known that there are some combinations or types of colours that the best systems still cannot reproduce.

Many people have assumed that the success of home shopping using colour catalogues would undermine the previous statement. However, catalogue shopping has often been expensive because of its returns policy. If a customer rejects a product and returns it because the colour of the garment or product they bought does not match the colour they thought they were buying, *then the cost of that failed transaction is carried by the company (and therefore the clients of the company).* It is a core inefficiency that undermines the profitability of home shopping. The Internet uses the restricted sRGB colour gamut to convey its photographic images. So it is hardly surprising that clothing does not sell well over the Net. If computer based digital imaging used a colour space that was derived from a fundamental understanding of the human visual system, then it would be possible for people to remotely sense stimuli that previously could only be experienced by direct contact. Doctors could diagnose many illnesses without physically being there, if imaging systems could convey the information they require. They use vision and haptic sensation in conjunction with question and answer interaction for many of their diagnoses. If they could do this remotely, then doctors could be telepresent almost anywhere and the time saved between accidents and treatment could save many lives. Military or police observers could use their highly trained visual acuities without exposing themselves to the danger of attack. Working on or above a location (in an aircraft) exposes many observers to danger as a normal part of their jobs. Bomb disposal experts, fighter pilots, submarine pilots, geographers, geologists, biologists, power line workers, safety engineers and countless other professions have to take risks to use their senses when it is now possible for them to do their work from a safe location. Presence research is currently the only route by which all of our sensory thresholds can be identified and realistically stimulated.

Why we need presence research: The case of body image distortion through photography

Poor colour discrimination is not the only problem caused by inadequacies in imaging technology. Recent research has indicated that conventional single lens photography is surprisingly poor at conveying images of people and objects with their correct size and shape information. Photographers are occasionally aware that scenes they see with their normal direct vision will change significantly in their 2D representations when they are seen on the page or screen. One of the best known, and most disconcerting changes, is the fattening effect of photography (Harper & Latto 2001). Models and actors are often chosen for slimness in the attempt to combat the fattening effects of photography. While looking "normal" in the photographic image, they often look frighteningly slim in real life. The reason for this is that cameras and people see the world from a single-point perspective. By the process of cyclopean vision (Julesz, 1971) we perceive through a "virtual eye" that generates an artificial viewpoint from a location mid-way between each real eye. However, recent experiments suggest that a single-lens camera cannot convey the way in which we can focus/fuse on an object with two eyes and simultaneously see both diverging and converging optical paths. Close-up objects viewed stereoscopically occlude less of the background than their monocular equivalents and therefor are seen as significantly slimmer. Failure to reproduce this perceptual geometry in a display medium is a major cause of the fattening effects and misperception in conventional photographic, film and television images.

The public health issues relating to these distortions were discussed in a major review document "Eating Disorders, Body Image and the Media" by the British Medical Association. One of the many findings of this report was that actresses on average are 20% percent slimmer than non-actresses of the same age. It is now generally accepted that that there has been an obsession with slimness in photography, film, television and print media. Yet recent work by Professor Rose Frisch of Harvard University has confirmed that body fat and fertility is intimately linked in pre-menopausal women. She says[1] "We can think of body fat as *Sex Fat*, because it provides the

[1] From the Observer newspaper interview, 26-0502 p.7. on the publication "Female Fertility and the Body Fat Connection."

energy required for reproduction. What I found as most astonishing is that there is a razor-thin borderline where losing just 3lb can tip a normal sized woman into infertility. And many women who maintain the body shape made popular on the catwalks are completely infertile. These women can, however, continue having periods and not guess that anything is wrong until they try, and fail, to get pregnant. If low body fat switches off a woman's reproductive system for too long, it may be too late to regain fertility simply by putting on weight. Women need to gain weigh at a younger age. A 5ft 3" girl must weigh at least 90lb if she is to become fertile, while a pre-menopausal woman of the same height needs to over 101lb if she is to continue to ovulate."

There is now widespread public health concerns regarding healthy people slimming to the point at which their health is seriously impaired. Early mortality is an inevitable outcome for large proportion of people who restrict their food intake in this way. It is probable that the fattening effect of photography has not previously been linked to anorexia nervosa and related illnesses because until the Harper and Latto studies, no one had proposed a mechanism by which it might occur. Their results suggest that anyone who cannot form a cyclopean view in the brain is likely to perceive images of people (with their direct vision) as fatter than observers with normal stereo vision. And surprisingly, many people are completely unaware that they do not have normal stereovision. Paradoxically, the photographic images that they see in magazines and television would appear to have a similar type of body image information as their own direct vision would give them. This could lead them to trust images that are already distorted because they have not been made orthostereoscopically. Orthoptists estimations vary on the amount of stereo acuity failure that exists within the general population, but figures of between one in six to one in nine are mentioned anecdotally. These people may have two eyes, each with excellent vision. But the mechanism for allowing them to fuse a stereo image in the brain is encumbered in various ways that effectively gives them zero measured stereo acuity. This leads to the important possibility, as yet not investigated empirically, that people with stereo acuity failure are more likely to develop a slimming disease. Presence in the media would allow remote viewers to perceive unhealthy levels of slimming as clearly as if they were in the presence of the real person, rather than using the "flattened and fattened" 2D image they see today.

5. A Vision of the Future

It is inevitable that new imaging and media technologies will be developed that will be more realistic than currently available systems. But effort and resources will be wasted on these developments until there is a fundamental understanding of how we perceive the world around us. Once we understand the thresholds, mechanism and interactions of all of the normal human sensory modalities, we can artificially deliver compelling and realistic experiences through the development of new media technologies. Current efforts to produce compression algorithms have failed to find solutions that do not throw away data vital to realism. This is almost certainly due to the fact that the psychophysical research on which their designs were based is outdated and inadequate. MPEG compression technology is now known to discard motion and colour data that it had been wrongly thought we could not see. Audio compression too has been hindered by the same theoretical failures, leading to the observation by audiologists that every compression system so far developed introduces artefacts that can clearly be perceived when compared to the original source. When the promised "unlimited bandwidth" data transmission systems arrive, heavy compression will be unnecessary and the prospects for presence will be significantly enhanced. Understanding how the brain produces such rich percepts through the seemingly impoverished sensory abilities given to us by nature will also allow for reductions in the amount of data required to reproduce reality.

The forces of motion itself can also now be effectively reproduced in a simplified immersive environment. Accelerations and forces can easily be recorded. And now they can be replayed without physically moving the subject. Galvanic Vestibular Stimulation (Scinicariello, Eaton, Inglis, & Collins, 2001) works by using tiny currents behind each ear to selectively stimulate the vestibular system. A prototype device can now "fine tune" sensation to make balancing easier, or "move" a subject so that they feel as if they are on a motion platform. Expensive motion platform style simulations could soon become commonplace once the need to "move the room" has been replaced by a tiny headset. Olfaction too is now on the list of things we can realistically simulate. DigiScents Technologies (now bankrupt) claimed to be able to record smells and replay them from a small palette of source scents that could be mixed to reproduce the original stimuli. The most pertinent challenges in this area are the reliable capture of smell (i.e. an artificial "nose") and the swift removal of the smells once they have been delivered.

Much of the travelling that people commit themselves to in a lifetime is out of duty rather than choice. Presence research shows that we do not need to travel to experience social presence or to use our senses to experience a remote stimulus directly. In the future, humans will physically travel to a destination because they want to, rather than need to do it. For instance, we would be able to attend conferences or interact physically and socially at a remote location without leaving our homes or workplace. It would be almost like the teleport technology so beloved by popular science fiction writers. However, it is likely that the software projection holograms found in science fiction films will not be easy to reproduce in the near future. In the film "Star Wars," a low-resolution holographic image of the Princess is projected in 3D space as a vital part of the plot. This simple effect is as science fiction now as it was 25 years ago when the scene was filmed. While some groups have reported success in capturing still holographic images without the need to use laser illumination (Marks, Stack, Brady, Munson, & Brady, 1999), no one has yet speculated on what type of display could reproduce the data or make it move realistically. The data rates required to do this are an order of magnitude above what is possible using envisioned developments of computer technology. Fortunately for presence research, many of the characteristics of software holograms have already been approximated using existing lens based imaging systems.

Realistic remote sensing and interaction is not the only possibility offered by presence research. Extended "hypersensory" sensation would also be possible too. Most lens-based imaging can easily have its sensitivities increased to include ultraviolet, infrared, and polarised light. It is feasible that these emissions could be presented within an immersive multisensory display so that engineers could sense stress patterns, hot spots and microscopic details that they could not see with direct sensory stimulation. Perhaps sensitivity could be extended to x-rays, gamma rays and other useful electromagnetic spectra. These could be added to expert computer systems that could be of great assistance to many professional disciplines.

In practice however, the most widespread use of presence technologies is likely to be in museums, encyclopaedias, entertainment, and the mass media industries. The enhancements that this technology could bring to people's lives are immense. It could become so pervasive that it might one day become difficult to imagine life without presence as part of our day to day reality. We might have life-long and intimate relationships with people that we have never physically met. The possibility to "travel" to places that are impossible for us to physically visit, or might not even exist any more is also enticing. The film industry is currently investing in the development of "synthespians." These are virtual actors who look, sound and appear to interact as if they were real. For studios that own the likenesses of long-dead Hollywood stars, the prospect of a "new" Marilyn Monroe, Humphrey Bogart or Bruce Lee film is too tempting to ignore. If this technology was coupled to a high-speed rendering engine and an interactive artificial intelligence program, then we could "meet" people in virtual environments who do not actually exist. The possibilities of this type of development are almost endless.

It could be argued that human perception is so sensitive that technology could never fool it into completely believing that even the finest artificial stimulus is in fact real. This argument has now been challenged by recent research using Transcranial Magnetic Stimulation (Cook & Persinger, 1997). By using tiny rotating magnetic fields (1 microtesla) a sense of human (and sometimes even "god-like") presence is induced in people who are in a sensory deprivation experiment. It works in over 80% of subjects at signal levels barely above the background noise in the brain. Subjects report that this sensed "physical presence" is vivid, exciting and completely believable. It really does not seem like science fiction to imagine a small TMS headset that one would wear while watching a next generation multi-sensory display. It could help us to suspend our disbelief that what we are experiencing is not real. It could also allow lifelike perception to become a reality, even in subjects who might be thought to be resistant or sensitive to artificial sensory stimulation. It might also allow less than lifelike resolutions, colour reproduction and sensory interactions to be perceived as completely "real." The combination of precisely gathered environmental information (that is conveyed in an uncorrupted way to the user) with this lowered reality threshold could make presence become a reality that would enhance the quality of life for millions of people. We simply need to do the basic research first and then creatively apply its findings.

References

Belton, J. (1992). *Widescreen Cinema.* Cambridge, MA: Harvard University Press.

Biocca, F. (2001). Presence of M.I.N.D. Invited talk at the European Presence Research Conference, Eindhoven, 9-10 October 2001.

Biocca, F. & Delaney, B. (1995). Immersive virtual reality technology. In: Biocca, F. & Levy, M. (eds.), *Communication in the Age of Virtual Reality.* Hillsdale, NJ: Lawrence Erlbaum, pp. 57-124.

Campanella, T.J. (2000). Eden by wire: Webcameras and the telepresent landscape. In: Goldberg, K. (ed.), *The Robot in the Garden.* Cambridge, MA: MIT Press, pp. 22-46.

Cook, C.M. & Persinger, M. (1997). Experimental induction of the "sensed presence" in normal subjects and an exceptional subject. *Perceptual Motor Skills 85,* 683-693.

Dennett, D.C. (1978). *Brainstorms. Philosophical Essays on Mind and Psychology.* Brighton, UK: Harvester Press.

Draper, J.V., Kaber, D.B., & Usher, J.M. (1998). Telepresence. *Human Factors,* 40: 354-375.

Ellis, S.R., Dorighi, N.S., Menges, B.M., Adelstein, B.D., & Jacoby, R.H. (1997). In search of equivalence classes in subjective scales of reality. In: Smith, M.J., Salvendy, G., & Koubek, R.J. (eds.), *Design of Computing Systems: Social and Ergonomic Considerations.* Amsterdam: Elsevier, pp. 873-876.

Freeman, J., Avons, S.E., Pearson, D.E., & IJsselsteijn, W.A. (1999). Effects of sensory information and prior experience on direct subjective ratings of presence. *Presence: Teleoperators and Virtual Environments 8,* 1 - 13.

Freeman, J., Avons, S.E., Meddis, R., Pearson, D.E. & IJsselsteijn, W.A. (2000). Using behavioural realism to estimate presence: A study of the utility of postural responses to motion-stimuli. *Presence: Teleoperators and Virtual Environments 9*, 149-164.

Harper, B., & Latto, R. (2001). Cyclopean Vision, Size Estimations and Presence in Orthostereoscopic Images. *Presence: Teleoperators and Virtual Environments 10*, 312-330.

Hatada, T., Sakata, H. & Kusaka, H. (1980). Psychophysical analysis of the "sensation of reality" induced by a visual wide-field display. *SMPTE Journal 89*, 560-569.

Hendrix, C., & Barfield, W. (1996a). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators and Virtual Environments 5*, 274 - 289.

Hendrix, C., & Barfield, W. (1996b). The sense of presence within auditory virtual environments. *Presence: Teleoperators and Virtual Environments 5*, 290 – 301.

Hoffman, H.G., Prothero, J., Wells, M.J., & Groen, J. (1998). Virtual chess: Meaning enhances users' sense of presence in virtual environments. *International Journal of Human-Computer Interaction 10*, 251-263.

Hurlbert, A. (2001). Colour perception in the natural world: From quanta to qualia. Invited talk at the 'Reality and our Sense of Presence" conference, IST 2001 Event, Düsseldorf, 4 December 2001.

IJsselsteijn, W.A., de Ridder, H., Freeman, J., & Avons, S.E. (2000). Presence: Concept, determinants and measurement. *Proceedings of the SPIE 3959*, 520-529.

IJsselsteijn, W.A., de Ridder, H., Freeman, J., Avons, S., & Bouwhuis, D. (2001). Effects of stereoscopic presentation, image motion and screen size on subjective and objective corroborative measures of presence. *Presence: Teleoperators and Virtual Environments 10*, 298-311.

IJsselsteijn, W.A., de Ridder, H., Hamberg, R., Bouwhuis, D., & Freeman, J. (1998). Perceived depth and the feeling of presence in 3DTV. *Displays 18*, 207-214.

Johnson, E.G. & Corliss, W.R. (1971). *Human Factors Applications in Teleoperator Design and Operation*. New York: Wiley-Interscience.

Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago: University of Chicago Press.

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2000). Development of a new cross-media presence questionnaire: The ITC-sense of presence inventory. In: W.A. IJsselsteijn, J. Freeman, & H. de Ridder (eds.), *Proceedings of PRESENCE 2000 - 3rd International Workshop on Presence* (CDROM), Delft, 27-28 March 2000.

Lodge, N. (2000). What can we learn from the cinema of the 1950's? Invited talk at PRESENCE 2000 - 3rd International Workshop on Presence, Delft, 27-28 March 2000.

Lombard, M. (1995). Direct responses to people on the screen: Television and personal space. *Communication Research,* 22: 288-324.

Lombard, M., & Ditton, T.B. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3(2). [online]

Marks, D.L., Stack, R.A., Brady, D.J., Munson Jr., D.C., & Brady, R.B. (1999). Visible Cone-Beam Tomography with a Lensless Interferometric Camera. *Science 284*, 2164-2166.

Minsky, M. (1980, June). Telepresence. *Omni*, pp. 45-51.

Prothero, J. & Hoffman, H. (1995). Widening the field of view increases the sense of presence. HITLab Technical Report R-95-5. [online]

Schubert, T., Friedman, F., & Regenbrecht, H. (1999). Decomposing the sense of presence: Factor analytic insights. Paper presented at the 2nd International Workshop on Presence, University of Essex, 6-7 April 1999.

Scinicariello, A.P., Eaton, K., Inglis, J.T., & Collins, J.J. (2001). Enhancing human balance control with galvanic vestibular stimulation. *Biological Cybernetics 84*, 475-480.

Sheridan, T. (1992). *Telerobotics, Automation and Human Supervisory Control.* Cambridge, MA: MIT Press.

Slater, M. & Usoh, M. (1993). Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators and Virtual Environments 2*, 221-233.

Slater, M., Usoh, M. & Steed, A. (1994). Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environment 3*, 130-144.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments 6*, 603-616.

Todd, N.P.M. (in press). Evidence for a behavioural significance of saccular acoustic sensitivity in humans. *Journal of the Acoustical Society of America.*

Todd, N.P.M. & Cody, F. (2000). Vestibular responses to loud dance music: A physiological basis for the "rock and roll threshold"? *Journal of the Acoustical Society of America 107*, 496-500.

Welch, R. B., Blackmon, T.T ., Liu, A., Mellers, B. A., Stark, L. W. (1996). The effects of pictorial realism, delay of visual feedback, and observer interactivity on the subjective sense of presence. *Presence: Teleoperators and Virtual Environments*, 5: 263-273.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7: 225-240.

APPENDIX C

Leonardo (2007), in press.

The Non-Realistic Nature of Photography: Some More Reasons Why Turner Was Wrong

Richard Latto and Bernard Harper

Richard Latto (psychologist), School of Psychology, University of Liverpool, Eleanor Rathbone Building, Bedford Street South, Liverpool, L69 7ZA, U.K. E-mail: <latto@liverpool.ac.uk>. Web: <http://www.liv.ac.uk/Psychology/DeptInfo/StaffProfile/RMLatto.html>.

Bernard Harper (photographer/psychologist), School of Psychology, University of Liverpool, Eleanor Rathbone Building, Bedford Street South, Liverpool, L69 7ZA, U.K. E-mail: <bernard.harper@liverpool.ac.uk>. Web: <http://www.liv.ac.uk/Psychology/postgraduate/bharper.html>.

ABSTRACT

The authors discuss the limitations of photography in producing representations that lead to the accurate perception of shapes. In particular, they consider two situations where the photographic representation, although an accurate reproduction of the geometry of the two dimensional image in the eye, does not capture the way human vision changes this geometry to produce a three-dimensionally accurate perception. When looking at a photograph, the viewer's uncertainty about the camera-to-subject distance and the fact that, unnaturally, a photograph presents almost exactly the same view of an object to the two eyes result in substantially distorted perceptions. These most commonly result in a perceived flattening and fattening of the three dimensional shape of the object being photographed.

This is the end of art. I am glad I have had my day.

—J.M.W. Turner, on seeing his first daguerreotype [1]

It is now accepted that photography is not a simple reproduction of the world and that the photographic process, which always allowed a degree of manipulation has, with the advent of digital photography, become as subjective an art form as any other [e.g. 2, 3].

There is however one part of the process where photography is still assumed to be totally objective. A high quality camera lens produces a geometrically accurate image [4] of the subject being photographed on the film or sensor chip. This in turn will generate an accurate representation of the image on the camera's display system (photographic print, VDU, etc). The representation differs in size from the original image but preserves exactly all its other spatial attributes. However, this apparent exactness is misleading because the representation of the image in the display system is different from the representation in our brains in a number of significant ways. As a result the *perception* of the photographic representation is distorted and is not an accurate depiction of the object or scene that generated the image in the first place. The problems arise mainly because the final representation created by the camera is on a two-dimensional surface of limited size. Representational artists have the same limitation but they have the freedom, until recently largely denied to photographers, to manipulate their depiction of the world until the perception it generates matches their perception of the original subject. As a result, the artist's representation can elicit a perception that corresponds more accurately to the real world than a raw photograph can. As Ciuffreda and Engber [5] have described, two-dimensional representations lack binocular cues to depth, resulting in a flattening effect unless they are viewed monocularly. This loss of stereopsis is well studied and well understood. However, there are other important distortions in our perceptions of photographs which are less well known (except as we shall see to Leonardo da Vinci) or, as in our second example, only recently discovered.

The Absence of Context

When we look at a three-dimensional object, it produces an image on the retina which is ambiguous in terms of the size and shape of that object. There will be an infinite number of other objects of different sizes and shapes that could have produced the same retinal image. The shape of the retinal image produced by even a two-dimensional object will depend on the *angle of view*. Normally this ambiguity can be resolved, even in a photograph, from the context and causes no problems. The retinal image or the photographic representation of the wheels of a car are usually elliptical but because of the context we see them as round.

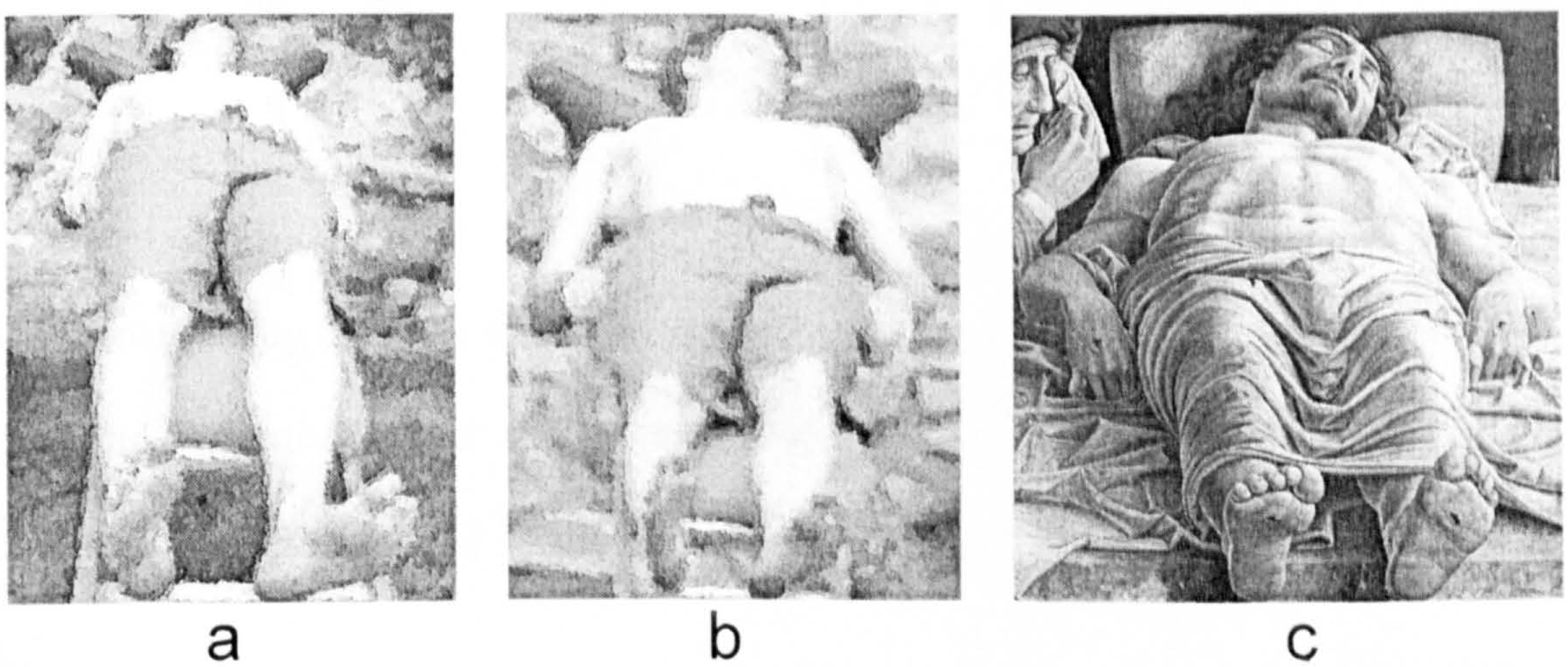
Both the size and shape of the retinal image will also vary with *distance*. For size, this is obvious. The magnitude of the retinal image will depend on the distance of the object from the eye. Less obviously, this is also true for shape. The shape of the two-dimensional retinal image formed by a three-dimensional object varies with distance. To resolve the ambiguities in size and shape, either we need to know the distance, so that we can scale the apparent size up or down accordingly, or we need to have a strong expectation as to what the correct interpretation of the image should be. For size, these expectations are particularly important. We scale our perception of the world and the objects within it according to our prior knowledge of the normal size of people, cars, houses, etc.. This makes knowing the distance less important [6]. With shape identifying the variations from the norm is more crucial, for example in recognising a person from their facial features. Consequently, expectations here, though still important, are less useful than for size.

All these problems of ambiguity also apply when we look at a pictorial image. We can usually resolve size from the context-defined distance and/or our prior knowledge. When we have no prior knowledge of size, the more cues to distance there are in the picture, the better we are at judging it [7]. Shape and, more generally, spatial layout is much more problematic and creates a substantial problem in photography.

A photograph is taken from a viewpoint that, to the viewer of the subsequent photographic representation, is not normally easily identifiable. A studio portrait is a good example of this. We usually have no way of telling the camera-to-subject distance. With direct viewing of a scene, an object subtends a visual angle on the retina that is correlated to its distance from the viewer, and our visual processing system ensures that the perceived shape of that object remains approximately the same irrespective of its distance (shape constancy). When we view a photograph, however, the viewpoint is often incorrectly determined and constancy mechanisms make incorrect adjustments to shape even when, as with a portrait, we know the size from experience. The result is a distorted perception: the perspectives in the photograph are perfect; the resulting perceptions are not. The apparently enlarged feet in a close up of a reclining person is a good example of this (Fig.1a). If we view

Fig. 1 a) & b). Photographic representations of a reclining figure taken with the camera at the same angle to the subject, but at two different distances, as measured from his feet: a) 1 metre; b) 30 metres.

c) Andrea Mantegna, *Lamentation over the Dead Christ* (trimmed to match the photographs), Pinacoteca di Brera, Milan. (Reproduced under licence from the Italian Ministry for Cultural Goods and Activities.)



an actual person from the same position as the camera, they do not look distorted because the representation in our brain, unlike the representation in the photograph, is adjusted to make allowance for the viewing distance. The perspective in the image on the retina is identical to that in the photograph in Fig. 1a, but because in the real world situation there are cues indicating that the person is being viewed from very close up the brain adjusts the representation of the image accordingly. This does not happen when looking at the photograph because the viewing distance is incorrectly interpreted as being greater than it actually was and shape constancy breaks down. Fig. 2a shows a representation of a manikin viewed from 1 metre which

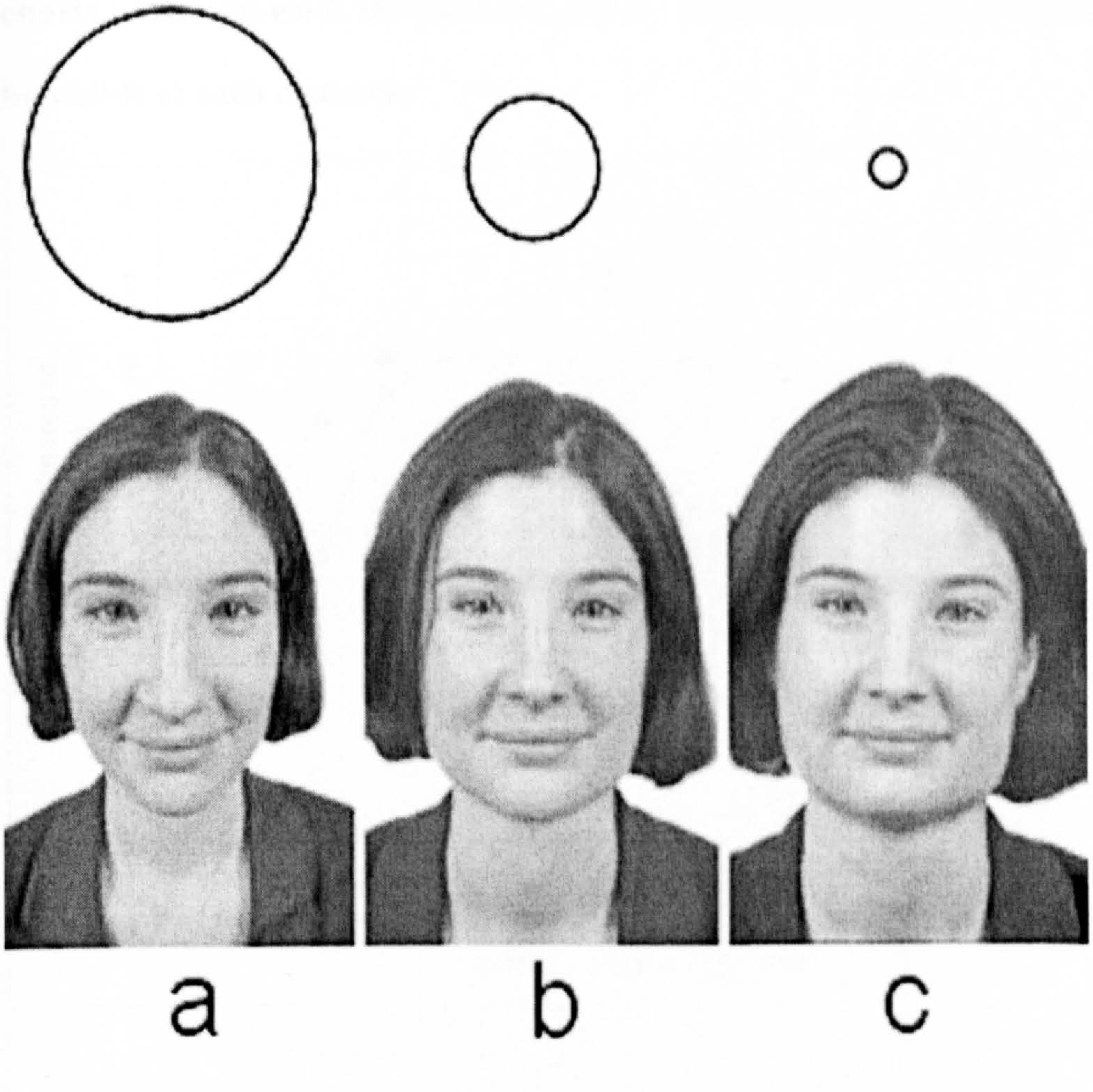
Fig. 2 Digitally generated representations of a manikin produced by a software package [8] that accurately reproduces perspective at two different distances: a) 1 metre; b) 30 metres.

was generated by software that calculates the correct perspective mathematically. The resulting perception appears distorted in the same way as in the photograph in Fig. 1a, confirming that the perspective in the latter is accurately represented.

This example of perceptual distortion when viewing a photograph taken very close up is well known, though it is often misinterpreted as being due to optical distortion by the camera rather than to a perceptual effect. Much less well known is the perceptual distortion of shape when viewing a long distance, telephoto photograph (Figs. 1b & 2b). Here, rather counter intuitively, the nearer part appears to shrink relative to the further part. The resulting representation is strikingly similar to Mantegna's painting of the dead Christ (Fig. 1c). As in the photograph of the man taken at 30 metres in Fig. 1b and the digitally synthesised manikin in Fig. 1b, the Mantegna painting is an accurate representation of the perspective of a body viewed from a distance. Why Mantegna should have done this is intriguing. He may have been able to map the shape of a distant object accurately, either using his own unaided perception or using some kind of grid system [9]. If the former, then his perceptual system must have been highly trained to see the perspective of the image accurately, perhaps from long experience painting ceiling frescoes designed to be viewed at a distance. Most of us, although we have much experience of looking at distant objects, are not aware of the perspective view they create. We see the object not the image on the retina and Mantegna's painting looks distorted. Indeed art historians discussing this painting commonly describe what they see as "inconsistencies of scale" [10], the tendency to "play down the relative increase in size of the nearest forms" [11] as a deliberate distortion to emphasise the more important parts of the body. This tendency is often cited as a forerunner of the deliberate distortions of 16th century mannerism [12]. In fact as Figs. 1b and 2b demonstrate, Mantegna's perspective is correct for a viewing distance of around 30 metres. As with the extreme close-up of Fig. 1a, it is because we are not aware of the artist's viewpoint that the body looks distorted. A life-size painting of a body with this perspective which was viewed from the artist's viewpoint of 30 metres would not look distorted. This is something that could perhaps be achieved in a large scale fresco, but not in a relatively small painting like *The Lamentation over the Dead Christ.*

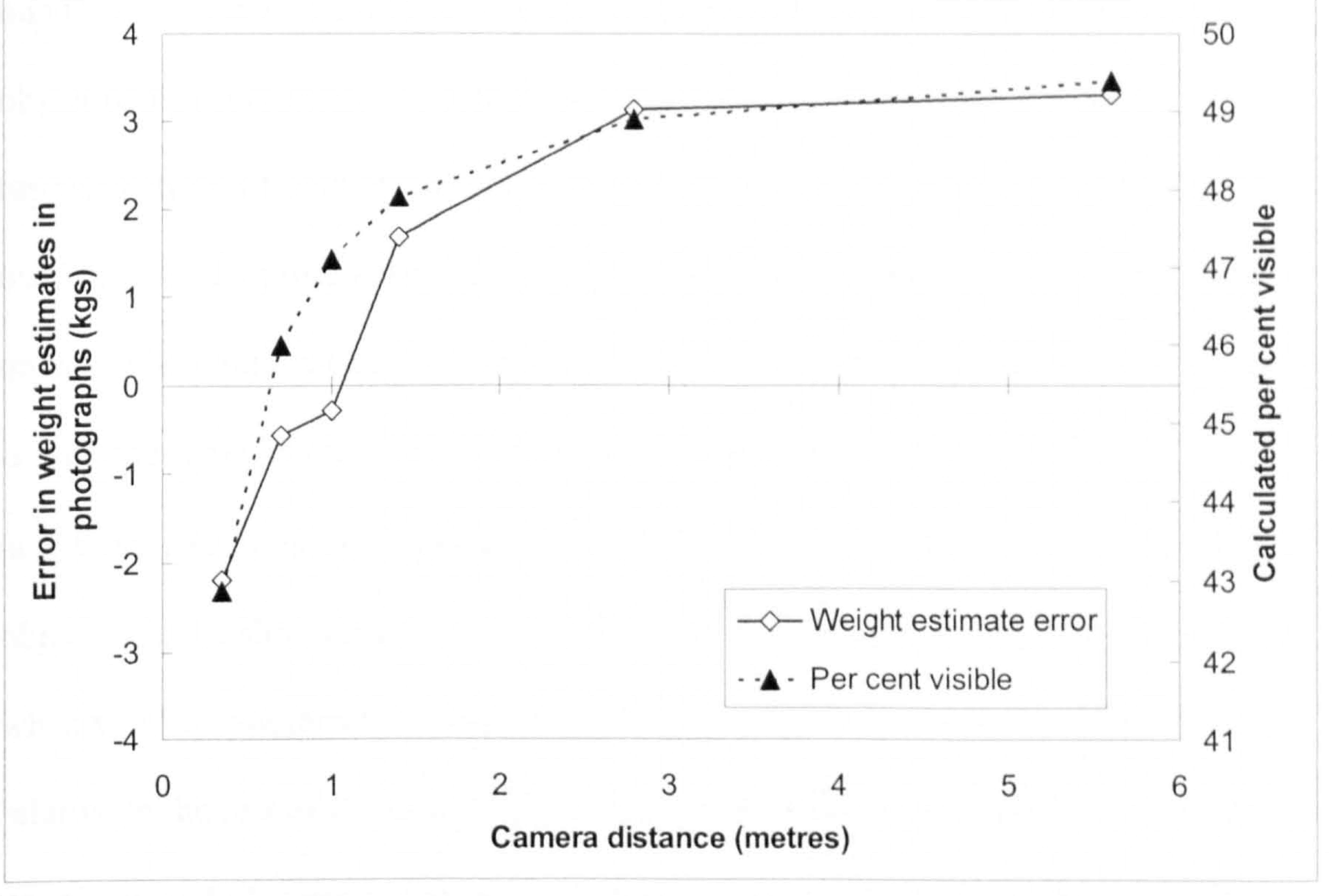
The flattening and fattening effect in photographs of distant objects taken using a telephoto lens is particularly striking when looking at people because we have visual systems that are highly tuned to discriminate small differences in the shape of features in order to facilitate recognition. Fig. 3 shows photographs of the same person in near

Fig. 3. Photographs, scaled to the same interocular distance, taken from: a) 0.32 m; b) 0.7 m; c) 2.7 m. The circles show the relative sizes of the image on the camera film (or the retina if the face had been viewed directly) at the three different distances.



identical poses taken from three different distances. As the distance increases the person looks fatter, an effect we have measured in terms of weight judgments (Fig. 4).

Fig. 4. Distortions in the perception of the weight of people in photographs: The solid line shows the errors in the mean estimates from 6 participants of the extent to which 12 people were over or under weight, plotted against the camera distance at which they were photographed. (The actual extent to which a person was over or under weight was calculated using standard body mass index (bmi) charts.) Also shown is the percentage of a 20 cm diameter sphere which would be visible at each distance.



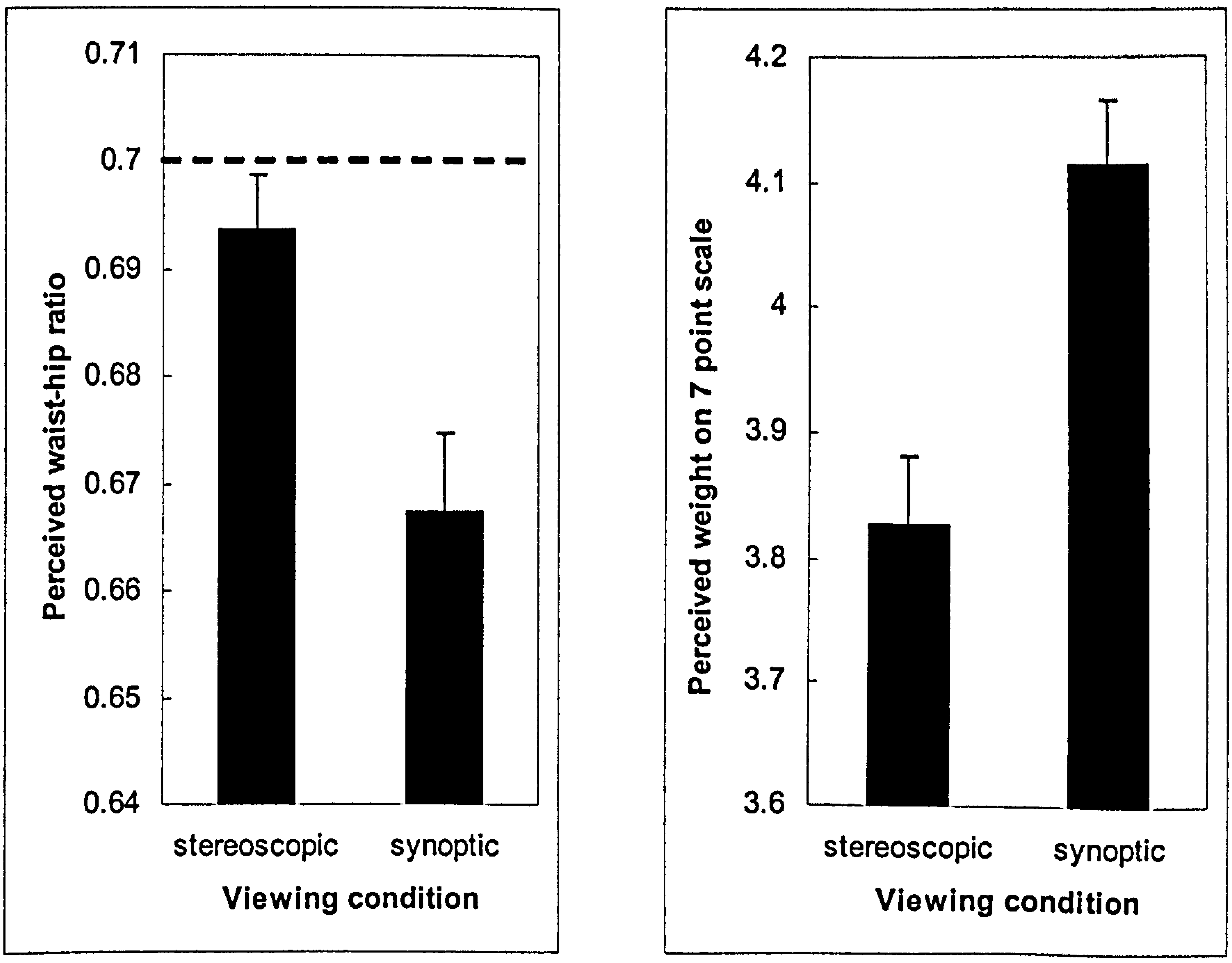
This effect can be understood in terms of Leonardo da Vinci's observation that "the further that a spherical body is from the eye, the more of it you will see" [13]. The dotted curve in Fig. 4 is derived from the calculated value of the proportion of a sphere approximately the size of a human head which is visible from different distances. The shape of this curve matches very closely the shape of the curve plotting the distortion in perceived weight.

The Synoptic View

When we view the three dimensional world with two eyes, we form a single perception from two slightly different images. When viewing a two dimensional photographic image we form images in the two eyes that are essentially identical, apart from some slight vertical disparity [14]. There is no significant horizontal disparity [15]. This synoptic (i.e. the same to both eyes) view of the world is quite unnatural. As with the telephoto view, it results in a perception which fattens the object being represented. We have demonstrated this effect in photographic representations of both abstract shapes and people [16]. For the abstract shape, we used projected representations of a large, computer-generated vertical peanut-shaped object. We compared estimates of 'waist-to-hip' ratio when viewed synoptically, that is with the same image in each eye, with estimates when viewed stereoscopically, that is with images in the two eyes differing slightly as they would when viewing an actual object. Fig. 5a shows that the peanut was seen as having a smaller waist-to-hip ratio when viewed synoptically, as in a normal photograph. The waist is seen as fatter relative to the rest of the peanut and, using the terminology used when describing people, the whole peanut is therefore seen as fatter. This fattening effect was also measured in photographs of people by comparing the estimated weights of models shown in life size photographic images when viewed synoptically or when viewed stereoscopically (Fig. 5b). The models too looked fatter when viewed synoptically.

Not all renaissance artists striving for reality in their representations were aware of the inability to represent accurately in a painting the view of the world generated by two eyes [17]. However, Leonardo da Vinci certainly was, observing that the two eyes are

Fig. 5. The mean estimates of perceived size and weight of photographic representations when viewed synoptically, as with a normal photograph, or stereoscopically, as when viewing an actual scene. a) The mean estimates by 20 participants of the ratio of the widest to the narrowest part of a vertically oriented, computer-generated peanut-shaped object. The dotted line shows the actual 'waist-to-hip' ratio. b) The mean estimates by 28 participants of the weights of ten female models photographed in swimsuits. For further details see [15].



sometimes looking at different surfaces on objects and even in the extreme at different objects [18]. This was one of the reasons why he was not as keen as some of his contemporaries on using grids for transferring perspectives from the real world onto a two dimensional surface [19,20]. However, as far as we know, he did not discover the perceptually fattening effect of so doing.

The Broader View

Perception is a complex, creative process finely tuned to the particular physical world in which we have evolved in order to create an accurate perception of that world [21]. We have given two examples of ways it does this which are not easily available to photography: adjusting shape representation to allow for changes in geometry with varying distance and the use of binocular vision in determining the three dimensional structure of objects. Unfortunately for those working in the visual media or subject to the paparazzi's telephoto lenses, in both cases the lack of these in photography results in flattening and fattening [22]. There are of course many other high level visual processes not available to the camera, from aspects of the perception of colour to the perception of human relationships. At a more basic level, photographs, as with the human eye, have varying focus at different distances. However, since the camera's depth of field rarely matches the eye's depth of field, this will give a misleading cue to depth, again probably facilitating perceptual distortions of the kind we have been discussing. At an even simpler level, the resolution of a photograph is normally uniform across its surface while the resolution of the eye falls off rapidly as we move away from the direction of gaze. (Both these phenomena are exploited by painters, particularly portrait painters, and film directors to control the centre of interest in a painting or a film.) The even resolution across a photograph should not matter with a life size representation. The pattern of resolution across the retina will match that when viewing the real world. However, with a reduced scale, this pattern becomes abnormal and again our perceptual system may draw false conclusions [23].

The fun, excitement, and importance of art derive in part from the attempt it makes to link perceptions of the objects it creates to perceptions of the world. Visual artists, like visual scientists, explore, and then exploit, the nature of visual perception. Their creations may strive for realism, as with Mantegna, or they may at the other extreme strive for aesthetic effects through exaggeration or abstraction, as with much art since the beginning of the twentieth century. In all cases though, as we have argued more fully elsewhere [24], the effect is the result of a careful articulation between the form of the creation and the nature of human perception. This is a process that uses the artist's own experience, sensory or otherwise, to create an object which will generate a particular perception in the observer. Visual scientists usually develop their hypotheses by recording and measuring the visual perceptions of others. Visual artists develop their creations, initially at least, by training and observing their own perceptions.

Photography is unreal; it alters values and perspective. Its cowlike eye stupidly registers everything that our eye has to correct and distribute according to the needs of the case.

—Jean Cocteau [25]

Acknowledgments

We are grateful to the members of the Visual Perception Group at Liverpool University for their valuable input, particularly to Luke Jones for preparing Fig. 2. The four referees also provided many helpful and thoughtful comments which we have done are best to incorporate.

References and Notes

1. J.M.W. Turner, quoted in: Links, J.G., *Canaletto and his Patrons.* (London, U.K.: Elek, 1977) p.104.

2. E.H. Gombrich, *The Image and the Eye.* (Oxford, U.K.: Phaidon, 1982).

3. C. Parkes, "A Loaded Paintbrush: Interview with David Hockney," *Financial Times Magazine* November 22, 16--17 (2003).

4. Terminology and its meaning can vary between, and even sometimes within, disciplines. The definitions of the main terms used in this article are given in the glossary.

5. K.J. Ciuffreda and K. Engber, "Is One Eye Better Than Two When Viewing Pictorial Art?" *Leonardo* Vol. 35, No. 1, 37--40 (2002).

6. R.N. Haber and C.A. Levin, "The Independence of Size Perception and Distance Perception," *Perception and Psychophysics* Vol. 63, No. 7, 1140--1152 (2001).

7. R.J. Miller, "An Empirical Demonstration of the Interactive Influence of Distance and Flatness Information on Size Perception in Pictures," *Empirical Studies of the Arts* Vol. 22, No. 1, 1--21 (2004).

8. MAXON Cinema 4D (MAXON, GmbH, Friedrichsdorf, Germany).

9. M. Kemp, *The Science of Art.* (New Haven, U.S.A.: Yale University Press, 1990)

p.43.

10. J. Martineau (editor), *Andrea Mantegna.* (London, U.K.: Royal Academy of Arts, 1992), p.155.

11. See Kemp [9], p.43.

12. The Modern Journal Project, *Andrea Mantegna.* http://www.modjourn.brown.edu/mjp/image/mantegna/mantegna.htm

13. From the anthology of Leonardo's writings collected by M. Kemp (editor), *Leonardo on Painting*, (New Haven, CT., U.S.A.: Yale Nota Bene, Yale University Press, 2001), p. 66.

14. Vertical disparity, although also the result of viewing the world with two eyes, is distinct from the more familiar horizontal disparity. See **binocular disparity** in the Glossary. In the experiment shown in Fig. 5, the stereo representations had the correct vertical disparity.

15. Strictly, there will be a very small degree of horizontal disparity unless the photograph is curved slightly to fall on the longitudinal horopter. See C.W. Tyler, "The Horopter and Binocular Fusion," in D. Regan (editor), *Vision and Visual Dysfunction, Volume 9 Binocular Vision* (Boca Raton, Florida: CRC Press, 1991) pp 38--74.

16. B. Harper, and R. Latto, "Cyclopean Vision, Size Estimation, and Presence in Orthostereoscopic Images," *Presence,* Vol. 10, No. 3, 312--330 (2001).

17. For example, Piero della Francesca wrote in *De Prospectiva Pingendi* (c.1474) "Painting is nothing but a representation of surfaces and solids foreshortened or enlarged, and put on the plane of the picture in accordance with the fashion in which the real objects seen by the eye appear on this plane." Quoted in I. Crofton, *A*

Dictionary of Art Quotations (London, U.K.: Routledge, 1988) pp. 143--144.

18. H. Ono, N.J. Wade and L. Lillakas, "The Pursuit of Leonardo's Constraint," *Perception*, Vol. 31, No. 1, 83--102 (2002).

19. See Kemp [9] pp. 170--171.

20. L. Wright, *Perspective in Perspective*, (London, U.K.: Routledge & Kegan Paul, 1983) p. 87.

21. R. Latto, "The Brain of the Beholder," in: R. Gregory, J. Harris, P. Heard, & D. Rose (editors), *The Artful Eye* (Oxford, U.K.: Oxford University Press, 1995) pp. 66--94.

22. It is a frequent assertion that television 'puts 10 pounds on your weight'. (See for example, L. Gunby, "Notes and Queries," *The Guardian G2,* p.16 (2000, September 14); L. Kelly, "Q and A," *Sunday Mirror Personal Magazine*, p.3 (1998, September 20); M. Warner, "Stealing Souls and Catching Shadows," *tate: the art magazine*, Vol. 6, 40--46 (1995, Summer). The common observation that noses look longer than expected when someone on television turns from full face to profile is another demonstration of the flattening and fattening effect of photography.

23. N. Wade, *Visual Allusions: Pictures of Perception.* (Hove, U.K.: Lawrence Erlbaum, 1990) gives a fuller account of some of these other differences between viewing pictures and viewing the real world.

24. R. Latto, "Do We Like What We See?" in G. Malcolm (editor), *Multidisciplinary Studies of Visual Representations and Interpretations* (Amsterdam, The Netherlands: Elsevier, 2004). pp. 343--356.

25. Jean Cocteau, *Le Secret Professionnel* (Paris, France: Librairie Stock, 1924).

Glossary

binocular disparity—the difference between the images in the two eyes. *Horizontal disparity* is the major binocular cue to depth and refers to the fact that the retinal image of objects not at the point of fixation will fall at different horizontal positions in the two eyes. *Vertical disparity* refers to differences in the vertical dimensions of the images in the two eyes. For example, when viewing a rectangle square on to the face, each eye will be viewing the rectangle at a slight angle. Consequently, the left eye will see a trapezoid with the left side longer than the right, while the right eye will see a trapezoid with the right side longer than the left.

image—the optical image falling on the retina, film, or photosensitive chip.

perception—the conscious experience resulting from the image and its representation.

representation—the coding of this image in the brain, photograph, or video system. This may vary at different stages of the process.

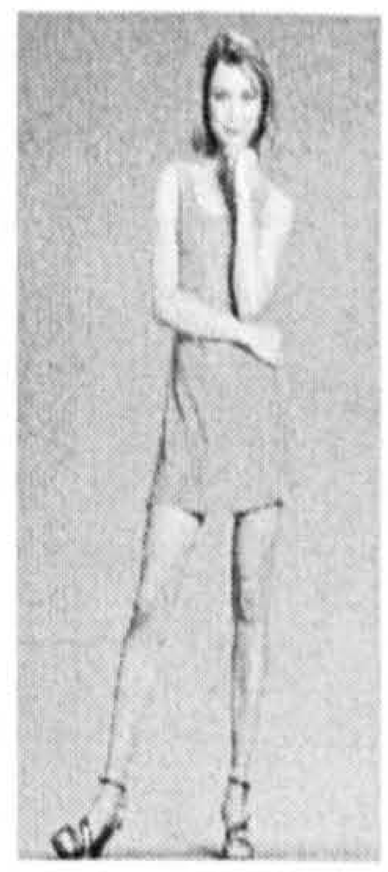
shape constancy—objects are perceived as having a consistent shape regardless of changes in the retinal image caused by different viewing positions.

size constancy—objects are perceived as having a constant size regardless of changes in the size of the retinal image caused by different viewing distances.

stereoscopic—the situation where the differences between the images in the two eyes generate depth information in the resulting perception (see also **binocular disparity**).

synoptic—the situation, for example when viewing a photograph, when the images in the two eyes are identical.

APPENDIX D



Thin, thinner, thinnest

No, this isn't really how Calista Flockhart looks. But how long before it is? And how long, Anita Chaudhuri asks, before somebody says enough is enough?

<u>The Guardian</u> Monday October 18, 1999

It hardly seems possible but Calista Flockhart, aka Ally McBeal, appears to have lost weight. The actress, whose body now looks to be constructed out of flesh-covered pipe-cleaners, has just had her holiday snaps published in the tabloids accompanied by headlines that purport to be concerned about her size. "Ally, the ocean waif" reads one. "Ally McMeal" quips another.

Cruel, yes, but celebrity weight-watching has become a spectator sport, fuelled by the ever shrinking waistlines of top female stars. In America, this week's People magazine carries a cover story posing the earnest question "How thin is too thin?" alongside gruesome photographs of Jennifer Aniston and Courtney Cox, Victoria Beckham and Helen Hunt, as well as Flockhart. "I swear I eat more now than I ever did in my life," Aniston says, not entirely convincingly. Liz Hurley was also recently lambasted in the press for her malnourished appearance. In this month's Elle, she talks openly about the self-imposed pressure for women in the public eye to lose weight. "If it's any consolation, I threw away two-thirds of my wardrobe and lost 15 pounds after first seeing paparazzi pictures of myself - the celebrity version of a vicious Polaroid."

But behind the headlines lurks a curious ambivalence. While anyone can see that these women are drastically underweight, it doesn't seem to occur to anyone closely involved with their careers that they might be putting their health at risk, never mind setting a poor example to young female fans. The photographs of Posh Spice are particularly alarming, showing her dressed in a chain mail halter top and tight black trousers, accessorised with jutting collar and hip bone. Clearly she has lost a lot of weight, but she is unimpressed at the suggestion that there's anything wrong. "It's irresponsible to say I'm anorexic. After a baby you are dashing about all day. I never sit down." This is not a common condition in new mothers but maybe being a famous new mother is different.

Flockhart is similarly defiant. "I don't think of myself as too thin. Am I anorexic? I guess my answer would have to be no." The actress, who once starred in TV docudrama The Secret Life of Mary-Margaret: Portrait of a Bulimic, is at pains to point out that she eats "whatever I want, whenever I want. I don't have a messed-up relationship with food." She describes her typical diet as a breakfast of egg whites and spinach, a chocolate chip cookie for lunch and chicken or sushi for dinner. To any woman who has ever worried about whether to save the second packet of chocolate HobNobs till after lunch, I'm sorry but her diet does sound seriously messed-up.

"It's not only irresponsible, it's deceitful," argues Deanne Jade, director of the National Centre for Eating Disorders. "These women talk about eating whatever they like and it makes young women feel like freaks because it doesn't seem to work like that for them. I go into classrooms and try to get the message through to teenagers that in a lot of cases these women are lying. They eat whatever they like, yes, then they go and throw up. Or else they take cocaine to keep their weight down. Or they're addicted to exercise and can afford personal trainers." It is ironic that these new photos of Flockhart appear just days after the death of Lena Zavaroni, the tragic manifestation of how fame can impact on a woman's self-image. To this we can add the other news that toddlers in Britain are officially "overweight", which can only mean that the pressure to diet is going to begin earlier and earlier.

"Half the problem with the girls I see is that not only do they have these skinny women in the public eye, but they have mothers who are on diets too. It's a powerful combination," Jade says. "For teenage girls, dieting becomes an endorsement of being female; talking about having to keep their weight down makes them feel like 'proper' women."

It is hard to see how any of these women, particularly Flockhart, can possibly be appealing role models, with their grey skin, lank hair and sad expressions. Not so long ago, Ally McBeal featured on a Time magazine cover alongside Gloria Steinem and Susan B Anthony, America's First Suffragette, under the headline "Is Feminism Dead?" Flockhart was not best pleased. "I was quite depressed. Ally McBeal is a woman who falls down, who throws her shoes. I'm afraid that's all I'll be known for." As her weight continues to shrink, we can count on her being remembered for something else entirely.

Thin end of the wedge

Today, the government's 'super-waif summit' will debate the causes of eating disorders in girls. Hadley Freeman, who was hospitalised for four years with anorexia as a teenager, reveals that while images of skinny celebrities did not cause her illness, they did delay her recovery

Hadley Freeman
Guardian

Wednesday June 21, 2000

In a way, Alexandra Shulman is absolutely right: you do not catch anorexia from Vogue magazine. At today's summit meeting, she and the other editors of fashion magazines can sit down, safe in the knowledge that the pages of their publications do not carry some kind of bacteria that will make their readers anorexic. Anorexia and bulimia are not trendy, zeitgeist illnesses. There are records of anorexics from as far back as the 16th century, so there is no point in trying to relate the cause of the illness to current pop-cultural issues.

No one would argue that alcoholism is caused by beer advertisements, and the concern that Trainspotting would lead to kids shooting up on the streets is now, at last, widely scorned. Nevertheless, lazy commentators still make the simplified assumption that too much Kate Moss is bad for your health. The truth is far more complicated. I know from painful experience - I suffered from severe anorexia through most of my teenage years, spending almost four of them in hospital. All because I, quite simply, stopped eating.

As studies routinely prove, the underlying causes of anorexia (as much as any psychologist can ascertain) are depression and childhood trauma. It can even be genetic. Far more complex, in other words, than the simple desire to fit into a size eight. Like alcoholism and drug addiction, eating disorders tear families apart. Parents, husbands, children, wives (an increasing number of men suffer from eating disorders) watch with a desperate helplessness as their loved ones slowly, seemingly wilfully, kill themselves. To suggest that all this misery is caused by a jealousy of Kate Moss or out of petulant vanity is insulting. So yes, the fashion industry can pat itself on the back: it is not from reading too many copies of a magazine that people decide to starve themselves to death. And yet. When models are becoming increasingly bony, when the ideal clothing size is diminishing every year, with actresses proudly showing off their hip bones and clavicles at this year's Academy Awards, when actresses and models who still need to wear a bra are called "curvy", there is something very unhealthy going on. To watch the current series of Ally McBeal on television is a particularly fascinating experience, with actress Calista Flockhart's tendons becoming more prominent each passing week. Yet she is being celebrated as an icon of feminine beauty and independence in at least two magazines this month. People in the fashion and entertainment industries are deliberately and belligerently missing the point when they storm that they don't "cause" anorexia.

An eating disorder is a mental illness. It is characterised by the sufferer's belief that they are too fat, that to survive on 500 calories a day is the norm, that doctors are trying to make them fat, that weighing more than seven stone is obese and unacceptable. So far, so paranoid. Yet the current culture of skinniness legitimises the anorexic's beliefs. That is where the danger lies. Once a person becomes severely anorexic, they are usually too locked into their own little world to care if Jennifer Aniston is now a size six, or to read about Jodie Kidd's protruding hip bones. But when they try to recover, it is very difficult to shake off these old beliefs when every other magazine cover seems to validate them.

Unquestionably, models and actresses are now expected to be thinner than ever before, and we are becoming inured to the situation. When these women become dangerously, bone-jutting skinny, they garner more publicity, more adulation, more success. This then becomes a vicious circle, with such images of skinniness being seen as the image of a Successful Woman. Calista Flockhart, Portia de Rossi, Lara Flynn Boyle, Sarah Michelle Gellar, Elizabeth Hurley, Courtney Cox and Jennifer Aniston were all, at most, moderately well-known actresses before their ribcages began to show. Now, particularly in the case of Flockhart and Aniston, they are to be seen marching down red carpets, clad in Versace dresses, beaming out from magazine covers, feted in high-fashion magazines which had previously ignored them. One US fashion magazine is said to have splashed out on a Marc Jacobs cashmere jumper to disguise what even it recognised as excessive skinniness during a cover shoot with Flockhart.

Thousands of models walk down the catwalks every year, yet it was Jodie Kidd who grabbed the press's attention when she strutted down, vertebrae visible to all. These women then become caught: the media focuses on them, with a mixture of fascination, adulation and voyeurism, so they are unable to put on any weight without making headlines. They are not to be attacked for being so thin; nor should they be blamed for making teenage girls push away uneaten plates of food, as Kate Moss was five years ago. They are simply manipulating (consciously or not) today's culture to their advantage, and the blame for this culture lies firmly on the heads of the media. Instead of pitying actresses and models who look pale, wan, ill, they are feted as symbols of wealthy, successful women. Incredibly, some fashion magazines have even made this a moral issue, claiming that they are defending a woman's right to be thin. Of course, no one should be criticised for their natural body shape. However, to look at some of the women in today's fashion and media industries is a painful experience. Everything about them, from their sunken eyes, thinning hair, bony arms and pigeon chests is a visual emblem of their self-denial and the misogyny propagated by this culture of thin. Women are not meant to look like women, they should resemble pre-pubescent boys.

My own anorexia kept me in hospital almost continuously for four years and I never imagined it was possible to be so unhappy. I weighed five stone and yet still exercised compulsively up to six hours a day. My day was built around how to avoid eating ("If I say I'm meeting someone for coffee, my mum won't make me eat breakfast"), how to exercise as much as possible (always take the stairs, never the lift) and just how to make it through the day. I was always cold - a cold that ripped through my increasingly visible bones, rather than one I could block out with my beloved baggy clothes. While my school friends, with whom I quickly lost touch, were going to parties and having their first snogs, I wasted these years crying on a hospital bed. I was always the youngest patient on the ward, and my closest friends were women and men 20 years older than me who had been ill for decades. At least three of them have died since my last admission. I took my GCSEs sitting cross-legged in a consulting room while a teacher from my school stood outside, asking the nurses to "keep the patients a bit quieter, please". This is not a part of my life of which I am particularly proud. However, I find it easier to talk about anorexia than to sit back quietly when people claim that it is entirely to do with models, or, conversely, that the current veneration of thinness is blameless. Both positions are simplistic. When I was in hospital, the other patients and I hardly ever talked about models. The doctors and my parents would try to assure me that all of my fears about fatness and food and my beliefs about the importance of thinness were the illness talking. Yet when I was finally discharged, I was amazed to find myself in a culture which seemed to disprove all of the doctor's assurances. These images had been irrelevant to me when I became ill, but, now that I was trying to recover, they fed into all of my thoughts.

I read articles in women's magazines that could have been written by me when I was locked on the ward at the Maudsley Hospital. There were models in magazines who looked like my former fellow patients. There was Elizabeth Hurley saying she would commit suicide were she as fat as Marilyn Monroe. I felt like the paranoid in a horror movie who realises that his worst fears are, in fact, real. There is something very wrong when society condones an attitude reminiscent of the beliefs I held when I was ill. For a while, I used this as an excuse not to get better. If models, actresses and other women didn't eat and obsessed about weight, then why should I have to change? Why were doctors picking on me? When I was finally discharged from hospital I weighed about six-and-a-half stone. Although most people commented on how thin I was, dozens of women used to do so with a tone of envy. I remember a group of women coming up to me in the street to congratulate me on my "sticky legs", counselling me not to put on any weight because "you can wear anything and look good when you're so thin". I went into a shop and cried. I finally began to recover when I realised that I could no longer bear being so unhappy, and accepted that this meant eating on a regular basis. I had to stop myself reading women's magazines, at least while I was putting on the weight, and stop desperately scanning magazine diet (or "health") articles and gazing at the models. When I look at these women now, I no longer feel the instinctive pang of envy, but rather one of pity. I remember, all too clearly, what it felt like to have that constant roar of hunger in my head and that feeling of being totally trapped and helpless. In their eyes, beneath the carefully applied eyeshadow and concealer, there is a familiar misery reflected back at the camera. The memory of that feeling girds me against ever letting myself slip back into the illness again.

Ironically, I now work in fashion journalism, which is why I feel even more strongly about this issue. Fashion does not need to present dangerously ill images of women. Fashion is about beautiful clothes, and fashion editors risk alienating their audiences and losing touch with any concept of reality with their needless promotion of the bony look. Models don't cause anorexia, but that does not mean that the fashion industry is off the hook and that we should maintain the status quo. We have to ask, why are women who weigh seven stone venerated as icons of beauty? The common reply from fashion magazines is that the sample sizes of clothing they receive are getting smaller, so they need to use models who can fit into a size eight or six. If this is the case, then they need to insist on sample sizes being bigger - this is not a difficult situation to resolve, and it needs to be resolved soon. Should we hang, draw and quarter magazine editors? Of course not. But we have to ask, do we really want to live in a culture in which the icons of beauty resemble patients in a hospital?

APPENDIX E

Virtual Stereoscopic Convergence: Projector Position Interpolation and production of varying interaxial stereo/synoptic images from fixed projector positions.
Phil Berridge, 02 December 2003

Initially images were created in the normal way. The shape was created in a 3d computer modelling package at full size. Virtual cameras were set up at 0, 30, 60, 120 and 240mm separations at a distance of 1680mm from the waist of the shape. The background was 1000mm behind the front of the shape's waist.

To enable distortion free projection of all images using slide projectors at 60mm separation a second stage was added to the rendering process. A flat plane, positioned where the front of the waist was located previously, replaced the peanut shape. A render created during stage one (above) was applied to the plane. This plane was rotated by a preset angle and the scene rendered using a virtual camera at the centre position. For example, the 120mm left image was rendered as normal in stage one. In stage two it was applied to the plane rotated by 1.02 degrees clockwise. The 60mm left image required no modification so was applied to the plane but the plane was not rotated.

The plane was rotated by the following angles for each of the stage two renders. In each case the corresponding stage one render was applied to the plane and the new renders were always generated from a virtual camera in the centre position.

L240	3.07 degrees clockwise
L120	1.02 degrees clockwise
L060	0.00 degrees clockwise
L030	0.51 degrees anti-clockwise
L000	1.02 degrees anti-clockwise
R000	1.02 degrees clockwise
R030	0.51 degrees clockwise
R060	0.00 degrees anti-clockwise
R120	1.02 degrees anti-clockwise
R240	3.07 degrees anti-clockwise

Mathematical description of geometry developed for the virtual projection technique.

Sin [angle from centre] = [separation / 2] / [distance from shape]

For example, 240mm separation images:
Sin X = [240 / 2] / [1680]
X = 4.09

Stage two rotation of plane:
[angle from centre – angle at 60mm]

For example, 240mm separation images stage two:
Sin Y = [60 / 2] / [1680]
Y = 1.02
Angle of rotation = X – Y = 3.07

APPENDIX F

Size Estimation Experiment

After filling in the top of the experiment form, please read this:

It has been noted that people can be remarkably consistent when making judgements about faces in photographs. Your first task in this simple experiment is to look at a series of twelve photographs and act as a visual "weighing scale." ***Please look carefully at each face, head and neck for the typical appearance of body fat*** **and sort the photos into three groups of slim, correct and overweight. Then, using the list of descriptors (above) put the photographs in order of thinnest to fattest (from left to right). Record the code for each image in order on experiment form.**

Please note that the size of faces in the photograph is irrelevant to the experiment.

Look at the list of descriptors again, and write the initials of the descriptor that most applies to each photograph in the box under the code for that photograph.

Your next task is to apply an actual weight value to each photograph in pounds or kilograms. A scale appropriate for the height of each person will be placed beneath their photograph to help you in deciding on an appropriate weight.

The photographs are of people from a very wide range of ages. Your next task is to put the photos in order of age (youngest on the left). Then estimate the actual age of each person in the photograph and write the figure in the box provided.

Your final task is to look at two monochrome photo sets (one male, one female) and repeat the size-ordering task as before. ***Remember to look carefully at each face, head and neck for the typical appearance of body fat*** **and sort the photos into three groups of slim, correct and overweight before putting them in size order.**

APPENDIX G

PHANTASCOPE
P.O. Box 3 Hayle, Cornwall, TR27 6YR – UK
www.phantascope.co.uk   info@phantascope.co.uk

The Cyclopter

CYCLOPS - From Greek *cyclos* (KÚKÂOς), "circle" + *ops*, "eye"

The Cyclops were mythical one-eyed giants of Homeric legend. Among other feats, they forged thunderbolts for Zeus. Their handiwork proved to be their own undoing, because Zeus is said to have used his thunderbolts to destroy them. In the middle of the 19th century Helmholtz used the term 'cyclopean eye' to allude to a hypothetical eye to explain identical binocular directions, combining both optical paths. About 100 years later in his 'Foundations of Cyclopean Perception' Bela Julez used the same term to indicate 'a central processing stage' of vision.

In 1907 Carl Zeiss patented a prismatic device that combined both optical paths. Calling it a 'Synopter', he marketed it as a visual aid, in order that the gallery-visiting public could appreciate depth in paintings. Being relieved of binocular conflicts and vergence by which the visual system fixes a surface and computes distance, the viewer was able to pass more easily through the picture plane and into the virtual space of a painting. In addition, a viewing device that eliminated peripheral vision allowed the various depth cues in paintings using formal perspective* to assert themselves more powerfully. Viewing paintings with two eyes partially suppresses the operation of some of these cues, allowing them to retain a degree of attachment to the picture plane. Because their paintings used systems of perspective, works by Piero della Francesca, Bellini, Uccello, Velazquez, Rembrandt and many others, respond well to this form of viewing. It is important to remember, however, that artists did not make paintings to create an array of visual data merely to produce a photographic or an objective truth, but an artistic truth. Not all depth cues should therefore be accepted at their face value. A shadow, for example, may have been intended by the artist as a moral statement.

It is possible that some spectators will have noticed that they could achieve an effect similar to that experienced through the viewer simply by looking at paintings and photographs with one eye closed – although it seems generally agreed that the effect is less strong, and the space less tangible than when seen through the viewer. Any device that superimposes the two optical paths, because it reduces disparities to zero, is impossible to calibrate for strength of effect, as one may with enhancement factors for Hyperscopes or Pseudoscopes. Considering the widely differing reports by users of a Cyclopter, it might be a useful and interesting exercise to calibrate the observers, and discover what personal factors determine the contribution different individuals make.

It is difficult now to know either what level of commercial success Carl Zeiss had with his viewer, or what impact or relevance it may or may not have had upon the appreciation of painting at the time. But even while Zeiss was obtaining his patent in 1907, the world of art was leaving behind notions that had for centuries challenged it to produce paintings that were mimetic or copied the natural appearance of the world.

The Cubists' revolutionary reorganisation of space denied the orthodoxy of perspective that had formed a visual referential frame since the Early Renaissance, and instead asserted an independent 'artistic' space.

While it is interesting to examine paintings with the Cyclopter, it is also instructive to see what changes it makes to viewing the real world. It has sometimes been referred to as Brobdingnagian vision, meaning to make things look bigger, and some users consider this description justified.

What is striking is the overall flattening effect, and the ability to resolve features involving depth, which binocular discrepancies hide or confuse. This unfamiliar clarity can be quite engaging if you are looking at a scene with intermediate partially transparent screens, such as through a window covered with raindrops, and you will notice that the visual feedback about your relative motion is different from your normal perception of movement. Painters who are involved in the discipline of life-painting, or portraiture, may find the removal of binocular conflicts both interesting and useful. Due to the use of both eyes there is a persistent stereo latency – an expectation of stereoscopic experience, though none materializes.

The prismatic instrument by Carl Zeiss (see Fig 1) used 5 hypotenuse prisms and one beamsplitting cube. It had a field of view of about 25°, and adaptations of it have been produced by various manufacturers for clinical applications, as well as microscopes and telescopes, for viewers who either need or prefer using both eyes. There is also a 4-mirror device, sometimes called an 'Orthoscope' or 'Monoscope' (see Fig 2) one of the mirrors being a beamsplitter, which also has a field of view limited to about 25°. The 2-mirror Cyclopter device shown below (Fig 3) has a field of view of about 30°, and allows far greater access to the cycloptic experience.

*Formal Perspective assumes a single vanishing point, and an eye that does not move. It is thought to have been invented by Brunelleschi in the early 15^{th} Century and refined by others during the High Renaissance.

TERRY POPE

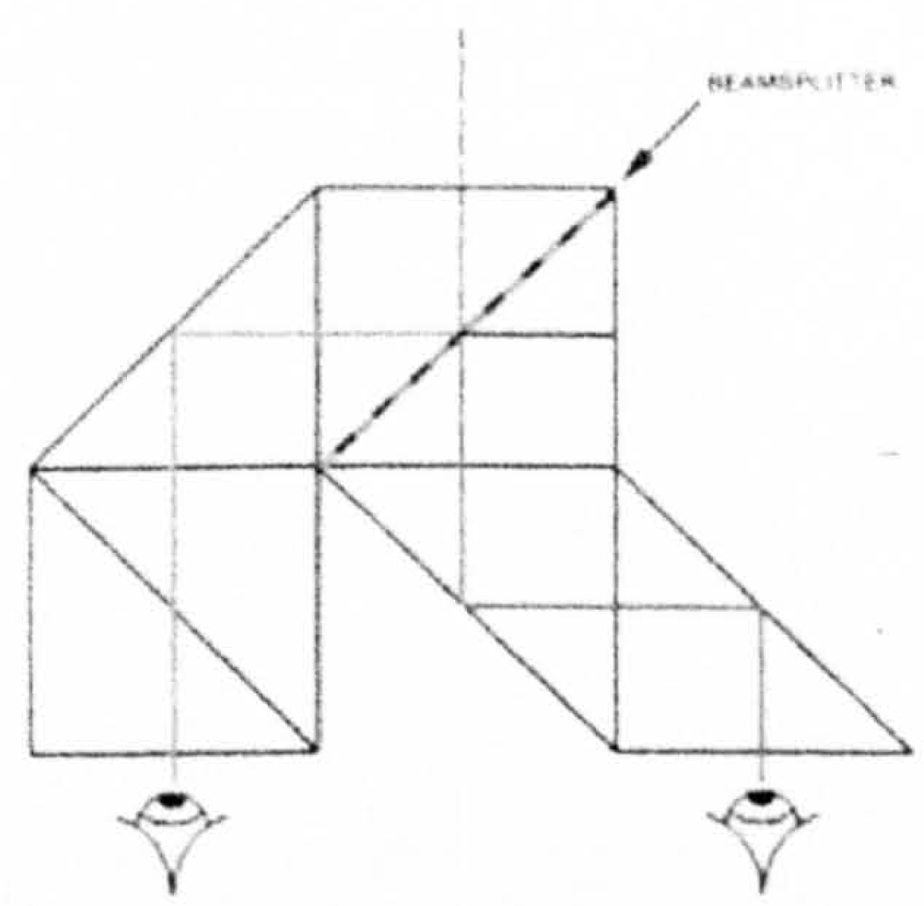


Fig 1

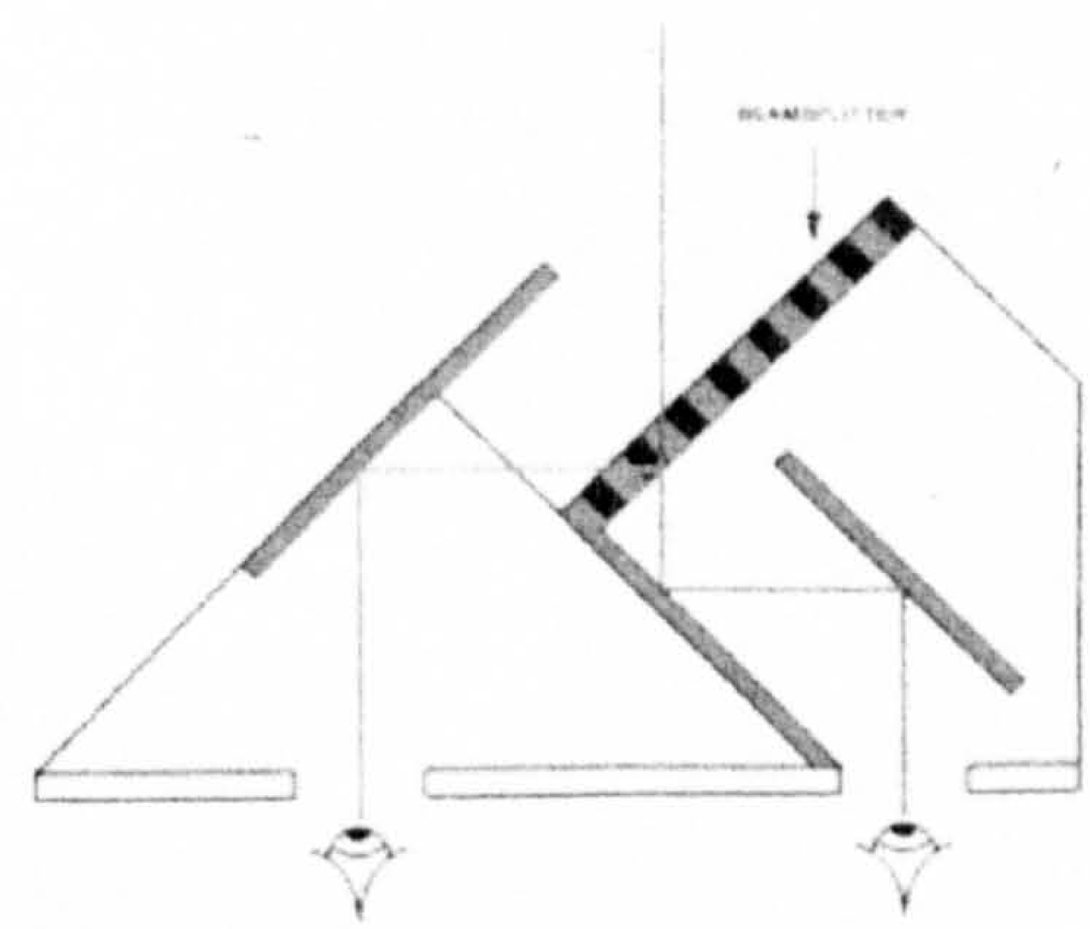


Fig 2

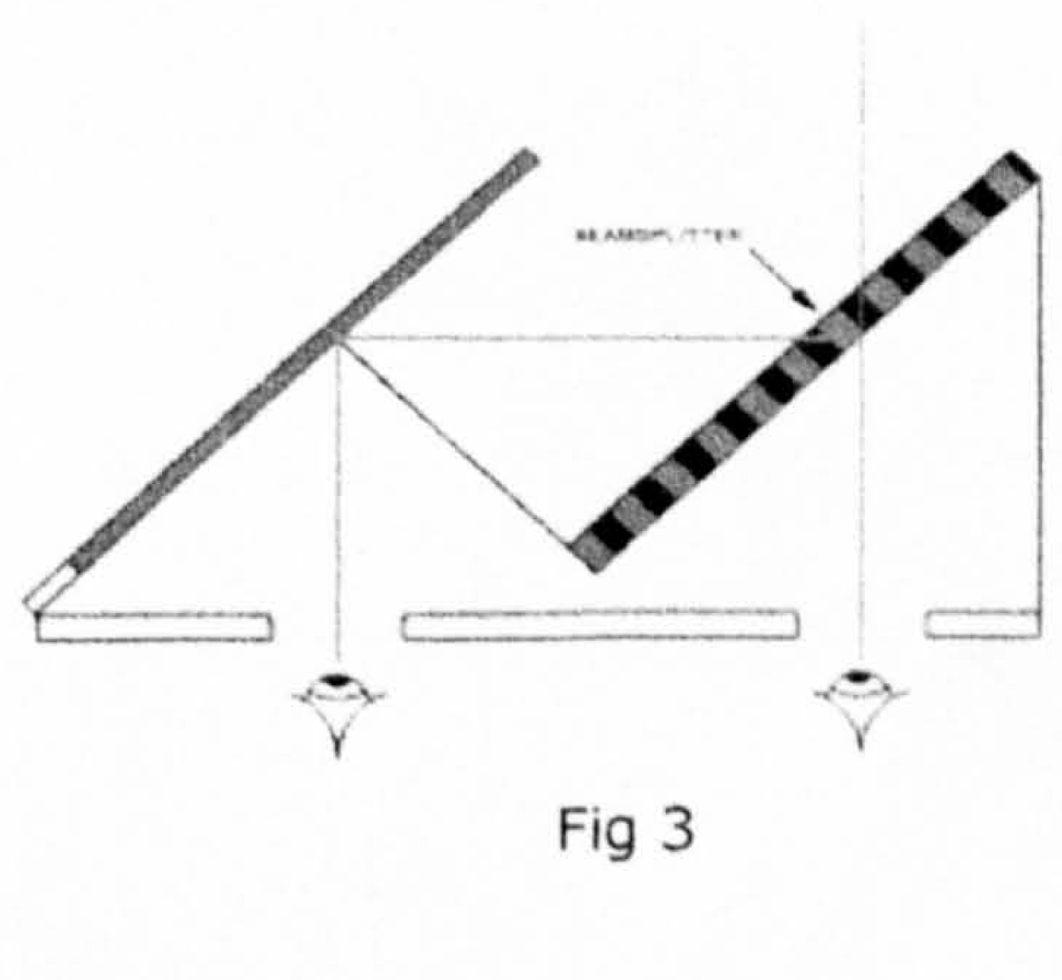


Fig 3

APPENDIX H

Appendix H

Pilot studies into perceptually distorted images and the limitations of Photographic Reality

H.1 Introduction

There are two primary hypotheses that underlie this research: first, a photographic image can be distorting because it lacks the volumetric (z-axis) depth information that we normally perceive with direct stereoscopic vision. Secondly, the magnification of a photographic image rarely matches the viewing distance of the eye when viewing the original scene, so it is not reproduced with life-like size, distance or magnification information. Thus, the original distance to the object is not conveyed and can undermine size constancy derived from a perceptual process combining perceived distance and visual angle. These factors seem to combine to form the characteristic flattening and fattening effect often observed in human portraiture and figure studies. This chapter describes the development of the methods used in the initial pilot studies, and how this experience was used to improve the methodology and procedures used in the subsequent psychophysical experiments.

H.2. Stereoscopic Body Image: First pilot study.

In order to investigate these theories, the first of a series of simple pilot studies was begun. The initial concept was to photograph people using full colour, high-resolution stereography and project their images life-size by polarized light projection. It was reasoned that if there were a fattening effect of 2D imaging, it would be simple to observe the effect by interrupting the image from one of the projectors tachistoscopically and judging if the image was changing shape in the transition from

2D to 3D conditions. Using readily available equipment, a simple 35mm stereo camera was improvised from a very basic schematic construction given in Michael Langford's 1989 Advanced Photography manual (fig 3.1).

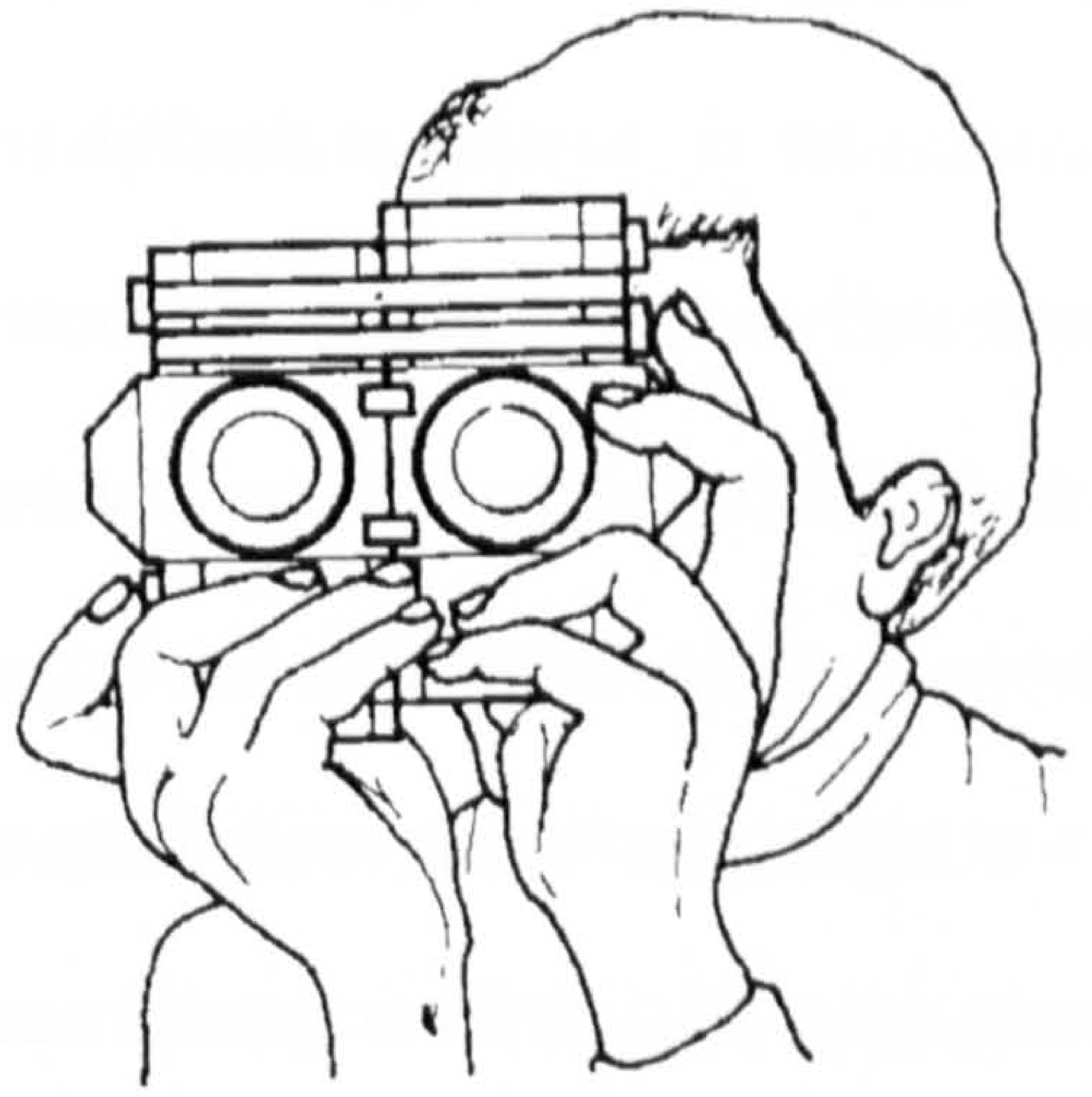


Figure .1. This figure is taken from Michael Langford's 1989 Advanced Photography manual (page 51) and was the basis for the first design of the stereo camera constructed for the first stereoscopic images in this study. Later versions of this design had an aluminium plate inserted between each camera that extended beyond the base and allowed them to be supported by a conventional photographic tripod.

H.2.1 Stereo Image Capture.

Two 35mm Pentax LX bodies, each fitted with 55mm f1.8 standard lenses were mated together at the base plate in the same manner as the diagram. These cameras were of professional specification and had hand operated film advance mechanisms with manual focus and exposure control. This produced an upright (portrait) format camera that was well matched to proportions of a freestanding model. The lens interaxial distance was measured using callipers as 64 mm. This seemed to be a good approximation of the average of 65mm human interocular distance that photographers most commonly use for making stereographs. The arrangement was then taped

together and elastic bands were used to ensure that should the tape fail during handheld photography, the cameras would still remain as a coupled unit.
A number of volunteers were asked to stand 1 metre in front of a background and the camera viewfinders were used to view them to see if stereoscopic photography was practical with this camera. It transpired that full-length stereoscopic figure study photography would be difficult to achieve, as the camera-to-subject distance for an average height male would be almost 3 metres. This is near the limits of useful stereovision and monocular cues may have priority over the available stereoscopic information. A closer position of 1.65 metres however allowed tall males and shorter females to be photographed from above head height to a crop line between the knee and thigh. This distance seemed to offer a good compromise between useful stereo information and close-up information of the important thigh, waist, arm, neck and face regions that convey the strongest information about body shape and fattening.

The arrangement was not without problems, as it restricted the variety of pose available to the volunteers. A particular difficulty was caused by the parallel lens axis of the stereo camera. It was clear from each cameras viewfinder that there were monocular only areas on either side of the subject that meant they could not stand face on to the camera with their hands on their hips, as their arms could not be imaged stereoscopically at the left or right edges of the picture. They would have to stand slightly side-on to the camera in order to be completely imaged in stereo. It was also noticed that viewing the model with the right eye through the left viewfinder, and the left eye looking at them directly, that they had different magnifications. It had been assumed that these 55mm "standard" lenses would produce life-size images in the viewfinder, but the images were slightly larger in the viewfinder than in reality.

A further problem was identified in the handling of the camera. The cameras tripod mounts were unusable as the cameras were mounted base to base. This meant they would have to be used hand-held, and flash photography would be required to ensure perfectly synchronised exposures in each camera after each shutter had been opened manually. So a decision was made to revise the stereo camera design. A metal base plate from a flashgun mount was used to attach both cameras to each other. This plate was eight millimetres thick and extended beyond the camera body to allow a tripod to be attached to it. The plate widened the interocular lens separation to 75 mm, but ensured that the camera could be safely supported by a tripod and accurately directed. It was noticed that if the camera bodies were converged so that each lens axis was aligned at the point of focus (1.68M), the lens centres now returned to approximately the interocular distance (measuring 66mm). This seemed like a good solution to the widening problem, as the human visual system (HVS) converges each lens axis automatically when viewing any close-up scene. Small wedges made from balsa wood were used to fix the convergence angle and the cameras were re-taped into their new position.

This type of improvised stereo camera is very difficult to use. Each camera has separate controls that need to be matched perfectly for each shot. Both 35mm cameras had no interconnection for their shutter release or electronic flash systems and so they had to be manually synchronised. Each were loaded with Kodak 100 ASA transparency film and then had to be strapped together with the elastic bands in order to insure that the cameras would not fall apart during use. The elastic bands made it impossible to load or unload the camera on the tripod as they held the film back down when firmly in place. Effectively, the system had to be assembled and disassembled every time one needed to load a film.

The test photographs were of staff members from Liverpool Universities Central Photographic Unit. It was felt that they might be more patient with the convoluted photographic process and be more understanding if a camera failure meant that they would need to be re-photographed. The stereo photographs were taken with electronic flash in a photographic studio that had a total daylight blackout system. This was essential because it was realised that long shutter opening times were to give the only reliable method of ensuring perfectly synchronised flash exposures on each film. If two flash systems (synchronised to each camera individually) had been used, slight differences to the manual shutter release time would give two time events on the subsequent stereograph and lighting that was laterally biased to each lens axis. This is something that would never happen with direct vision so it was decided that a single light would be used to illuminate the subject. Each shutter would be opened together by hand as the photographer held the cameras. The flash unit would be fired in the middle of their one-second exposure time by the model or an assistant, using a remote flash trigger. This would allow a single time event to be recorded in both cameras. The method was flawed in that total darkness was required to allow the cameras to be open for one second without environmental light fogging the film in addition to the flash exposure. The subjects had to be posed and focused in room light before each exposure was completed in total darkness. If they were to move slightly between posing and going to blackout and firing the flash, then the image would probably be unusable as it would be out of focus or badly composed.

Fifteen people were eventually photographed using the method described. A very high wastage rate due to technical failures meant that only ten people's images in one pose could be used at the next stage. No attempt was made to get them to wear clothes more appropriate to the study, and they were not posed in a systematic way. The resulting stereo pairs were mounted into 35mm transparency mounts and viewed using two GEPE slide viewers taped together to form a simple binocular viewer. These small-sized reproductions were impressive, but could not be considered to be as accurate as a life-size reproduction and, it was not obvious as to how the stereographs could be used in this form. So it was decided to improvise a stereoscopic projection system using another schematic diagram in the from the Advanced Photography manual (figure 3.2) so that the images could be seen at same-size (S/S) reproduction.

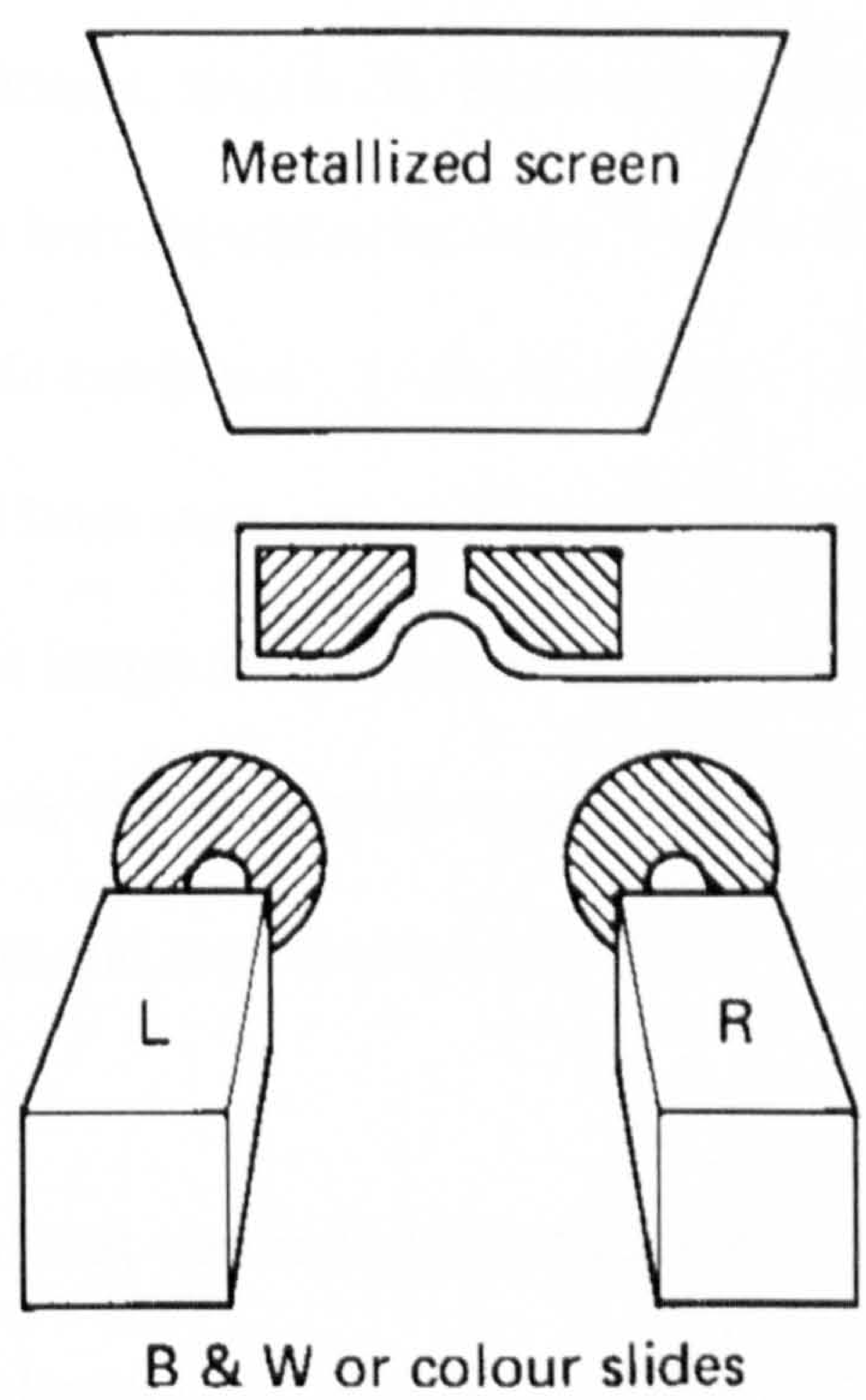


Figure 2 This figure is from Michael Langford's 1989 Advanced Photography manual (page 53) and shows how full colour stereoscopic projection can be achieved using readily available photographic equipment, such as 35mm projectors, linear polarizing filters and hi-brightness metalized projection screens. For viewers wearing a pair of dedicated linear polarized spectacles (or Polaroid sunglasses), correctly aligned and polarized 35mm left and right camera transparencies can be seen with full stereoscopic depth.

H.2.2.Stereo Image Projection.

The equipment required for full colour stereoscopic projection is simple in comparison to the stereo camera. Two identical 35 mm Kodak Carousel projectors fitted with 100mm lenses were placed side by side on a projector stand. Two large linear polarizing filters were then placed over each lens and each projector is directed to the centre of a large metalized projection screen. The projectors were switched on and the images were viewed (from an above centre position) while wearing standard 45^0 linear polarizing spectacles. These spectacles are like Polaroid sunglasses in that they have linear polarizing filters mounted at 45^0 to the vertical. One can align each filter on the projection lens precisely to that of the filters in the spectacles by rotation. When correctly aligned, the left projector image would appear dark to the right eye and bright to the left eye (and vice-versa for the right image). The separation between the lenses was almost 300mm, despite the fact that matching the convergence of the cameras would require a lens separation of only 110 mm. This was impossible to achieve with the available equipment, as the projectors could not be mounted any closer together. When 35mm stereo pairs are loaded into correctly-aligned projectors, a full colour stereoscopic image should be seen by viewers wearing the polarizing spectacles. This is because the left camera image is separated for the right image at the screen by polarization and seen exclusively by the left eye, and similarly for the right channel too.

It was considered that viewing comfort should have a high priority in the presentations. There are limits (Panum's fusional area) to how far out of horizontal or vertical alignment binocular stimuli can be before there is loss of fusion and diplopia or suppression of one image (Howard & Rogers, 2002). It was decided that as the point of focus for each camera would coincide with the convergence point for each camera lens, the lens axis of each projector should converge equally. The centre of interest would then be reproduced as a point of zero disparity in the display. This alignment was most likely to give comfortable viewing because when the points of each camera's focus are horizontally aligned in the display, the centre of interest (a face, for instance) appears as a single image. Zero separation in the display (no double image at the centre of interest) means that relatively flat objects can be viewed successfully without polarizing spectacles. Typically, this condition has a high degree of 2D compatibility as only the out of focus areas are not aligned at the screen. Polarising spectacles allow the viewer to separate these areas into discrete channels by which they can then perceive the original scene depth.

H.2.3 Stereoscopic Body Image demonstration.

The first successful stereoscopic test image (see fig.3.3) seemed to work very well on first inspection using the polarizing spectacles. After some horizontal and vertical adjustment to ensure that the images were correctly converged, a surprisingly pleasing 2D image could be observed without the polarizing glasses: the face and form were well represented and the overlapping stereoscopic information gave the impression of a sharp, minimal depth- of-field portrait, rather than the distressing doubled image usually seen when one views 3D movies without wearing the appropriate spectacles. Viewing the images with the polarizing glasses, the stereo image appeared to be bright, have high resolution and a remarkably natural appearance. The flat photographic background too appeared to be flat and approximately the correct distance away (1.5 metres behind the model). The model for this image was asked to view the picture and give her impressions. She thought that the image was "Very spooky. Like looking at a waxwork of yourself in Madame Tussauds."

Right camera image   Left Camera Image

Figure 3. A monochrome free-view stereo pair of the first stereo colour transparencies made for these studies. They are laid out in the free-view format with the right camera image placed to the left of the pair. This is so that they can be viewed using the cross-eyed technique whereby one holds up a finger approximately 50cm in front of the white border between the images. When ones eyes are focused on the finger, the images on the page should align behind it. Lowering the finger and focusing on the page should allow the viewer to fuse and focus the image into a stereo pair. If focusing is difficult, move closer to the image while retaining a fused image. This should happen at approximately 30cm from the images.

Despite the fact that the model was wearing winter clothes (allowing only her face and hands to be seen clearly) the stereoscopic image seemed to be slim and natural looking. Each 2D image (given when the light path to one projector was interrupted) was distinctly fatter in comparison. The other models too were invited back to view the set of images. All agreed that the stereographs were remarkably natural in their appearance and that the 2D versions were inadequate in comparison. The only model photographed in slim fitting clothes was particularly taken with her own image as the projectors were switched between 2D and 3D. She commented that it looked as if tummy is getting smaller and bigger, like as if she were breathing. When other people had their attention drawn to this effect, they too agreed that her waist appeared to get slightly smaller in the 3D condition. Her neck too appeared to become slightly slimmer. Her hips and face however, appeared to stay the same width. Remarkably, this suggested that the slimming effect was acting on the slimmest parts of the image only, yet no journal or photographic manual I could find made any mention of a slimming effect in stereoscopic images, or that the image might change its apparent proportions with the change of viewing conditions.

H.2.4 Stereoscopic Body Image Study.

It was decided to run a simple study with these stereographs to see if the fattening effect was observed by people who did not know the models and were not aware of the reason why the photographs were made, or had their attention drawn to the slimming effect. It seemed obvious that these participants would have to be of varying age, sex and have good general vision. A stereovision test for these participants was devised after it was realised that the first test image (fig. 3.3) had subtle shape information in the shoulder pads worn under the woollen jumper that could only be revealed binocularly. With a single projector, the shoulders appeared to be normally shaped and the background appeared to be immediately behind the model. So two simple questions should reveal if the participants have good stereovision: How far is the blue background behind the model? And can you see anything unusual about the shape of her left shoulder?

It was less obvious as to what questions should be asked to find out if indeed the stereographs were slimmer than the 2D projections. Eventually, it was realised that using a small number of body weight descriptors and asking the participants to apply one of them to each of the images as they are displayed in 2D or 3D should reveal if the condition changed between 2D and 3D presentations. Seven descriptors were decided upon and they were printed on a large card to be placed next to the projection screen. Each descriptor also had an acronym so that in the subsequent recording process, the participant would not have to write out the full descriptor. They were:

VERY OVERWEIGHT.	V.O.
OVERWEIGHT.	O.
SLIGHTLY OVERWEIGHT.	S.O.
CORRECT	C.
SLIGHTLY UNDERWEIGHT.	SU.
UNDERWEIGHT	U.
VERY UNDERWEIGHT	V.U.

Each of the twelve participants (drawn randomly from visitors) sat in a position below the centre line between both projectors (so that their heads would not enter the light path of each lens) behind a small table. In front of them was a small writing pad with the numbers 1-10 down the left hand side. They were told that they were being asked to look at each image and be a "visual weighing scale" and assess the bodyweight of the people as they were projected. During each presentation, they were to look at the seven descriptors and decide which one most closely approximated their estimation of the persons bodyweight and to record the acronym next to the correct numbers. The slides were shown in 2D first (using the right image projector) and the subjects did not wear the glasses. They then put on the spectacles and the stereographs were then shown in the same sequence.

The twelve observers all saw the stereo test image after the presentation and reported that the background was well behind the model and that her jumper had shoulder pads. This was a surprise as anecdotal evidence from a colleague at the Department of Orthoptics at Liverpool University had predicted that at least two people in a group this size would have poor stereo vision and might be unable to complete the task. Some of the responses were spoiled in that the participants could not write the response code correctly next to the slide number because it was too dark for them, they were not wearing their close-up spectacles or they did not understand the task. Those who completed the task were asked to look at their response sheets. All were surprised at the outcome, as they believed they had made the same quality judgements in both presentations. However, the responses showed that less than half of the judgements were the same in both conditions. Where was a change, it was slimmer estimation in the 3D condition. There were approximately nine times as many slimmer estimations than fatter in the 3D condition. Two respondents had score sheets that had no common estimations in 2D and 3D, and all of their judgements were shifted to a slimmer estimation in the 3D condition. These responses, coupled to the slimmer appearance of the stereographs suggested that there was a systematic slimming effect with stereoscopic portraiture and that the 2D condition was the more unnatural of the two conditions.

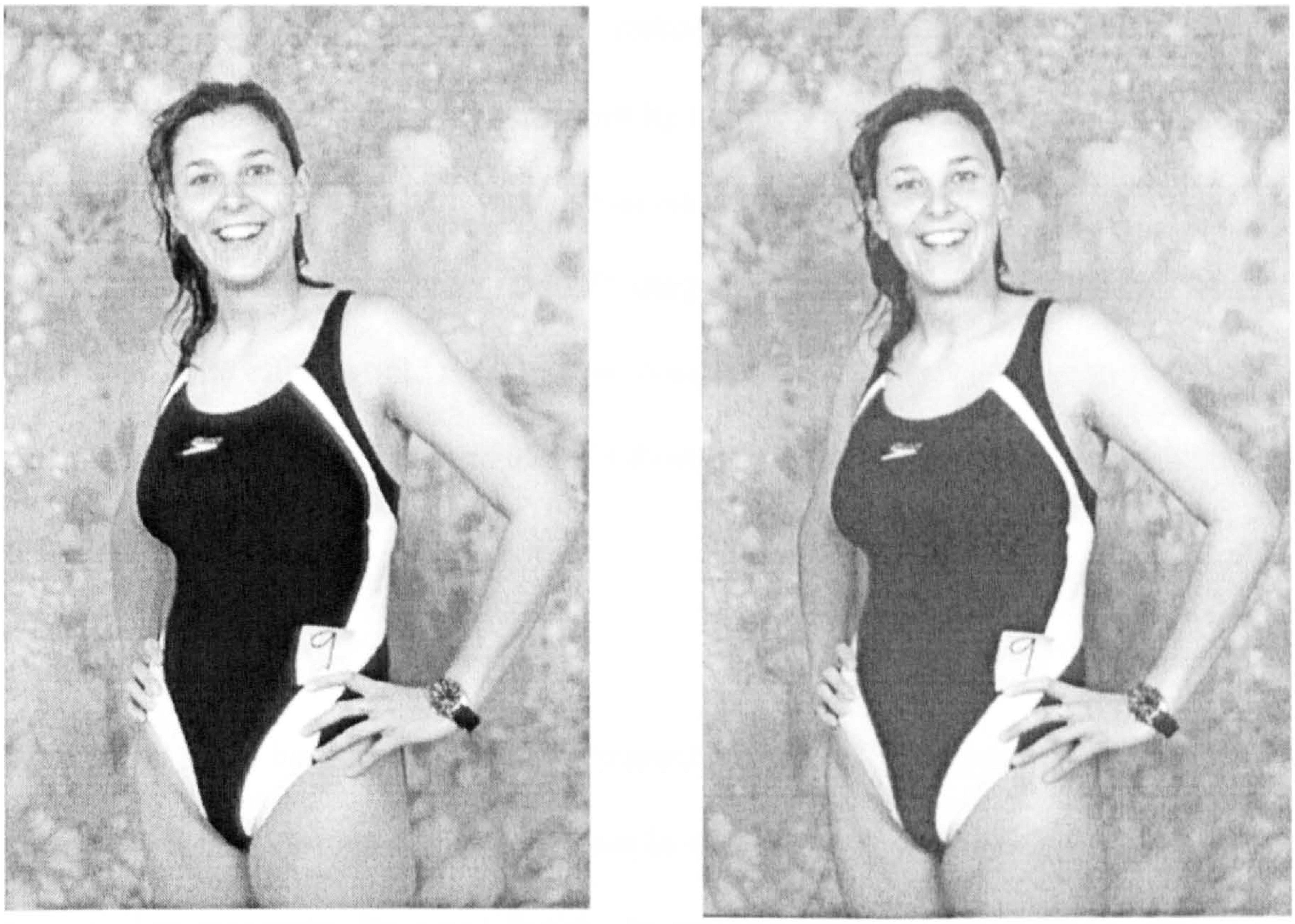
A number of guests from a photographic and psychology background were invited to view the images. They too saw the slimming effect as clearly as the other participants. Many of them sat through the experiment to help form their opinions on what might be causing the effect. They were asked if they were aware of any literature that could be helpful in explaining it, but no-one could not offer any sources for further investigation. Most were impressed by the slimming effect, but concluded that it was unlikely that this was a new perceptual observation, even though they could not think of any text or existing theories that could explain the effect. They speculated that the cause was some kind of conceptual, methodological or procedural error that was directing people's responses in towards slimmer estimations. Typical concerns expressed were that the sample size was too small, the glasses were only worn for the stereoscopic images, the models clothing tended to disguise their body shape, the slimming effect might be an artefact of the unnatural geometry of the cameras and projector, as they did not match the ideal specifications. Further suggestions were that there might be order effects on the data due to the presentation always being 2D first and 3D second.

H.3.1. Stereoscopic Body Image: Second pilot study.

It was decided that a larger study was required using the existing camera system in a more controlled way. As the slimming effect seemed to be more obvious on the female images from the first study, it seemed sensible to concentrate on them in the next. And as the main confounding factor for those images was the clothing, the suggestion was that women in swimming costumes would provide more consistent information. Also, it was realised that the camera itself was impractical for accurate image making in its existing configuration, as the posing had to be completed in complete darkness. A dual shutter-release cable was sourced that could fire the cameras simultaneously. This would allow one of the cameras to trigger a professional studio flash system that could be used in bright ambient lighting.

Professor David Brodie of the Department of Sports Science (who had proposed the original Body Image Distortion study) arranged for an area adjacent to the swimming pool at Liverpool University to be made available for photographing the new subjects. The same background was used and two studio flash units with umbrella reflectors were positioned on either side of the camera. This lighting gave a very diffuse quality of light without strong shadows that could give depth cues independent of the stereoscopic depth information. A higher resolution film (Kodak 64 ASA Transparency 36xp) film was used and the set-up approximated the original camera-to subject distance and magnification. A lighting stand was used as a target to check the convergences and focusing. Below its centre an X was marked on the floor to show the ideal place for the models to stand. The photography was timed to coincide with the women-only swimming session and a female assistant was used to approach the swimmers and ask for their participation. They were asked to write their names and contact details on a sheet were given a phone a number so that they could arrange to come and view the 3D images later if they wished to do so. Twenty-four volunteers were photographed, un-posed during the two-hour session. However, camera and flash misfires meant that only twenty-one of the stereo pairs were useable. The photographs were processed, mounted and displayed using the techniques described in section 3.1. A sample stereo pair from this sequence is reproduced in figure 3.4.

The models were invited to view the images and the five that returned agreed for the images to be used in any subsequent study. Fifteen participants, drawn from a group of colleagues and university students to view all the images. They were shown the stereo test image first (figure 3.3) to confirm they could see the depth information. Each observer saw the images individually and viewed the 2D and 3D presentations while wearing the polarizing spectacles. The presentations were balanced, alternating between seeing the images in 2D first for the odd numbered participants and 3D first for the even numbered participants. The same seven descriptors from the first study were used and the subjects responses to each image were self-timed and self-recorded.



Right camera image   Left Camera Image

Figure 4. This free-view image from the poolside photography used in the second pilot study. It shows a typical swimsuit style that appears to have been designed to give a slimmer appearance.

H.3.2. Second pilot study conclusions.

The responses confirmed the conclusions of the first pilot study. 300 separate estimations of bodyweight were made by all of the participants. Of these, 161 were the same in both conditions, only15 were fatter in 3D and 124 were slimmer in 3D. Eight of the fifteen participants made all of their revised choices in the expected trend direction (slimmer in 3D) and of the other seven participants, their averages of responses were all in the expected direction. Only one respondent (Anne) made more fatter than slimmer judgements of the 3D images. The pilot study suggested that there was a persistent perceptual effect at work, but could not confirm whether it was simply an artefact of the techniques used, or which type of image (2D or 3D) gave the most accurate representation. Some additional confounding issues that were not addressed in this study were the lack of control over the appearance of the models. Specifically, the swimsuits that were worn by the models were their own and not standardised in any way. There was also an issue of "human presence" in that it could be the slimming effect was due to the 3D image looking more lifelike than its 2D equivalent. The effect would therefore be more convincing if it could be shown with no human presence in the images and if a thorough statistical analysis was made of the data.

H.4. Peanut Shape Pilot Study.

It had been noted by the author that the slimming effect of the binocular disparity seemed to disproportionately affect the necks and waists of the models in the projected stereographs. So it was decided to produce stereographs of shapes that had waisted areas that did not look human in any way. It was hoped this lack of human presence in the image would reveal if the effect were genuinely dependent on perceived shape, rather than some kind of human bias. Initially, the plan was to make a human-sized cylinder and then slim the central area in some way. This object would then be photographed using the improvised stereo camera described in section 3.1.1. Viewers would then be shown the same shape stereoscopically on a large screen and given a hand-held 2D photo of the same object and asked to identify which was slimmer. The problem with this idea was that there was no obvious way of how to make the model, until a solution was discovered accidentally. [1]

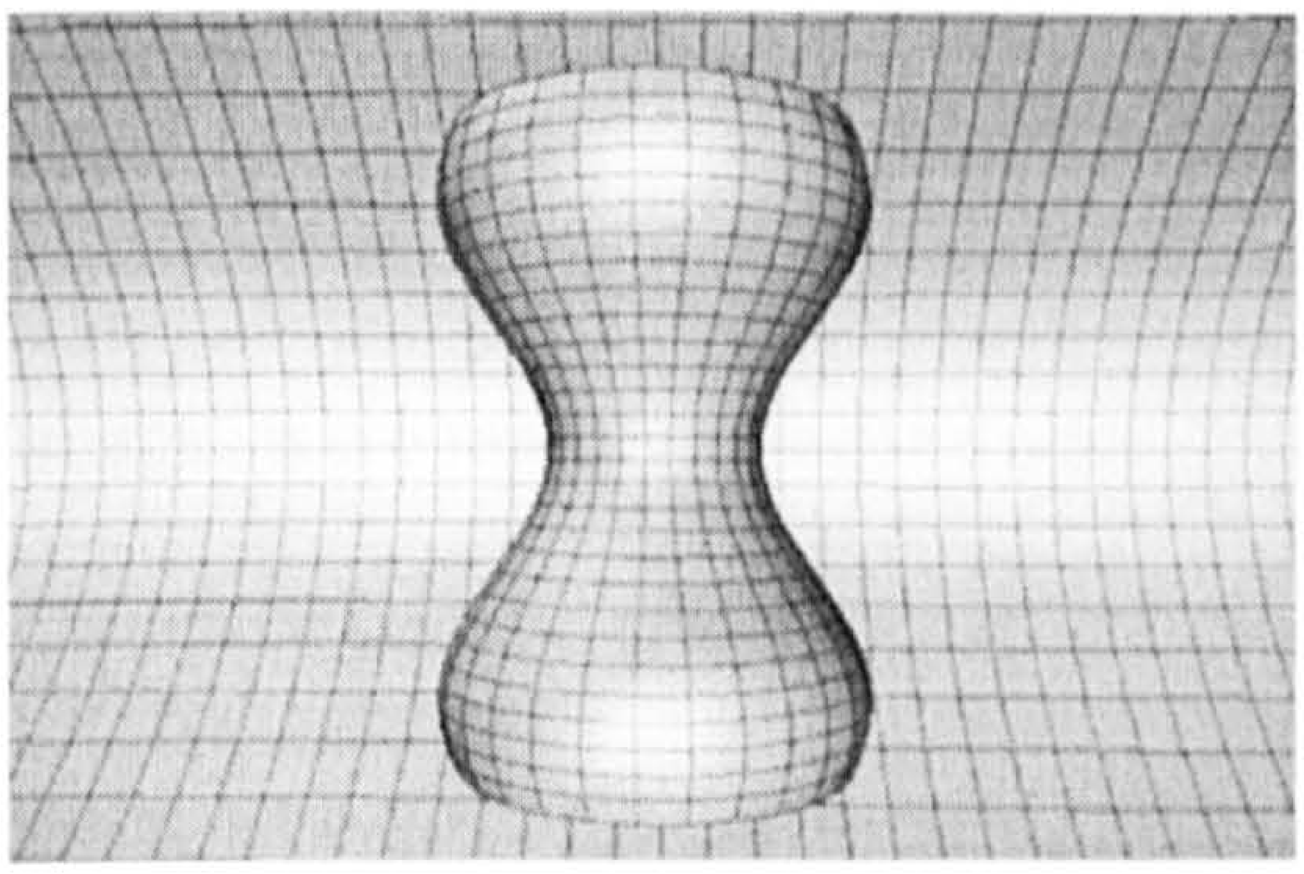
---

[1] Liverpool University Chemistry Department used their own method of 3D molecular modelling. Their version of the modelling program ChemDraw was unique, in that it showed the molecules with stereoscopic depth information. It did this using model rotation coupled to LCD field-sequential (shuttered) viewing glasses. The technique used was to model the shape in the normal way and to computationally rotate the model through 1.5-3.0^0 about its geometric centre. The two images were identical except for their rotation in relation to the viewer and were in effect a converged stereo pair. The images were then alternately flickered at high speed and the LCD glasses would allow the first images to be seen only by the left eye and the second only by the right. The images were then perceived from a conventional TV screen as a full depth stereo image.

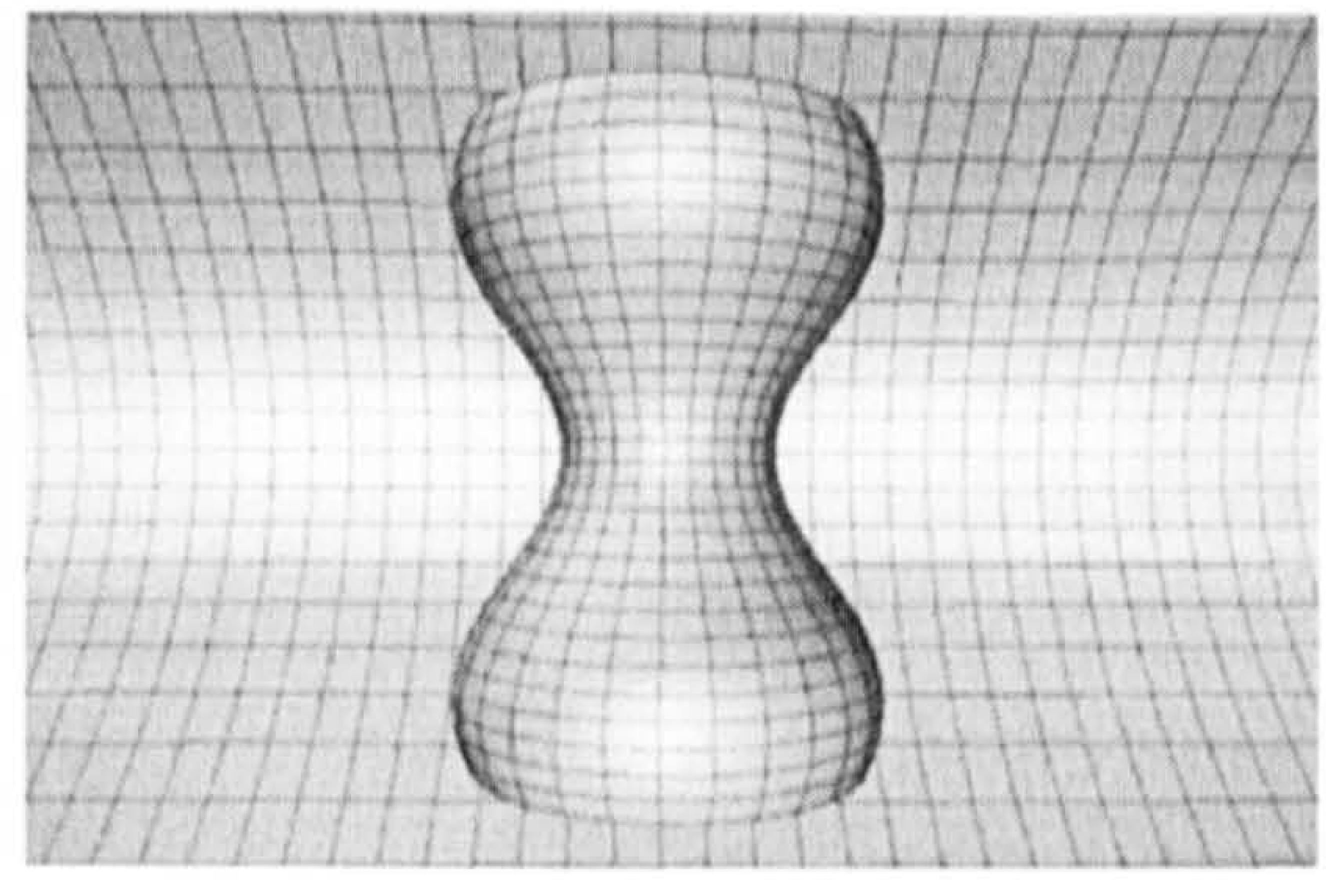
It was realised that the "waisted" cylinder stimuli could be mathematically modelled with the appropriate software. However, instead of using LCD glasses to view the images, each rotated image could be exported to a Lasergraphics 35mm colour laser printer and the resulting transparencies could be presented by the polarised stereo system previously described in section 3.1.2.

H.4.1. Peanut Shape Design.

The stimuli was designed by a mathematics post-graduate (Dr. Gordon Fletcher) using his own computer program that used "wire-frames" for generating 3D shapes. The model was given a shape that had a waist area that was 50% slimmer than the "shoulder and hip" regions. The software had a simple surface and lighting programs to give the finished model the appearance of a solid, illuminated object. He was asked to render six images in all: the first two images were of the "cylinder" from a straight-on, central viewpoint. These two images would be loaded into the left and right projectors to produce a 2D image without requiring the viewers to remove their polarized glasses. The two other images were 0.75^0 left and right (to make a 1.5^0 converged stereo pair) and a 1.5^0 left and right image (to make a 3^0 converged stereo pair). All of the cylinders had the same wire-frame background that maintained a consistent geometric relationship to the cylinder in all the images. The centre of rotation for the image was set to revolve the centre of the cylinder. It was realised that the cylinders required a more accurate descriptor for their shape, and the designer suggested that they should be labelled as "peanut" shapes (fig. 3.5).



Right CGI Virtual Camera image



Left CGI Virtual Camera image

Figure 5 The stereoscopic computer generated peanut shaped stimuli.

H.4.2 Peanut Shape Study.

The peanut shapes were projected stereoscopically using the method described in section 3.1.1 to viewers after they had correctly responded to the stereo test image. Each participant was given an A4 print of the stimuli that was made from the central viewpoint. The print was enlarged so that when held at arms length, it appeared to be a similar size to the projected 2D image. The five participants were then shown a sequence of the 2D pair (central viewpoint image), the 1.5^0 stereo pair and the 3.0^0 stereo pair on three occasions, making a total of nine presentations. As each new pair was presented, the participants were asked to compare it to the print of the image held at arms length and report whether the image on the screen was fatter, slimmer or the same shape. In total, 45 individual responses were recorded to the images. On only two occasions out of 15 did a participant fail to match the monocular shape correctly. The 1.5^0 shape was judged as the same shape in seven of the responses and slimmer by eight. The 3.0^0 stereo pair was judged as being slimmer in thirteen responses and the same by two. None of the stereo pairs were judged as being fatter than the printed image on the card, and this fact supported the subjective observation made by everyone who saw these test images that the object was slimmer in the 3D conditions. In the post test debriefing, the observers were all surprised when told that the image did not change shape at any time, (in that the dimensions of the projected shapes were identical) and only the amount of stereo information was varied. It was only when the images were rapidly switched between the 2D and 3D conditions was it clear to the observers that the effect was real and that the object did appear to have a slimmer waist area when the depth information was added.

H.4.3. Peanut Shape Study: Conclusions.

The simple descriptive statistics used supported observer comments that the slimming effect was stronger in the peanut images than they had seen in the previous stereoscopic images of people. Despite this powerful effect, there was still the possibility that there may be an underlying artefact driving the participants judgements to be slimmer in the stereo conditions. One possibility was that the participants could have been "square counting." This was because the wire-frame background behind the peanut shape showed different patterns of occlusion by the waist of the peanut, depending upon which image was being observed. The stereo images presented more background information because one could see and count the wire-frame squares in these images that cannot be seen from the central viewpoint 2D images. It was possible that if a background had been chosen that described a flat plane without the users being able to count squares that appear in some images and not others, then the slimming effect would be shown without "monocular cues." These are cues that could possibly allow someone using one eye only, or suppressing the view from an eye to complete the task in a similar way to someone using normal stereovision. Although possible, this seemed unlikely as the images were viewed quickly and without much contemplation from the subjects. More significantly however, it was also noted that an artefact in the rendering software gave a different stereoscopic position for the surface of the object than the wire-frame stereo coordinates. The artefact could not be seen in the 2D conditions the software was designed for, but was evident in the stereoscopic versions of the images. This meant that the object had two stereoscopic shapes: One for the wire-frame surface and another smaller one for the lit surface. Perhaps this smaller stereo image was what people were seeing when they made the slimmer judgements. A further confound was that the wire frame image had no "real world" dimensions. Therefore it could not be built from raw data to repeat the demonstration with a life-size object made to the same dimensions and photographed by the stereo camera.

All attempts to establish the "real size" of the computer generated peanut failed to find useful real-world information that could allow the peanut to be built from scratch. It was also not possible to establish the virtual camera to subject distance in the images so that the focal length of the virtual lens could be calculated. This was because the program was designed as a mathematical visualization aid and not as a CAD CAM (computer aided design, computer aided manufacture) system. These software packages use real world measurements to produce "solid" objects using surface rendering algorithms coupled to virtual cameras with virtual lighting. This type of software can convey very accurate renderings of objects that could be built to an almost identical appearance. If the peanut had been built with an architectural CAD CAM program, then it could have been accurately reconstructed using real materials from the design instructions found in the program.

H.5. Shape perception and lens design.

The inability of the peanut software program to reveal the virtual camera to subject distances, and therefore the angle of view of the virtual lens was disappointing. This was because the angle of view of photographic lenses is known to have effects on the appearance of shape and relative size. These distortions are perceptual, in that almost all modern lenses are considered to be distortion free in normal use (Langford Advanced Photography p.24-25). Photographic lenses can be considered as simple meniscus (magnifying glass) lenses in that their convex shape causes light to bend in a way that creates an enlarged virtual image in close-up or a reduced and inverted "real image" at its focal plane. In the early days of photography, it was found that these simple lenses had a number of optical aberrations that caused visible distortions in the proportions of the image. So-called pincushion distortions can make the lines of a square appear to bend in towards its centre. Barrel distortion has the opposite effect and will make the straight lines of a square bend outwards, away from the centre. These curvilinear distortions can clearly be seen on photographic images made using simple lenses, yet are almost immeasurable with the modern professional prime lenses that were used throughout these experiments. Of more concern, however are the unnatural perspective distortions of wide-angle and telephoto lenses that are optically correct, but can distort ones perception of size, shape or distance.

Photographers use a number of different types of lenses in the task of image making. Of particular interest to professional photographers is the concept of the "standard" lens (Langford Basic Photography 1983 p76). This lens in its simplest terms is defined as neither magnifying the image, like a long-focus telephoto lens or miniaturising it like a wide-angle lens. The working definition of a standard lens is that its focal length should approximate the diagonal measurement diameter of the film format. In 35mm cameras, the film the diameter is 44mm and the standard lens is considered to be 50-55mm. The figure of 50mm was arrived at for practical reasons, as it allowed the lens to be far enough away from the film plane for a large reflex mirror to sweep a clear path inside the body and provide a large, bright image. When fitted with wide-angle lenses, these cameras had to lock the mirror up before the lens was attached. Only later, with the arrival of expensive retro-focus lenses were the rear elements of wide-angle lenses far enough away from the film plane to allow the mirror to be used. For the larger 6x6 cm film format, 80mm is considered to be the typical focal length for a standard lens. This figure more accurately matches its measured diagonal of 84mm and is considered to give a more appropriate angle of view for a standard lens. As far as the author is aware, no major 35mm format manufacturer has ever produced a 45mm standard lens that could have been used in these experiments.

Wide-angle and telephoto lenses have very different focal lengths and angles of view: In 35mm photography, wide-angle lenses are typically in the 8mm to 50mm focal length range and have angles of view from 180^0 (fisheye lens) to 46^0 angle of view of a 50mm standard lens. Telephoto lenses are considered to be all focal lengths above 55mm. A typical telephoto lens would use a 135 mm focal length that had a horizontal angle of view of 18^0. Super-telephoto lenses have a horizontal angle of view of 6.9°at 300mm, up to 2.6°at 800mm. Focal lengths beyond these ranges are very rarely used by photographers and are considered to be "astronomical" in terms of their likely usage and indeed cost of purchase.

If a 35mm camera (fitted with a standard lens) is held vertically to ones left eye, one should see through the viewfinder that the image in the right eye is very similar in both magnification and perspective. Background and foreground objects should also have the same size and distance relationships in the viewfinder as can be seen with the right eye directly. It is easy to fuse these images stereoscopically if the cameras lens axis is directed to the gaze point of the right eye, even though the viewfinder image will be darker and have a lower resolution. However, although the camera image is the same size as reality, it has a much narrower field of view (FOV). The right eye will have a vertical and horizontal FOV of about 150^0 (180^0 binocularly) but through the camera this will be reduced to 46^0 vertically and 30^0 along the horizontal. Cameras can capture wide-angle images of over 150^0, but they cannot do so while reproducing the same magnification as the human visual system (HVS). Thus the standard lens offers a good monocular approximation of the HVS, but as if viewing the scene through blinkers that obscure the full FOV. Stereo photography using standard lenses is therefore geometrically similar to direct human perception, but at the expense of a natural FOV.

H.5.1. Practical considerations in lens use.

This thesis investigates the subjective impressions one gets when viewing photographs of people and objects that they differ in important respects from their appearance in reality. While the lack of stereoscopic information in a photograph is an obvious area of perceptual difference, a more subtle effect could be found in the use of wide-angle and telephoto lenses. Photographers rarely use wide-angle lenses for portraiture as they have a number of characteristics that make them unsuitable for this purpose (see Langford 1983 p.75). Of particular concern is the problem of subject proximity, whereby the camera has to be very close to the subject in order for it to fill a satisfactory area of the available field of view. While this is fine for inanimate objects, people tend to feel threatened or uncomfortable if a camera lens is too close to their faces. Also, very wide-angle or fish-eye lenses are known to have curvilinear distortions that can give objects an appearance of excessive curvature. This is why many photographers believe that wide-angle lenses make people look fat, even though there is no data to confirm this. So photographers tend to use telephoto lenses that allow them to be much further away from the subject. They also impart other technical benefits in that the background behind the subject is smaller (and easier to light) and can be made to appear out of focus when using the largest lens apertures. This effect of "differential focus" using defocused foregrounds and backgrounds is considered to be an important element of conventional portraiture techniques. Further benefits can be found with extended camera to subject distances in that the camera is too far away to occlude any lighting angles and camera noise (especially with moving image cameras) will not be picked up by microphones used to record the subjects voice.

A typical portrait lens used in professional 35mm photography would be from 85-100mm in focal length. However, these lenses are considered to be telephoto when compared to the 50mm standard lens that most closely approximates the human angle of view and magnification. Yet telephoto lenses are known to compress perspective to give a pronounced flattening effect to the perspective in any scene. For instance, if one wished to make the rising moon appear to be very large in relation to a foreground object or person, one would use a telephoto lens. This technique will also make the foreground object appear to very close to the background. If one wished to make the moon seem smaller in relation to the foreground, one would use a wide-angle lens. This would have the effect of making the moon appear to be relatively smaller, and with a fisheye lens appear to be a star-like point source. These simple effects of manipulating perspective to alter the appearance of natural size constancy are normal everyday techniques used by experienced photographers. They are especially appealing in that most of these effects are cryptic to the casual viewer or professional photographer, as it is often very difficult to judge exactly what type of lens has been used to take any conventional photograph.

With portraiture, certain questions arise from the use of different types of lenses in terms of recording the subject accurately. An aspect of telephoto size constancy noted by Cutting (2004) was that objects dimensions measured from the print or screen appear to be "close to their actual sizes." It may be that many photographers have thought the flattened "plan view" image provided by telephoto lenses is therefore a more accurate representation of the scene to be photographed. Yet this interpretation does not take into account the most important effect of magnification by telephoto lenses, which is that they can show details that cannot be seen with direct vision. This ability to enlarge distant detail is completely unnatural when compared to normal human vision. It also allows us to see compressed distance perspectives that alter the appearance of natural size constancy in ways the wide-angle HVS cannot normally see. Yet lenses that have an angle of view as wide as the HVS, such as fish-eye lenses, often have curvilinear distortions that look very unnatural. Within these extremes are a range of focal lengths used by professional photographers that are either free of curvilinear distortions or capture distant detail with similar resolution to what we can see from the camera position. Typically, this range is from 20mm super wide-angle to 300mm telephoto.

H.5.2 Vari-Focal pilot study.

An important question investigated by this study is which of these lenses is the most accurate for portraiture, when the subject size is kept constant in the viewfinder as the camera is moved closer or further away from the subject. A simple experiment was devised to see if people appeared to have different apparent bodyweights in portraits taken with lenses of varying focal lengths. To do this, eleven professional quality lenses were borrowed that represent the most common focal lengths used in professional photography. These were of 19, 24, 28, 35, 50, 80, 100, 135, 200, 300 & 400mm. focal lengths. The closest of these photographs was taken at a distance of only 340mm from the models faces. The furthest was taken from approximately 5.4 metres away.

A very large white photographic background (3x3 metres) and two Multiblitz studio flash heads were set up in a 15 metre long room that had full blackout curtains and screens. A centre-line running the length of the room was marked on the floor using masking tape. The lights were set up 2 metres from the white background equidistant to the centre line. Each light was fitted with a diffusing umbrella and directed at a chair placed on the centre line, one metre from the white background. An Olympus OM1 camera was mounted on to a tripod placed directly over the centre line and fitted with a 19mm lens. A male and female model of average heights and weights agreed to be photographed using this equipment. Their task was to retain the same pose and expression for each of the 11 photographs. They were helped in this task by an assistant.

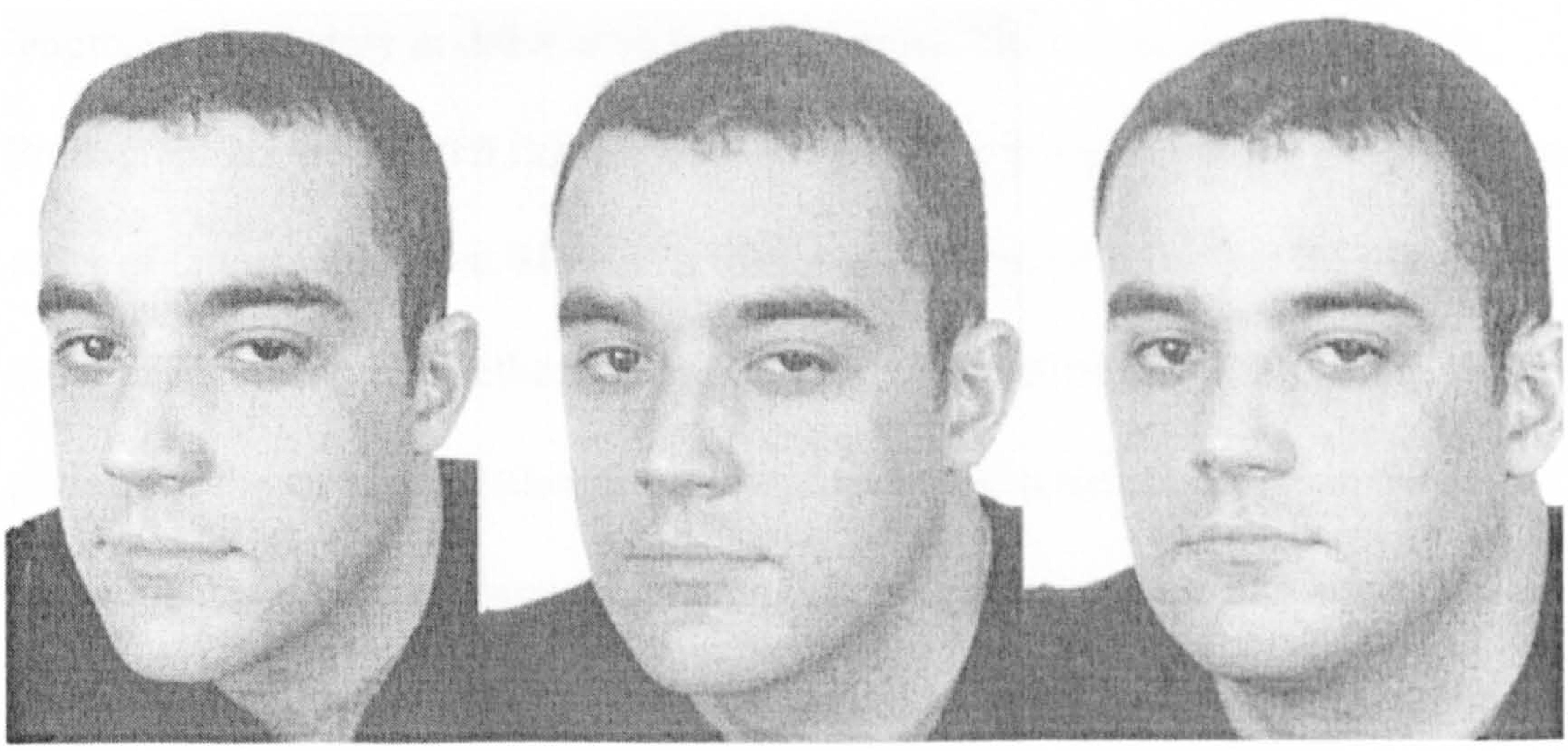
Using a video camera mounted on a tripod and linked to a TV screen, it was possible for the assistant to mark their head position and eye centres on the screen with wipe-board pens. This ensured that when they changed places, it was comparatively easy to ensure their pose and expression was consistent in each image. The pose was face-on towards the video camera, but with their gaze directed to the still camera. Both models were photographed with each of the lenses from wide-angle to extreme telephoto in increasing focal length. As each lens was changed, the camera was moved further away. The correct distance for each shot was established by using marks on the viewfinders focusing screen that matched the required eye separation. This was an approximate guide, as correct sizing of the prints would be achieved in the darkroom by slightly varying enlargement factor of each print. Figures 3.6a and 3.6b on page 27 represent the range of photographs taken for both of the two models.

Figure 6a

19mm wide-angle lens. 50mm standard lens 400mm telephoto lens

Figures 6a . The three photographs above and below are of the female (a) and male (b) model from a sequence of eleven photographs taken with a range of professional quality photographic lenses. These were of 19, 24, 28, 35, 50, 80, 100, 135, 200, 300 & 400mm. focal lengths. The closest image was made from a distance of approximately 300mm, and the furthest 5.38 metres away. The complete set are reproduced in section 6.1.1.

Figure 6b



19mm wide-angle lens. 50mm standard lens 400mm telephoto lens

H.5.3. Vari-Focal images: presentation procedure.

Twenty-two 5x3.5 inch prints were made from each of the best eleven negatives of the two models. The prints were marked on the back with the type of lens used and their number in the sequence from 1-11. The prints were separated into male/female and randomised. Ten participants were asked to look at the prints and place them in order from thinnest to fattest for the male and female sets. Remarkably, almost all of the participants found this task very easy to do and there were very few transpositions that were out of sequence. The appearance of gradual increases in apparent body size from slim to fat was so profoundly obvious that it was realised the task was too simple: The effect of changes in body size when one can compare images of the same person under different conditions (figs. 3.7) was so obvious that the only other question that was asked of most of the people who viewed the images was to indicate which print looked the most natural. Most people said that print four or five looked the most natural. These images were from the slightly wide-angle (35mm) and standard (50mm) lenses and confirmed the authors own aesthetic preferences. An interesting conclusion to this study was found when the photographs were shown to local professional photographers. They were asked to put the pictures in order of focal length, or alternately in order of thinnest to fattest. The results were surprising. Photographers were better than most of the previous group at putting the photos in order of fattest to thinnest. However, they were very poor at putting the photos in order of focal length, whether or not they did this task first or second. They also re-performed the task with little improvement, even when the ideas behind the experiment had been explained to them. This suggests that most photographers are probably untrained, uninterested or incapable of identifying the approximate focal length of a lens used in simple portraiture (from cues in individual photographs).

However, they are acutely aware of the appearance of body fat. It may be that this awareness underlies their apparent preference for photographing very slim models.

The images is this study were not without problems. The ¾ pose used for the models was arrived at because it was decided that it would be more useful if the images looked as much like conventional portraits as possible. However, this meant that in the widest-angle lenses there was a slight distortion of the faces due to the proximity of the lens that some viewers found to be disturbing. Additionally, it was realised that if we wanted to measure the proportions of the faces and the way these vary with different lenses, then the pose would have to be face-on to the camera.

A final problem was that the camera to model measurements were incorrectly measured from the tripod centre to the chair centre. This was unfortunate as it was realised that the important variable in this experiment was the distance from the rear nodal point of the lens (its optical centre) and the point of focus, which was the eyes of the models. This is important because camera lenses of differing designs can be described as being of a specific focal length (e.g. 50mm) and yet have slightly different angles of view. So a photograph from the same position using two different types of 50mm lens could give slightly different degrees of magnification. If one were to use viewfinder markings to scale the image from two different lenses to the same apparent magnification, one would inevitably have to move the camera to do so. This would then render two different perspectives with each lens from what would appear to be the same magnification on the same focal length. *As different lens designs will always record the same perspective from the same viewpoint, (regardless of the type of lens design or focal length), camera to subject distance and not focal length would be the important variable in any subsequent experiment.*

H.6. Conclusions of the Pilot Studies.

Each of the four pilot studies demonstrated that the reported fattening effects of 2D photography were reproducible in monoscopic/stereoscopic comparisons, and with 2D images where the perspective and camera to subject distance were varied as the magnification remained constant. The broad conclusion to be investigated further in this thesis was that *whenever photographs are made that diverge from the natural conditions of human stereoscopic vision, consistent and predictable body image distortions are likely to occur.* However the four studies reported could not be described as experiments, as typically they used improvised methodology (without strict repeatability or consistent data recording) and used only the most basic of descriptive statistics to demonstrate the strength of the effects. The participants who viewed the stereo images did so without any accurate measurement being made of their stereovision, so there was some doubt whether all of the participants actually had normal stereoscopic perception. Also, no extensive background research into human perception was made to ascertain if any of the effects observed were artefacts of the methods used or could be ascribed to existing perceptual theories. Additionally, it was noted that the models used in almost all of the photographs had not signed release forms allowing their images to be used in any future publications. These procedural, analytical and theoretical deficits would need to be addressed in the methodology of any further experimentation if the claims of a new perceptual observation were to be substantiated. It was also clear that any presumption that the perceptual abilities of the author were generalisable to the study participants could only be established by regular preview demonstrations seen by colleagues and invited viewers. This procedure should indicate if any of the effects were too subtle to be seen by first time viewers and only observable by experimenters experienced with the stimuli.

H.7 The limitations of Photographic Reality

In my experience, there seems to be a contradiction in the beliefs of amateur and professional photographers. Amateurs tend to regard any unusual image rendering as a "lens distortion." For them, a fisheye lens distorts reality because straight lines are rendered with extreme curvature and near objects are massively over enlarged.

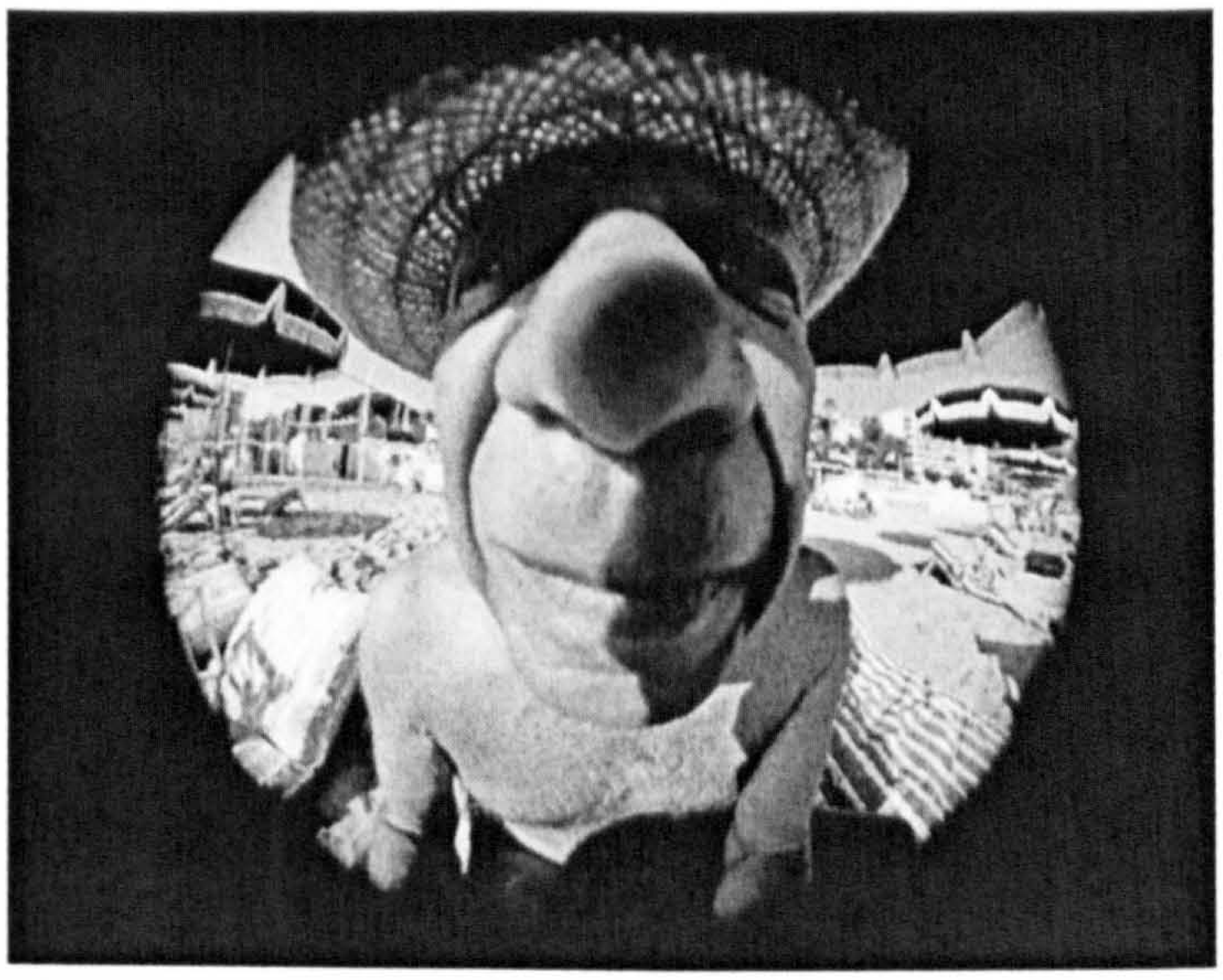


Figure 7 A fisheye lens portrait.

To most viewers, the picture above is highly distorted. But for most photographers, the picture is an almost perfect representation of what you would see if your eye was as close to this face as the camera is. Professional fisheye lenses are designed to be "mathematically perfect" and capture perspective in a consistent and predictable way. This effect is so predictable that computer programs like Apple's QuickTime VR can re-render images like these to make the straight lines from the original scene straight again and restore the original proportions to a more conventional appearance. The most important thing for the layman to understand is that all lenses capture perspective with the same level of accuracy, regardless of whether the lens is the widest of wide-angles, or the longest of telephotos. If one enlarges the centre of an image made with fisheye lens to the same degree as a telephoto lens image made from the same position, the perspectives rendered by both lenses are identical.

The perceptual problem for most photographs is not caused by the methods of image capture, but by the methods of reproduction. If one were to fit a fisheye lens to a 35mm camera and point it straight up at the night sky, one could photograph the positions of the stars with great accuracy (using high resolution transparency film) in relation to the pole star and the horizon. But if one were to reproduce that image onto a page in a book, these relationships would appear to be highly distorted. A way to recover the original perspective would be to mount the same lens on to a slide projector pointed straight up. The projection screen would need to be a very large white hemispherical dome, with its "equator" set at the same height as the top of the projector. The 35mm slide one had made of the night sky could then be projected into the dome so that its equator becomes an artificial horizon. If the viewer were to then stand next to the projector and look in any direction, he would see a very accurate reproduction of the night sky imaged by the camera. This is the basic principle behind the techniques used for traditional planetarium designs. But it was also used by the IMAX corporation for films made in its Omnimax format, and is now most often used for 360 degree webcams and QuickTime images on the internet.

Professional photographers however are not necessarily free from criticism of their understanding of "photographic" distortions. For example, they are taught that it is acceptable use a 200mm telephoto lens to photograph a scene that requires a 400mm lens to fill the frame. All that is required is to enlarge the image to the required magnification and to crop-off the unwanted parts of the image. As perspective is the same from any viewpoint, the only downside to this technique is a very small loss of image quality due to enlarged film grain (or pixels) and lens aberrations that would not normally be seen. Viewers of the final photograph are rarely, if ever aware that large amounts of an image may have been cropped away like this and the centre magnified. But they are likely to be aware when circular objects are reproduced as ovals, rectangles are seen with curvilinear distortions or that or objects have been stretched anamorphically. So the photographers faith in the rendering abilities of their finest lenses may blind them to fact that the images in their photographs may indeed have subtle perceptual distortions. Photographers may get enough information from their photographs about size and nature of real objects (as they appeared to them when the photograph was taken), that they may be unaware of any perceptual changes. *A viewer who did not see the original scene or subject may have a subtly different appreciation of it, simply because they were not present when the photo was taken.*

H.8 The "perfect" camera and projection system.

The "perfect"camera and projection system is easier to imagine and than it is to build. If the design criteria are to reproduce exactly what one can see from a fixed view-point, one could theoretically do it using an orthostereoscopic technology like this: One would build two micro-cameras and fit them with fisheye lenses. These cameras would be mounted side-by-side with a lens inter-axial separation of 65mm. The camera sensors would have a resolution matching the resolving power of the human fovea, yet do this across the whole of their surface. Each pixel would also be sensitive to the whole range of colours, brightness and distances that we can see, and be able to convey each of those values to an equivalent pixel in a matched projection system. The projectors would have each individual pixel matched to the performance of its partner sensor pixel in the camera. It would modulate white light falling on it to output the exact value recorded in the camera. It would then steer the modulated light through a fisheye lens mounted on the projector. The projector would then be mounted in a matched pair and fitted with linear polarizing filters, similar in principle to the design in figure 3.2. However, they would be mounted vertically to point straight down onto a semi-silvered mirror mounted at 45 degrees to the screen. The screen would not be a flat metalized material. Instead it would be a vertically mounted dome coated with retro-reflective material designed to return light directly back to it source with very high efficiency. The diameter of the dome would be at least 4 metres wide so that it could exceed the limit of differential focusing by an individual eye.

The viewer would see the projected orthostereoscopic scene from a chin rest to place their eyes on the ideal viewpoint for the retro-reflected image. If the screen was perfectly optimised, the viewer would only be able to see a bright image when their eyes occupied the ideal viewing position. Cross-talk (co-channel interference) would be almost zero and off-axis viewers would not see any image at all. The user would need to wear linear polarizing spectacles with mini cameras in the frames and vari-focal lenses. The cameras would detect the vergence of each eye so that a computer can determine where the viewer is converging their gaze, and therefore each eyes point of focus. If the viewer were to gaze at a target at optical infinity, the display would project each image with 65mm of separation and change the dioptre setting in the spectacles to allow each eye to focus at infinity. If the viewer were to change their gaze to a target in virtual space that is much nearer, the spectacles would allow his eyes to accommodate on the target as if it were in real space in front of his eyes. The images would also be realigned so that the convergence of each eye perfectly matches the angles we would use to see the real scene. This process would not only retain the linkage between vergence and accommodation (which is lost in conventional stereoscopic images) but would also allow the point of zero disparity in the original scene to be accurately reproduced. To do this, it would be imaged in virtual space at the same size, distances and convergence angles as would be used to view the real scene from the stereo cameras position in real space. Combined with life-like colour, brightness and resolution, the stereoscopic information should be able to reproduce the content of the original scene with enough accuracy that it might be impossible for a viewer to discriminate between an artificially mediated scene like this and reality.

The sci-fi nature of the display described above would make it seem that lifelike reproduction and stereoscopic displays are unavailable to photographers or researchers wishing to conduct psychophysical experiments using stereoscopy. However, it was clear in side by side comparisons when viewing the real subjects standing next to their projected images, the system used in the pilot studies was remarkably accurate at producing a life-like "virtual twin" of everyone photographed. While no available system could ever offer perfect reproduction, the potential for an improved version of the first stereoscopic system seemed obvious to all who saw the images produced from it. The reasons for its potential were all based on the fact that it was comparatively easy to optimise the components, specifications and film used to that point. Also, the improved image capture and projection specifications (such as magnification, convergence, lens aperture, camera to subject distance, film exposure/brightness) could all be permanently locked into an optimised state. The resulting projection system would allow experiments to be run whereby in 2D projection conditions, the display would offer a conventional photographic image. But when the display was switched to its orthostereoscopic mode, it would produce an image that would be more accurate than possibly any seen before. And it should offer a unique opportunity to quantify the perceptual differences between conventional 2D and orthostereoscopic photography in well-designed experiments for the first time.