Jason Drummond
UCL
Engineering Doctorate for Virtual Environments, Imaging and Visualisation

Towards the use of Visual Masking within Virtual Environments to Induce Changes in Affective Cognition

To Luca, Caroline, Mum, Dad, Marie & Jean Louis.

Acknowledgements

I would like to thank my supervisors Anthony Steed and Nadia Berthouze for their wonderful support throughout this work.

I thank my wife, Caroline, for her wholehearted, unfailing support and understanding throughout this long process.

I would also like to acknowledge David Drummond and Jean Louis Pujol. I am grateful for their financial support, without which this thesis would not have materialised.

I would also like to thank Mel Slater for taking the time to review the work at various stages and offering sound advice. I would also like to acknowledge and thank all of the many participants for their time and effort and the various students and staff of the Centre for Doctoral Training for their invaluable advice.

This work was funded by the Engineering & Physical Sciences Research Council through its EngD VEIV Centre for Doctoral Training programme at UCL. I am very grateful to all involved. I am proud to have been, in some small way, a part of this programme.

I, Jason Drummond confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Thesis Abstract

This thesis concerns the use of virtual environments (VEs) for psychotherapy. It makes use of VE properties that go beyond real-world simulation. The core technique used is based on research found within perception science, an effect known as backwards visual masking. Here, a rapidly displayed target image is rendered explicitly imperceptible via the subsequent display of a masking image. The aim of this thesis was to investigate the potential of visual masking within VEs to induce changes in affective cognition. Of particular importance would be changes in a positive direction as this could form the foundation of a psychotherapeutic tool to treat affect disorders and other conditions with an affective component.

The initial pair of experiments looked at whether visual masking was possible within VEs, whether any measurable behavioural influence could be found and whether there was any evidence that affective cognitions could be influenced. It was found that the technique worked and could influence both behaviour and affective cognition. Following this, two experiments looked further at parameter manipulation of visual masking within VEs with the aim of better specifying the parameter values. Results indicated that the form of visual masking used worked better in a VE when the target and mask were both highly textured and that affective effects were modulated by the number of exposures of the target. The final pair of experiments attempted to induce an affect contagion effect and an affect cognition-modification effect. An affect cognition-modification effect was found whereas an affect contagion effect was not.

Overall, the results show that using visual masking techniques within VEs to induce affect cognition changes has merit. The thesis lays the foundation for further work and supports the use of this technique as basis of an intervention tool.

Table of Contents

Tables

Figures

Acronyms

ART	Attention restoration theory	203
CAVE	CAVE automated virtual environment (recursive acronym)	56
CBM	Cognitive bias modification	16
CBT	Cognitive behavioural therapy	16
CFS	Continuous flash suppression	61
DPS	Direct parameter specification	76
EMG	Electromyography	100
FEB	Facial expression blindness	97
FFA	Fusiform facial area; located on the infero-temporal lobe/occipital lobe junction of the brain	93
FFH	Facial feedback hypothesis	113
FRU	Face recognition units (Bruce & Young 1986)	96
FSE	Facial superiority effect	93
GAD	Generalised anxiety disorder	103
HMD	Head mounted display	49
IAT	Implicit association test	24
ISG	Inter-stimulus gap, a spatial variable	69
ISI	Inter-stimulus interval, a temporal interval	69
NAT	Negative automatic thoughts: a theoretical construct underpinning cognitive appraisal theory, sometimes used to explain the robustness of affect disorders	46
NCE	Negative compatibility effect	67
NHS	The UK National Health Service	17
NICE	National Institute for Health & Clinical Excellence, a UK Government body evaluating NHS treatments	48
NRT	Nature restoration theory	203
PANAS-X	The Positive and Negative Affect Schedule – Expanded Form	33
PI	Place illusion, seen as a component of presence	50
PIN	Person identity nodes, (Bruce & Young 1986)	96
POMS	Profile of Mood States	33
Psi	Plausibility, seen as a component of presence	50
PTSD	Post traumatic stress disorder	48
OSE	Object superiority effect	73
RDS	Random dot stereogram	72
RSVP	Rapid serial visual presentation	68
SAD	Social anxiety disorder, in this thesis	103
SOA	Stimulus onset asynchrony	68
STA	Stimulus termination asynchrony	71
TOM	Theory of mind, conspecific assumptions about another's cognitive state, includes affect	97
V1	Striate cortex in the occipital bun, associated with early vision processing	76
V2-Vx	Neocortical areas associated with increasingly complex vision processing	77
VE	Virtual environment: a virtual world created on a software platform such as a game engine	5
VR	Virtual reality: a mixture of certain hardware and software paradigms that can lead to a sense of immersion within a VE can produce an experience of VR	43

1 Introduction

This thesis looks to the prospective use of VEs for psychotherapy, using non-simulation aspects of VEs. The main technique used is based on research found within the psychology of perception, an effect known as visual masking. The aim is to investigate the potential of visual masking within VEs to create a psychotherapeutic tool that can induce a positive change in state affect, either directly via affect contagion or indirectly by cognitive modification. The intended agents of change are virtual objects that carry an affective valence. Previous studies have shown human faces to be particularly potent carriers of affect signals and that these may have a priming effect on emotion or mood in human observers.

The work in this thesis therefore rests upon three, main research topics:

1. The use of VEs in psychotherapy, including the benefits and limitations that VEs bring;

2. Visual masking studies including parameter issues, general priming and priming using visually masked facial imagery;

3. Affect and related cognitions research, in particular mood cognitions.

The benefits of placing masked facial stimuli within a VE will be discussed more fully in the Literature Review, Chapter 2, but are outlined in the background introductions to the three topic areas below.

1.1 Background

1.1.1 Virtual Environments and Psychotherapy

One approach taken by the mental health profession to the treatment of affect disorders is to prescribe some form of psychopharmacology to stabilise the condition followed by psychotherapy of some type to attempt a longer-term amelioration (NHS Choices: Depression Treatments 2010). For example, a person diagnosed with depression might be prescribed a selective serotonin reuptake inhibitor (SSRI), such as Prozac, followed by an appropriate course of cognitive behavioural therapy (CBT). Although the CBT would typically last for anywhere between eight and twenty six weeks, its effects are expected to last for a lifetime.

Would it be possible to address the question of recipient affective state prior to or alongside CBT without the use of psychopharmacology? There are a number of reasons as to why this may be desirable. A person may not want to take drugs or it may not be possible to find the right prescription for that particular person. GPs and psychiatrists often have to try several different drugs or drug types with a patient before a helpful one is found. Could a change in state affect induced by other means also act as a useful preparation or adjunct to psychotherapy? For example, would a person experiencing severe dysthemic disorder be more open to CBT intervention following a session of positive mood induction or cognitive bias modification (CBM) (see section 2.2.1 & 2.4.3 below) for instance? Would the particular properties of a VE be of use in bringing about mood induction or cognition modification? These questions are drivers for this thesis.

By accompanying psychotherapy with appropriate mood induction or cognition modification techniques, it may be possible to increase the chances of short-term success for each individual. We may also find that the longer-term prognosis is better also. Instead of periodic CBT 'booster' courses, it may be possible that only the associated mood induction or cognition modification method need be applied. This raises the wider question as to whether repeated changes in state affect using a VE could induce a longer-term trait affect change. In other words, could whatever method used for positive mood induction or cognitive modification constitute a therapy all of its own? This may have particular relevance to milder cases of affect disorders deemed treatable by initial interventions from the UK National Health Service (NHS) stepped care programme. Specific psychotherapy techniques do not offer a 'one size fits all' approach, indeed some practitioners attempt to integrate many disparate strands in order to meet what are seen as the varied client needs (O'Brien & Houston 2000). Therefore, widening the spread of techniques available can only prove helpful.

VEs have been researched and put to use in psychotherapy for specific disorders such as certain phobias (Carlin *et al.* 1997, Klein 2000, Mühlberger *et al.* 2003, Mühlberger *et al.* 2006) and post traumatic stress disorder (Rizzo *et al.* 2005, Difede 2002). Virtual characters have been investigated for their behavioural and affective effects on participants within VEs (Pan *et al.* 2008). Virtual characters and VEs may also be holding out some promise on the understanding, diagnosis and treatment of eating disorders (Ferrer-García *et al.* 2005, Ferrer-Garcia & Gutiérrez-Maldonado 2009, Gorini *et al.* 2010). For instance, it is claimed that modulating a user's sense of 'perceived subjective presence' in a VE will affect their modelling of avatar eating behaviour (Fox *et al.* 2009). Avatars have also been used as consciously perceived, emotion induction stimuli (Causse *et al.* 2007).

There appears to be a trend towards using VEs to simulate some aspect of the therapeutic process as it stands in the real world. For example, virtual spiders are used as stimuli to treat arachnophobia (Carlin *et al.* 1997). There are some deviations from this such as the snow-bound, fantasy world used by Hunter Hoffman for pain distraction (Hoffman *et al.* 2001). Mood change is generally not an explicit aim of such simulations although an exception can be found with the Engaging Media for Mental Health Applications (EMMA) project's 'mood induction devices' (Alcañiz *et al.* 2003). VEs do provide us with the opportunity to immerse people in an experience that can evoke strong affective reactions (e.g. Meehan *et al.* 2002). People do seem to behave in VEs as they would in a similar, non-virtual situation, although sometimes in a less intense way (Slater *et al.* 2006).

VEs, however, allow us to present and control perceptual effects not possible in the real world. This may hold the promise of further extending the use of VEs for therapeutic use, especially if an effective, consistent and reliable state affect change were derivable from such an effect. If long-term mood or cognition shifts were possible by using VEs on clinical populations then an adjunct or alternative to some psychotherapy or psychopharmacology may be feasible.

Extending the research to produce a useful therapeutic tool, ideally we would need a VE that reliably produces a desired effect, consistently produces the same controllable level of effect and produces little in the way of side effects. The latter is important because some confounds may have a negative impact upon any therapeutic process, undermining any gains from using such a tool. Questions of reliability and consistency are perhaps difficult to address as it can be argued that the presence or level of any mood induction or cognition modification can be viewed as having a subjective component. It may be possible to introduce some measure of standardisation: inventories, physiological measures against a baseline etc. However, currently we may have to recognise and accept a variance in this.

However, before a tool such as that suggested can be engineered, the science must establish whether the effects it would rely upon are possible. What effects are possible within a VE that could give rise to the desired changes in affect of cognition and what advantages would a VE bring over other stimuli?

1.1.2 Visual Masking

As noted, VEs offer more than just simulation. We could use perceptual techniques from psychological research, for instance, altering the presented visual field and investigate for an affective reaction. One technique that could be applied is visual masking. Here the rapid presentation of one figure can be used to render a second figure effectively invisible, at least to conscious perception. Although a masking stimulus can be used to prevent conscious visual awareness of a target stimulus, the target may still be processed non-consciously. There is a paucity of research using visual masking within VEs at present.

In one form, known as backwards masking, an image (the target) may be presented onto a screen for a very short duration, in the order of milliseconds. This is then replaced by a second image (the mask) which stays onscreen for a longer duration, enough to be consciously perceived. This second image is then removed, leaving the original screen background. This process is then cycled repeatedly, leaving the viewer with the impression of a flickering mask image only, the target being rendered invisible. If the mask were not present then the target may be perceivable, depending on the time that it was onscreen. The presence of the mask appears to interfere with the conscious perception of the target.

Despite a lack of *conscious* awareness, there is evidence to show that affective reactions to the target still occur: not only are overt behavioural changes measurable (Winkielman & Berridge 2004) but facial muscle reactions to masked expressions are reported (Dimberg *et al.* 2000). Not only can visual masking techniques be used to affect behavioural processes but non-conscious, affective processing generally (Whalen *et al.* 1998, Ohman & Soares 1994, Dimberg *et al.* 2000, Ohman 2002, Wong & Root 2003). It can be argued that the latter overlaps with the functional domain of psychotherapy (Hassin *et al.* 2005) to some extent. One of the criticisms that can be levelled at CBT is that it might not be addressing those implicit processes that it claims to, instead remaining on a purely explicit level. However, on reaching implicit affect processes, visual masking would not be subject to this criticism.

There is evidence to show that people consciously modulate their own display of emotional facial expression (Panksepp 1998, Tedeschi *et al.* 1971, P. Ekman and W.V. Freisen 1969) depending on such factors as social norms expected within a given culture; what Ekman calls 'display rules' (Ekman 1972). In other words they manage their overt state affect, at least as far as display to others is concerned and possibly their own explicit affective state itself, as supposed by the facial feedback hypothesis (McIntosh 1996). By using visual masking techniques it may be possible to circumvent this cognitive, top-down interference and effect a change on a deeper, non-conscious level. If we could reproduce masking effects in a VE then we may be able to develop methods to alter mood or cognition on an implicit level leading to a better therapeutic outcome. Such methods may prove useful in developing future, virtual interventions. The advantages of using VEs to present visually masked stimuli are bound up both with the programmable flexibility of such stimuli and the notion of presence and plausibility. These will be discussed more fully in Chapter 2, below.

The use of visually masked facial expressions with an affective valence have proved effective in inducing affective reactions of one sort or another. It may be particularly useful from a psychotherapeutic point of view that such stimuli are related to social processing: they may directly affect a viewer's internal affective state, via a contagion effect, or do so indirectly via implicit self-esteem processing (Leary & Baumeister 2000). However, the stimuli used in such research tend to be static, planar images not presented within a VE. They therefore lack any of the advantages that VEs can bring.

A VE may present a plausible enough world to engage participants and so allow social acceptance signals from masked virtual agents to effect a positive change in mood. This could also happen directly via affect contagion or indirectly via possible cognitive modification of social processing. Any change in either or both domains, though, could lead to a virtuous circle as one acts as positive causal agent for the other, resulting in a positive impact on implicit self-esteem. If these effects were realisable then they could form the basis of the intervention tool touched upon above.

1.1.3 Mood, Cognitive Bias and Implicit Self Esteem

The terms affect', 'mood' and 'emotion' will be discussed more fully in Chapter 2 but here it is suffice to say that, for the purposes of this thesis, mood and emotion are seen as types of affect. Therefore, the term 'affect' will be used generally throughout the text acting as an umbrella term for the other two. All affect is regarded as comprising of component processes, a model consisting of multiple elements such as cognitive and physiological events etc. (Schere 2005). It is also important to note that, within this thesis, affective processes are considered to be types of cognition. They may be given their folk psychology labels, such as ‘mood’ for instance, to distinguish them from other cognitive types or may be referred to more fully as, for example, ‘mood cognitions’.

There are several mood induction techniques used within research to induce an analogue of a natural mood in a controlled way (Gilet 2008). We would need to consider mood induction techniques appropriate to the objectives developed in this thesis. One method may be found by utilising affect contagion, in exposing participants within a VE to positive facial expressions such as an agent’s smile directed at them. Emotional contagion (Hatfield *et al.* 1994, Saxe & Baron-Cohen 2006) and mood contagion (Neumann & Strack 2000) are related types of affect contagion. Here, affective valence changes appropriate to the stimuli are spread from person to person, communicated by various modes such as facial expression, body posture, vocal intonation etc. Facial expression has been shown to be instrumental in facilitating general affect contagion (Wild & Bartels 2001). This thesis is partly concerned with mood contagion. Moods can be seen as related to emotions but are regarded as of lower intensity and longer duration (Fox 2008, Table 2.2).From this it follows that mood contagion depicts low intensity, long duration affect valence change spread from person to person via various behavioural modes. Some affective responses to stimuli do not appear to need instigation by ‘higher’ cognitive processes (Zajonc 1980). Affect contagion may be describing one such direct route of influence, from affect to affect. It may not rely on cognition outside of affective processes, beyond initial perception. For example, seeing a very frightened face may be enough to generate fear in a viewer, regardless of immediate threat judgements. However, other routes exist which lead to changes in affect.

Research has shown that just as affect can causally change cognition (Gasper & Clore 2002, Dolan 2002), so too can other cognition cause affective changes (Smith & Lazarus 1990, Beck 1967, Mathews & MacLeod 2002). Modifying the cognitive biases that are thought to support affect disorders may prove equally valuable in causing a change in underlying affect (MacLead *et al.* 2002). Here, the aim would be to instigate a virtuous circle where positive cognition caused positive affect that caused positive cognition and so on. Current cognitive modification techniques, such as CBM, tend to be explicit in nature i.e. the participant is aware of the agent of change and can possibly guess whether they are in the experimental or control condition. Despite the great deal of promise held out by CBM techniques (Koster *et al.* 2009) they are therefore subject to demand characteristics that may distort their effectiveness or measurements of their effectiveness. Cognitive modification via masked stimuli within a VE may circumvent these difficulties. An example of how useful this might be can be found by considering the functional domain of social processes such as implicit self-esteem. By effecting changes in this domain, we may be able to create a lasting impact on the lives of sufferers of disorders related to these processes. We may be able to circumvent explicit cognitive control (arguably identifiable with 'conscious resistances' or 'defences', depending on the model of psychotherapy) to modify implicit, disorder supporting biases.

Low self-esteem is considered one of the indicators of both dysthemic disorder and certain categories of major depression (DSM-IV-TR 2000). It is also prevalent in certain anxiety disorders such as social anxiety. The facial expression of others directed towards us can act as indicators of social acceptance that influence implicit self-esteem (Baccus *et al.* 2004, Dandeneau & Baldwin 2004). Implicit self-esteem is not concerned with the way people describe themselves to others nor the way in which they consciously perceive themselves. Both are described by explicit self-esteem and are thought to be subject to certain cognitive biases arising from social and cultural pressures (Suls & Krizan 2005, p79.). These biases can distort the picture of a person's actual self-image. Implicit self-esteem is thought to be a construct based upon automatic concept-attribute associations regarding the self. As such, it is not necessarily available to conscious self-examination and needs to be revealed via activities such as the Implicit Association Test (IAT) (Greenwald & Farnham 2000). In associating the concept of implicit self esteem with the IAT, Greenwald & Farnham are not only claiming that the IAT captures implicit self esteem but further suggest that this is 'true' self esteem, free from what are termed impression management (Tedeschi *et al.* 1971) and self-deception (Gur & Sackeim 1979).

If the subjective experiences of positive facial expressions of others, directed at ourselves, have an effect on implicit self-esteem then a positive impact on affect conditions, like social anxiety or dysthemic disorders, may be possible. However, as noted, it would be necessary to find a way to alter state affect from, or below, the level of cognitive distortions that have a negative impact on the therapeutic process. To do this we may be working with implicit self-esteem processes.

1.2 Main Hypothesis and Research Questions

The overall hypothesis of this thesis is that a VE can be constructed using visual masking techniques that can be used to bring about changes in affective cognition. Three sub-hypotheses can be made regarding this:

a. VEs generally attempt some form of simulation of the real world when used for psychotherapy. The non-simulation attributes of VEs remain largely untapped. Nonetheless, we may be able to use them to disrupt negative, non-conscious, affective processes and cognitions.

b. Visual masking provides one possible way in which to exploit such non-simulation attributes to reach implicit processes. However, visual masking still needs to be shown to work within a VE.

c. In order to disrupt negative, non-conscious, affective processes it is suggested that visually masked stimuli that carry a positive valence be used. An example of this would be a smiling faces directed towards a participant. There may be a direct effect via affect contagion or an indirect effect via cognitive modification. It is currently unknown whether this could bring about a mood change or not.

From these points, several research questions arise:

a. Initially, the problems to solve will be technical ones: can visual masking be made to work within a VE? There are aspects of the masking effect that may present difficulties here such as inhibition by stereoscopic depth disparity.

b. The next question will be to ask whether it is possible to elicit a response to emotionally valent, masked stimuli in a VE. Will participants respond in a measurable way to visually masked, emotive expressions? At this stage, some initial modulation factor exploration may be useful.

c. If effects are found, the following work should be again of a technical nature: what parameters, if any, could be manipulated in order to modulate any effects or increase the reliability or consistency of those effects. Here masked, expression dynamics could be introduced within a VE. Will these be understood as carrying the valence intended? Other parameters should include the surface qualities of masks and targets such as colour saturation or texture legibility tested against affective salience.

d. The final problem will be how to use visual masking in a VE to change a person's affective state, specifically altering their mood in a positive direction, directly. Alternatively, can cognition be changed implicitly, with the expectation that this will, in itself, lead to a change in affect.

1.3 Experiments

This section outlines the six experiments undertaken which form the core of this thesis. The abstracts are presented alongside their adjoining hypotheses. All experiments had full ethical approval from UCL.

1.3.1 Experiment 1

Visual masking techniques produce strong effects but the stimuli used are often co-planar and with no internal depth disparity. Prior research has shown that external depth disparity, the distance between target and mask, tends to inhibit any masking effect. This study investigated not only whether a virtual object's internal depth disparity inhibits masking but also whether it is possible to achieve visual masking within a VE in the presence of either depth disparity type. Non-planar, virtual objects were used as targets and planar objects were used as masks within a stereoscopic VE. Forced choice tests were administered to measure any masking effect. An effect was found that showed it was possible to construct target objects with internal depth disparity in a VE. In addition, the effect was tolerant of some external depth disparity.

This experiment was designed to answer the first research problem.

Hypothesis 1a

It is hypothesised that visual masking can be made to work within a VE. If the participants do not explicitly detect the masked, target objects within the VE then this will support the proposition that, although the objects are being displayed, they are not being seen consciously.

Hypothesis 1b

Similarly, where participants claim to see the target objects, they will not identify the object type correctly. Having to identify the object type will act as a check on the participants' responses and would lend further support to the proposal that visual masking techniques can be successfully applied to VEs.

Hypothesis 1c

The position of the target objects will be detected implicitly by participants. Despite participants not being explicitly aware of the position of the target objects, their choice behaviour could indicate an implicit awareness.

See Chapter 4 for further explanatory detail of this experiment and a discussion of its results.

1.3.2 Experiment 2

It was investigated whether masked, non-planar facial expressions could have an effect on participant choice behaviour within a VE. The VE set-up was similar to Experiment 1, using the same basic virtual world. The participants were free to move through the world choosing which doorway to pass through to enter the next room. However, each doorway was masking a non-planar, static face with one of three expressions: smiling, angry or neutral. Each pair of doors contained different expressions on either side. There was also a modulating factor introduced: a conflict or non-conflict condition created by varying instructions to participants. The results showed a significant difference between expressions but the modulating factor produced no significant interaction with facial expression nor a main effect.

This experiment was designed to answer the second research problem.

Hypothesis 2a

There will be a significant interaction between the Threat variable and the Expression variable. An interaction is expected because it is thought that a participant's response to a particular facial expression would differ depending on their emotional state. In the presence of a threat, for instance, it is possible that more resources would be given to processing negative expressions than positive.

Hypothesis 2b

If no interaction is present then there will be a significant main effect of the Expression variable. Even if the threat variable did not produce an interaction or main effect of its own, the expression type on an implicitly perceived, target face may produce a behavioural effect in itself.

See Chapter 5 for further explanatory detail of this experiment and a discussion of its results.

1.3.3 Experiment 3

The third experiment extended the work of Experiment 1 and 2 by looking at the effects of altering various properties of visually masked, non-planar objects within a VE. Two aspects of surface quality were investigated: texture image fluency and texture colour saturation. In addition, the saliency of the masked objects was introduced as a variable. The aim was to look for disruptions to the masking effect and to see if any interactions were present. On encountering flashing doorways, which sometimes acted as masks for various target objects, participants were required to state if they could perceive any objects and, if so, indicate their type (from a reference sheet provided). Targets that may or may not have been in the doorways included various affect neutral domestic objects and highly salient facial expressions. This experiment found a significant interaction between surface colour and texture fluency. Saliency did not interact with either of the other two variables nor produce a main effect.

This experiment was designed to answer the third research problem, in part.

Hypothesis 3a

Target salience will interact significantly with saturation. Salience could be expected to show more than an additive effect in the presence of surface quality variables because of the increase in attention brought by the detection objects of high affective salience. Increased focus on an object, even if implicit, might be expected to heighten the differences between effects wrought by the various surface qualities such as saturation.

Hypothesis 3b

Target salience will interact significantly with fluency. As above, it is expected that salience would heighten the differences between the effects of produced by different levels of fluency.

Hypothesis 3c

Saturation will interact significantly with fluency. Certain levels of saturation or fluency could in themselves act in a similar way to that expected of salience in that those levels implicitly gain more of the perceiver's attention than another level.

Hypothesis 3d

As all of the variables are expected to interact with each other, it is expected that an overall interaction will be significant.

See Chapter 6 for further explanatory detail of this experiment and a discussion of its results.

1.3.4 Experiment 4

The fourth experiment looked at whether it was possible to form masked targets from dynamic (i.e. changing) objects. Facial expressions were used as non-planar targets that changed from neutral to a full expression over several cycles of target and mask presentation. Each time a target was flashed onto the screen it would change slightly from before, similar to a frame of animation. The intention was to see if these dynamic targets were read, implicitly, as animated expressions or as a series of still images. In the VE, each dynamic, masked expression was juxtaposed with an affect neutral object to which the participants had to give an attractiveness rating. A comparison control used static masked expressions. Results showed that the dynamism and the expression variables interacted but that the static scores produced the highest variance, indicating modulation by an exposure effect. This contra-indicates that the dynamic expressions were being read as animated faces.

This experiment was designed to answer the third research problem, in part.

Hypothesis 4a

An interaction will occur between the Expression and Dynamism scores.

Hypothesis 4b

The difference between the two dynamic group scores means will be greater than difference between the two static group score means. This hypothesis extends 4a in that it states that dynamic expressions would heighten a participant's reaction to that particular expression when compared to the static version.

Hypothesis 4c

When scores are aggregated into a smiling group and an angry group, the mean of the smiling scores will be higher than the mean off the angry scores. This would be expected if the implicitly perceived, smiling faces altered a participants evaluation of a visible, juxtaposed object.

See Chapter 7 for further explanatory detail of this experiment and a discussion of its results.

1.3.5 Experiment 5

The intention of this experiment was to discover if an affect-laden VE could have an effect on either subjective mood assessments, affect-driven cognitions or both. Two independent, between-participant factors of the VE carried an affective valence: the environment and visually masked, facial expressions. The environment factor had two levels, one designed to be mildly pleasant and the other mildly unpleasant. The expression factor utilised two kinds of expressive, masked faces placed in doorways throughout the VE, the expressions being either angry or smiling. The intention was to measure the affective responses in a manner suggested by the component process model (Schere 2005), specifically subjective feeling and cognition. However, physiological measures were not appropriate as mood was to be measured, not emotion. Explicit subjective feelings were measured prior and post experiment using the Profile of Mood States (POMS) and the Positive and Negative Affect Schedule, Expanded (PANAS-X) respectively. Cognitive responses to the affective stimuli were measured using an emotional Stroop reaction test. This also introduced a further independent factor, word valence that was within-participants in form. Results showed that, although a direct effect on explicit feeling was not found, both environment and expression did interact significantly to affect reaction times. Word valence was found to have significant effects, interactive or otherwise. When environment was factored out, taking account of expression and word valence only, a similar result was obtained: explicit feelings were unaffected but expression had a significant effect on reaction times with word valence, once again having no effect. These results suggest a direction for further developing the use of visually masked, affect-laden VE stimuli to bring about mood state change: a direct, affect-contagion route is less likely to yield results than a more indirect, cognitive approach. In addition, the effect of the environment factor also suggests the form of VE in which such a state change might best be effected.

This experiment was designed to answer the fourth research problem, in part.

Hypothesis 5a

There will be no significant interactions or effects between the POMS scores of treatment groups. This would be a necessary condition prior to treatment.

Hypothesis 5b

There will be significant interactions or effects between the PANAS-X scores of the treatment groups. If this hypothesis were supported it would indicate that an explicit mood effect has taken place.

Hypothesis 5c

There will be significant interactions or effects from the Emotional Stroop scores. Such an interaction would show evidence of the effect of treatment group differences on reaction times for processing emotionally valent text. This would support the argument that the form of a VE, the environment factor, could act as a modulator for effects produced by the expression factor.

Hypothesis 5d

Where no interactions are found in the Emotional Stroop scores, the word valence within-factor will show a significant main effect.

Hypothesis 5e

Where no interactions are found in the Emotional Stroop scores, the environment between-groups factor will show a main effect.

Hypothesis 5f

Where no interactions are found in the Emotional Stroop tests, the expression within-groups factor will show a significant main effect. Even if the two factors environment and expression did not interact, it is expected that they will each affect the reaction scores significantly.

See Chapter 8 for further explanatory detail of this experiment and a discussion of its results.

1.3.6 Experiment 6

This experiment was designed to test cognitive bias modifications using masked expressions within a VE. The intention was to prime a bias away from negative expressions, towards positive expressions. Participants were instructed to follow a route through a virtual world in the form of a series of garden hedges, similar to a maze. At each junction, they could choose to go through one of two doorways, each of which was blocked by a masked facial expression. In one condition, only the doorway with the smiling expression would lead to the next area, passing through an angry expression would lead to a dead end and the participant would have to retrace their route back to the other doorway. A second condition reversed the expression placements. Participants were required to complete the route as quickly as possible. They were then asked to complete an adapted CBM training task that was originally designed to modify attentional bias. This task used response times to attentional bias towards smiling expressions and away from negative ones. Results showed as significant difference between group times: participants who experienced the smiling prime condition generally displayed faster reaction times when compared to those who experience the angry primes. This indicates a transfer from implicit to explicit learning.

This experiment was designed to answer the fourth research problem, in part.

Hypothesis 6

The reaction times of the smiling prime group will be significantly lower than those of the angry prime group. This would be expected as the CBM task induces a learned bias to orient attention towards smiling faces. Participants already primed with this bias by their VE should therefore perform better in the earlier stages of the CBM task than those primed with angry expressions. This would lead to significantly lower mean reaction times in the smiling group than the angry group.

See Chapter 9 for further explanatory detail of this experiment and its hypotheses.

1.4 Contributions

The contributions are presented as being either methodological or substantive.

1.4.1 Methodological Contributions

1.4.1.1 Contribution One

From Experiment 1, creating masked objects with internal depth disparity within a VE forms the first contribution for this work. Addressing the technical issues, we know that internal depth disparity of a target does not disrupt the masking effect within a stereoscopic VE, provided that the mask and target remain at approximately the same depth. An initial attempt at a modulation variable produced no significant result.

1.4.1.2 Contribution Two

From Experiment 3, the best masking results, using non-planar targets within a VE, were found using high fluency, greyscale textures for both the target and mask. The affective salience of a target does not appear to affect the masking effect in either a helpful or a disruptive way.

1.4.1.3 Contribution Three

From Experiment 4, an exposure effect that modulates the responses to masked facial expressions forms a contribution. Alongside this lies the finding that dynamic targets are not read as animated, certainly not in a way that can affect evaluative conditioning.

1.4.1.4 Contribution Four

A standard, explicit CBM training paradigm can successfully be adapted to provide a measure of prior, implicit attentional bias modification.

1.4.2 Substantive Contributions

1.4.2.1 Contribution Five

From Experiment 1, despite not explicitly perceiving the visually masked, non-planar objects, participants nonetheless showed a positional awareness of those objects within a VE.

1.4.2.2 Contribution Six

From Experiment 2, showing choice behaviour influenced by masked, non-planar, affective facial expressions within a VE forms the sixth contribution of this work. This experiment indicated evidence for choice behaviour being influenced by visually masked faces, with depth disparity and emotionally salient expressions, via implicit affective processes.

1.4.2.3 Contribution Seven

Experiment 5, showed a link between VE form and participant mood state as well as an affect-driven, cognitive response effect.

1.4.2.4 Contribution Eight

The last experiment has shown that non-planar, masked facial expressions within a VE can implicitly alter cognitive attentional bias. This has important implications for the future use of CBM as therapeutic tool. Some current CBM training paradigms suffer from two problems: they are explicit and therefore subject to demand characteristics, and similar confounds, and they can be tedious to undertake. The latter can affect not only a participant's willingness to engage with a particular therapeutic task but may also affect the efficacy of the task itself, tedium acting as an effect disrupter.

In addition, learning transfer, from implicit to explicit processes, appears to have taken place in this last experiment. Visual masking within VE holds the promise of enabling the circumvention of potentially confounding cognitions.

1.5 Scope

This thesis sits at the convergence of several areas of science including virtual environments, perception, affect and cognition. It is principally based upon the affective effects of masked, virtual facial expressions. The work aims to combine visual masking with virtual facial expression to investigate ways in which a state affect shift might be attempted, either directly or indirectly. The specific state affect shift targeted is a positive mood change.

1.5.1 Virtual psychotherapy

This thesis is not intended as research directly dealing with psychotherapy nor to lead immediately to a new therapy or therapies. However, it proposes the possibility of constructing an affect or affect-cognition altering tool that may be of use in therapeutic work.

1.5.2 Visual masking

Visual masking itself will not be studied of and in itself but will instead be applied as a central part in the investigation. However, it is envisaged that some conclusions will be reached as to how a visual masking effect can be engineered within a VE, thereby tangentially provide work of interest to the study of visual masking in its own right.

1.5.3 Facial expression, affect induction and cognition modification

Affect science forms an important strand in this work, both as a foundation and a recipient. This work will derive principles from the study of faces and facial expressions as well as add something to it. If the study is successful, a link will have been demonstrated between the non-conscious processing of dynamic facial expressions and either participant mood state or cognitive bias. The work should sit alongside other mood induction or cognitive modification techniques.

1.5.4 Affect contagion and self-esteem

Both self-esteem and affect contagion research were used as background and to drive the various hypotheses, particularly with regard to facial expressions and social cognition. However, this work will not add significantly to the study of these two areas. No definite conclusions will be drawn as to the mechanisms underlying any affective changes.

1.5.5 Cognitive biases

The final experiment in this study overlaps CBM research in that it adapts a standard CBM paradigm as a measure. It shows that is possible to alter implicitly cognitive attentional bias. Added to this, the theme of perceptual bias pervades the thesis as a whole.

1.6 Thesis Structure

Chapter 2 will focus on specific areas of relevance to this thesis already introduced by the background section and will be structured into the same base topic areas. As well as giving a general background to relevant research, arguments will be developed which lead to the research questions upon which this thesis is based.

Chapter 3 sets out the experimental methodology for this thesis and includes all of the methodological resources common to all experiments. Points specific to individual experiments are discussed in the chapters on those experiments. Where an issue of methodology applies to one or more experiments (but not to all) it will be discussed in the chapter on the first experiment in which it arises.

Chapters 4, 5, 6, 7, 8 and 9 describe the six individual experiments, one per chapter. Each chapter will begin with a restatement of the appropriate abstract found in this introduction (section 1.3). Next, a brief background will be set out although most of the theoretical underpinnings of any given experiment will have been covered within the Literature Review (Chapter 2). A methodology specific to that experiment will then be presented, giving details of design, participants, materials, measures and procedures specific to that experiment. This will then be followed by results and a conclusion.

In Chapter 10, a final discussion will be drawn out through general thesis conclusions. This will also include the impact of the work, based on the thesis contributions (section 1.4), and suggestions for future research work. The thesis closes with the final three parts being the Thesis validation (Chapter 11), References and Appendices.

2 Literature Review

2.1 Introduction

This review is comprised of three main topic sections. The intention is to begin each section with a broad view of the topic before narrowing towards points specific to this thesis. In this way, it is hoped that a sound basis for the experimental work will be presented.

Section 2.2 will introduce the use of VEs and virtual reality (VR) for psychotherapeutic intervention. Beginning with a brief look at computer-based therapies in relation to cognitive approaches to intervention, the section then moves on to consider the advantages and limitations of VEs and VR for psychotherapy. Finally, further possibilities for these kinds of therapies will be expanded upon. Section 2.2 will serve to place this thesis within the broad context of computer-based psychotherapeutic interventions and the narrower context of VE psychotherapy.

Section 2.3 concerns visual masking. It begins with a basic taxonomic review and brings in issues of perception, attention and priming that might affect the intended use of masked stimuli presented in this thesis. Following this, the application of visual masking to masked priming using facial stimuli will be considered in sub section 2.3.5. This sub section is of particular importance to this thesis as it discusses research that underpins the use of masked facial imagery to effect a change in an observer. Here, face and expression processing are examined at length and the implications for experiment design are drawn out. This then culminates in a discussion of studies that appear to show implicit processing of facial expressions taking place and possible modulation factors of such effects.

The third topic section, 2.4, will deal with affective cognition and includes cognitive biases, mood induction, social processing disorders and implicit self-esteem. This section is intended to provide an examination at some depth of the possible theoretical bases underlying affective responses to the proposed stimuli.

The review will end with a conclusion that links together various points made in the topic sections. At this point, the implications for the current thesis will be drawn out and the general basis for the experiments will have been established.

2.2 Psychotherapeutic Intervention via Virtual Environments

2.2.1 Introduction

This section of the Literature Review is intended to not only place the work of this thesis in the context of computer-based therapy but also to demonstrate where further research is needed. Many different forms of therapy have been ported into a computer-based form and many interaction paradigms have been used, VEs included. Indeed, it will be shown that VEs are particularly useful in this area, despite some limitations. However, this section will argue that what computer-based therapy generally lacks to date is a way of addressing non-conscious processes that could have an impact on the therapeutic process. It will also be argued that VEs can have a role in addressing this, not just because they can simulate the real world to some extent but that they can also be made to act in ways that reality external to a VE cannot.

Mental health interventions can lend themselves to digital form. Evidence for this can be seen in the plethora of examples from the overlap between computer science and psychotherapy including computer-based therapies, cybertherapy conferences and journals. The reasons may be many and varied but some must surely lay in the flexibility of form and convenience of delivery that computer generated stimuli possess. Computer based mental health interventions take on many forms ranging from text-based to image-based therapies and beyond to interactive systems and virtual worlds.

Two interventions that lend themselves well to digital form, and therefore benefit from the flexibility and convenience that computers allow, are introduced here. They are the long-established CBT and the newer CBM. With CBM, patients are trained to ignore stimuli that are thought to activate certain cognitive biases that may underpin forms of affect disorder. For example, it is believed that an attentional aspect of social anxiety disorder is manifest in a sufferer's bias towards both searching for, and attending to, faces displaying negative expressions (Schmidt *et al.* 2009, Beard *et al.* 2011). CBM training attempts to implant the habitual ignoring of such expression within the processing of visual data from a social gathering or similar using rapid, repetitive computer-based tasks. The assumption is that by learning to ignore such expressions then the triggering of cognitions such as threat or value judgements can be avoided. These cognitions can be equated with a 'negative automatic thoughts' (NAT) found in cognitive therapy theory (Wells 1997). NATs are seen as leading to distressing affect such as the anxiety that can be associated with social situations (Clark 2001).

CBM therapies display some similarity to CBT as both assume a causal path from cognitive to affect processes; a bias in the former causes a disorder in the latter, in the case of affect disorders (Hertel & Mathews 2011). For many years, though, this path was an assumption based on correlation evidence from cognitive research (Browning *et al.* 2010). It was not until the introduction of CBM techniques that this causal line was demonstrated within controlled conditions (Mathews & MacLeod 2002). However, there is some criticism that CBT sits entirely within the level of conscious awareness and cannot directly affect non-conscious processes (Schartau *et al.* 2009). This seems justified when we consider that an aim of CBT is to make explicit 'buried' NATs and deal with them consciously, using reasoned challenge and behavioural techniques. Once deemed to be sufficiently 'reprogrammed', the modified NATs are ‘reburied’ by using repetition to form new habitual responses to the initial trigger. This ‘reburial’ may be nothing of the sort and may merely be the introduction of yet another process to be triggered on top of those NATs already in place, the latter remaining unmodified and awaiting triggering by a slightly different stimulus.

CBT does have a good name for efficacy when compared to other therapies, given the time and costs involved (Layard *et al.* 2006). Its efficacy, however measured, is sometimes placed on a par with psychopharmacology but delivering a better long-term prognosis. However, it would be fair to point out that CBT is not perfect and efficacy rates can always be improved upon. The proffered inability of CBT actually to effect change on a non-conscious level could be one such reason for any failures encountered. CBM appears to work in different way. Where CBT attempts explicitly to challenge NATs, CBM can be seen as attempting to directly retrain them or even inculcate a habitual ignoring of them. A person experiencing CBM is not necessarily aware how they are forming a bias modification or even that a change is taking place, even where they are explicitly aware off the stimuli. The implication here is that CBM does work on the same level as NATs and may deliver results that CBT cannot. CBM therapies inevitably followed the discovery of the techniques and CBM research is currently undergoing something of a rapid expansion: see section 2.4.3 and Chapter 3 below, for further discussion of these points.

Computer-based interventions could be seen as reflecting a progressive zeitgeist within the psychotherapy world: earlier work, driven by psychoanalytic theory. See, for instance, Joseph Weizenbaum’s now-famous ELIZA program (Weizenbaum 1966). This was a somewhat humorous attempt to simulate the natural language processes of ‘talking cures’. This approach has given way to more cognition-based perspectives, recent examples of which are the text-based therapies of Fear Fighter (Marks *et al.* 2003) and Beating the Blues, recommended by the National Institute for Clinical Excellence (NICE) (NICE 2008). In turn, the limitations of CBT theory are now themselves being highlighted by the use of CBM techniques. Yet another therapeutic effect can be derived by using of computers to distract. This, along with CBM, steps away from traditional therapeutic practice of engagement with disorder processes.

Interactive, 2D computer programs have been shown to be of some therapeutic benefit. For example, playing 'Tetris' immediately after viewing a traumatic scene may provide a 'cognitive vaccine' effect to an extent, controlling subsequent post-traumatic, flashback memory (Holmes *et al.* 2009). There may be nothing inherently helpful in the game Tetris itself and, as the authors suggest, any similarly absorbing game or other activity could have similar beneficial effect following a possible, post-traumatic stress disorder (PTSD) inducing event. Self-esteem has also been shown to respond positively to simple computer games, using smiling faces and a task that directly links these to a user's personal data (Dandeneau & Baldwin 2004). However, different mechanisms to distraction may be at play here such as social processes.

2.2.2 VR & VE Psychotherapy

David Coyle's work takes the computer therapy paradigm a step further by setting a CBT structure within a 3D gaming VE (Coyle 2009, Coyle *et al.* 2009). The work is aimed primarily at adolescents who, it is believed, would be more likely to engage with such a format rather than dry text. Others have also looked at integrating VEs and cognitive approaches to therapy (Przeworski & Newman 2006, VEPSY 2009).

Further benefits may be gained from using a VE in this way. For instance, a sense of immersion within a virtual environment may be achievable, depending on the hardware platform: Hoffman's 'Snow World' is a computer game that is designed to distract from the painful procedure of periodic wound dressing, experienced by some burn victims (Hoffman *et al.* 2001). Patients are immersed, via a head mounted display (HMD), into an icy world where they have to engage in various activities intended to distract attention away from the sensations of the real world. Although not directly a mental health intervention, an indirect benefit may be gained: it is understood that there is a high co-morbidity of depression and the chronic experience of pain (Korff & Simon 1996), and that causal paths may run in both directions. Cognitive factors leading to depression in burns victims, such as the anticipation of treatment pain, may be lessened through use of Snow World and its ilk. We would then expect to see longer-term affect benefits although it currently does not appear to be the case that such a study has been undertaken. A further point of note is that, as with the 'Tetris' example above, the content of the Snow World VE may be incidental to the effect. The idea that images of snow, implying cold, can be used to counteract burns pain, implying heat, may be a pleasing one. However, any distracting content might work just as well. Counter to this, though, we could argue that content implying heat may well remind some patients of their injuries, so bringing attention back to the painful procedures happening 'outside' of the VE.

Aside from engagement and distraction, are there any other benefits that could arise from using immersive VEs as a tool for psychotherapeutic intervention?

2.2.2.1 Advantages of VEs for Psychotherapy

When asked about the advantages of using immersive VEs for therapeutic intervention we might reply in terms of the accurate repeatability of the stimulus, agents never tiring etc., or other forms of control that may be had over the content and subsequent experience of the user. However, standard arguments of this type have two problems: to begin with, they could apply to many other forms of stimuli, without necessarily providing a unique justification for the use of VR. The second problem, linked to the first, is that such arguments are apt to miss a genuinely useful aspect of immersive VEs: presence.

Arising from Minsky's early conception of 'telepresence' (Minsky 1980), it has been argued that the notion of presence has become a little muddled (Slater 2009). However, recent work has lead to the resolution of presence into two orthogonal factors that may be of much use, from an intervention engineering perspective. It has been suggested that 'presence', as applied to VEs, can be resolved into two components: Place illusion (PI) and plausibility (Psi) (Slater 2009). PI, the sense of being in a place depicted by a VE, is a robust effect and is considered a product of the sensory motor contingencies supported by an immersive system: if the immersive system correctly matches the appropriate contingencies to the particular sensory modalities, PI will follow. In other words, if I turn my head to the right, within a VR world, the scene should swivel in an expected way to depict correctly that part of the world on my right.

Psi is given as the extent to which participants believe that what is happening in the VE is really happening. It is antagonistic to a participant's belief that they are witnessing a non-real environment. Factors that are thought to be important for Psi in an immersive VE include realistic illumination and the use of a responsive virtual body (Slater *et al.* 2010). The VE must also respond to the participant's actions as if real. Work has been carried out to establish that participants will respond realistically within a particular VE, provided that Psi is high and PI exists (Slater 2009). The presence of a responsive virtual body, which the VE responds to realistically, is considered an important factor: here, PI and Psi can combine, leading to realistic participant behaviour (Slater *et al.* 2010).

PI can be a very effective way to block out distraction and could lie behind the success of Snow World (Hoffman *et al.* 2001, see above) and other distraction tools. If basic sensor processes are fully engaged with information from the World then other, *incongruous* data, such as pain signals from wound dressing, might find it harder to integrate themselves into conscious experience. This may also open the way for increased engagement of higher cognitive processes, for instance allowing a person to become more easily 'hooked' by the game, thereby further screening out 'exterior' pain signals.

PI may make VR, regardless of content, strongly distracting and engaging when compared to other stimuli. Even if we accept this, however, it still does not give immersive VEs a unique justification for therapeutic use. For this, we need to turn to Psi. Belief that the event encountered within an immersive VE is real could be extremely useful to psychotherapeutic intervention. Examples of this are the use of VR for the treatment of various phobias (Carlin *et al.* 1997, Klein 2000, Mühlberger *et al.* 2003, Mühlberger *et al.* 2006) or PTSD (Rizzo *et al.* 2005, Difede 2002). These kinds of intervention may be relying on the effect that at some neural level the content of the VEs are processed and responded to as if real. If they were not then this sort of exposure therapy would not work. Content can be controlled and presented gradually to increase the perceived threat to a point qualitatively comparable to a real life encounter. An extremely fine grain of graded exposure may not be realistically achievable with other stimuli types such as photographs or stock video. We might argue against this, though, giving the example of a television screen being set into a wall. Participants are told that this is a window and that the events depicted on the screen are actually happening beyond the wall. However, it would very quickly, if not instantly, become obvious to the participants that the screen was not obeying several motor contingencies or, for those looking with two eyes, stereoscopic depth. Belief would be lost along with any ‘perceived exposure’ benefits.

It may not even be necessary to extinguish all disbelief in order to work. The participant would not necessarily have to be completely fooled, if this were even possible, into thinking that they were no longer experiencing VR when they were. Realistic response behaviour within an immersive VE has been recorded many times, even in graphically limited stimuli (Pan *et al.* 2008, Slater *et al.* 2006). There is even some debate as to the fidelity of graphics required. To create PI, at least, even relatively low fidelity worlds can be employed (Zimmons & Panter 2003) although higher realism can be helpful (Meijer 2009, Slater *et al.* 2009). Psi may be similar (Garau *et al.* 2005) although a more complex relationship between realism and response may exist. For example, some work shows that the realism of a virtual character’s rendering must match the realism of its behaviour, in order to be responded to realistically, however these variables are measured (Garau *et al.* 2003, Vinayagamoorthy 2004). The very idea of belief may itself need to be effectively resolved into functional components before this debate can progress. It may be that activity within disparate parts of our neural anatomy imply differing levels of belief, meaning that realistic response behaviour from a participant does not imply an overwhelming belief in the reality of the virtual experience, merely that a part of the brain is responding as if the events were real. This would explain why VR exposure therapy, to treat arachnophobia for instance, could work despite the receiver knowing that they are only experiencing a virtual spider. This may also explain why participants can behave in an analogous but dampened fashion when comparing reactions within a VR experiment modelled on the infamous Milgram experiment (Slater *et al.* 2006) with actual study (Milgram 1963). The latter study confounded a kind of visceral belief with cognitive belief that the former study was able to separate.

A further benefit derived from Psi can be had when regarding transferability of effects. Within learning research, near transfer and far transfer are phenomena encountered when testing the effects of stimuli taken out of their immediate experimental context (Nokes 2009). We can hypothesise that the effects produced by a stimulus in one scenario many then be observed when a participant is presented with another, and test accordingly. For example, a patient may be proscribed a computer-based, text driven intervention for social anxiety by their GP. We could, for example, ask whether any 'progress' (however measured) made within the program itself were transferable to later, online, text-based encounters within a social network. Under controlled conditions, we might ascribe similar levels of progress to the transfer of effect from program to online encounter. This would be an example of near transfer as the nature of the stimulus and subsequent test are very similar. Far transfer effects can be valuable from an intervention point of view as they place fewer limits on subsequent use. We can, in effect, measure how transferable an effect is from one context to another by calculating relative effect sizes, although whether near or far transfer has taken place is perhaps more loosely defined. It is important to note here that transferability in this context does not mean that the same effect can be encountered in many other paradigms or using many different types of stimuli. Here, it means whatever the effect changes for a participant in one situation, those changes will carry through into another: it is the difference between those situations that will indicate the label ‘near’ or ‘far’. As such, transferability can be seen as a learning phenomenon (Nokes 2009).

It can be argued that Psi, as encountered within VR, has far transfer properties and that this is similar to the way in which a real life encounter might have a lasting effect. Learning to be calm on encountering a virtual spider means it is highly likely an arachnophobe will be calm seeing a real spider in much the same way that learning to be calm encountering a real spider can help an arachnophobe be calm on seeing a later spider. The same cannot be automatically applied to other stimuli types such as photographs of spiders for instance. As previously noted, a VE experience need not be identical to its real life counterpart in order to elicit similar behaviour although PI fidelity and matching Psi to expectations, based on PI, may be important to elicit the similar *levels* of behaviour to those seen in real life situations. Matching an agent’s expected behaviour to its level of fidelity is one way, already noted, of meeting the latter criterion. Of course, if a stimulus is completely life-like then exposure effects will inevitably be the same as an actual exposure. It may be difficult to define quite what the term ‘completely life-like’ means or how to measure it. However, what actually matters here is what aspects of the real world need to be modelled within a VE in order to effect good transferability to the real world (see Leberman *et al.* 2006, Hertel & Mathews 2011 for discussions on transfer effects).

2.2.2.2 Limitations of VEs for Psychotherapy

There are limitations that may lower the effectiveness of using a VR experience as an intervention tool. A break in presence describes an effect whereby the VR collapses for an individual and they merely experience a VE, perhaps also noticing other distracters in the laboratory. It has been shown that PI component is relatively robust, easily created in the first instance and easily restored if broken (Slater 2009). Visual PI appears to dominate other sensory modalities (Burns *et al.* 2005). This is perhaps fortunate, as other modalities, with the exception of audio, can be difficult to model with any dynamic accuracy. Haptics, for instance, can currently only provide an approximation of the physical sensations encountered in the everyday world (Saddik 2007) despite exciting future promise of such techniques as ultrasound haptics (Iwamoto *et al.* 2008). PI, though, can remain robust, despite this. Psi , appears to be different, however. It may not be so easy to achieve strong Psi and, once broken can be very difficult to restore (Slater 2009). Indeed, creating an extremely plausible VR illusion, indistinguishable from reality, is beyond our current ability. It is likely that a patient would always be aware that they were experiencing VR, possibly lowering treatment efficacy for some intervention types.

Even supposing (extraordinary) lengths could be gone to in order to deceive someone into believing that they were experiencing a real situation when they were not, ethical questions would be raised, perhaps preventing widespread use of such techniques. However, behaviour is not automatically dampened just because it is generated within VR. The interaction interface can go some way towards increasing believability. For example, motorists within a driving simulator that uses an actual car as an interface will often display spontaneous responses comparable in type and level to those found with actual driving behaviour and such simulators can be regarded as an effective training aid or driving behaviour prediction tool (Pasetto & Barbati 2011).

Finally, it should be pointed out that the above matching behaviour to belief is based on an assumption. If behaviour within an immersive VE matches that found in a real life counterpart, both in terms of type and level, then it is assumed that the VR experience carries the same level of plausibility as the real life situation. This should be open to question as it is possible for some neural processing to occur within the VR session that does not affect immediate behaviour but nevertheless impacts negatively on treatment efficacy. Of course, a reverse, positive effect, is also possible. Whether such impacts were down to Psi levels of not would need to be investigated.

2.2.3 Further possibilities for Virtual Therapies

The increasing graphic and animation capabilities of personal computers, somewhat driven by the video gaming industry, provide much promise. Whilst there may currently be a difference between behaviour in fully immersive, CAVE(tm)-like VEs and that on basic desktop systems (Pan & Slater 2011), there is an interest in the possibilities that lower end solutions provide. Health authority budgets may not allow for high-end platforms such as CAVE(tm)-like facilities or power walls within hospitals. However, low cost, desktop systems can still deliver an immersive experience via HMDs incorporating stereoscopic, head tracked vision and lifelike audio. If hand or gesture tracking is also incorporated then PI and a strong Psi may be had by giving some thought to the virtual scenario. If the user remains seated or standing in one position, for instance, a full body, seated or standing avatar representing the user would provide a good point for PI and Psi to meet, strengthening overall presence. Virtual reflections would also modulate Psi positively (Slater *et al.* 2010) as would real-time, global illumination and correct dynamic shadowing (*ibid*). The door is therefore open to all manner of complex virtual interactions and situations.

Virtual exposure treatment for various phobias has been in use since the early 1990's (Roy 2003) and continues to this day. Research into the use of VEs has also expanded to other areas such as social interaction (Pan *et al.* 2008, Klinger *et al.* 2006). Realistic, social interactions of the sort that may be useful for tackling various forms of social anxiety or depression are now possible on desktop/HMD systems. The software, again driven somewhat by the gaming industry, is similarly within reach of modest budgets. Immersive desktop systems, such as those used by various military bodies to treat PTSD (Rizzo *et al.* 2005), are becoming affordable for widespread deployment in GP surgeries. It is conceivable that at some point virtual therapies may be integrated alongside a current standard treatment for affect disorders: psychopharmacology and CBT combined (Layard *et al.* 2006). Delivery of content onto patients' own, home systems may also be desirable but it is expected that there would be a great degree of variability in any available hardware.

Much of the expanding field of VE and VR therapy research relies on some form of simulation. The simulation of real world therapeutic processes includes exposure techniques for physical phobias such as arachnophobia or aerophobia (Carlin *et al.* 1997, Klein 2000, Mühlberger *et al.* 2003, Mühlberger *et al.* 2006), post traumatic stress disorder treatments (Rizzo *et al.* 2005, Difede 2002), social disorder psychotherapy (Glantz *et al.* 1996, Anderson *et al.* 2003), social anxiety (Klinger *et al.* 2006) and eating disorder treatments (Gorini *et al.* 2010, Ferrer-Garcia & Gutiérrez-Maldonado 2009, Fox *et al.* 2009, Ferrer-García *et al.* 2005). Some VE therapies attempt to simulate social processes in order to enable clients to achieve new learning akin to cognitive therapies (Coyle 2009, Coyle *et al.* 2009).

With VEs, there is scope to examine and utilise not only the similarities between a VR experience and reality but also the differences. In other words, what can we do in VEs that we cannot realistically do in real world settings? Perhaps there are ways in which a simulation might be crafted which, although not simulating the real world directly, might nevertheless amplify therapeutic gains. Some work can already be seen as addressing this, such as Hunter Hoffman's Snow World, but there may be much untapped potential.

A common problem acknowledged within the therapy profession is sustaining engagement (Beard *et al.* 2011). Treatment programmes can have high dropout rates (*ibid*). VR simulations can be written which use similar techniques to the games industry to 'hook' players into returning to the simulations repeatedly. These techniques were borrowed in turn from behavioural psychology and concern reward schedules (Ferster & Skinner 1957). A variable frequency reward schedule describes a ratio of action to reward in which the number of actions needed before a reward is given varies. This schedule can prove just as irresistible to humans as it did to the original pigeons that experienced it (*ibid*). Reward schedules and other gaming constructs could be built into simulations to encourage repeated use therapeutic tools. For example, one study utilising CBM found that participants considered the task 'boring' and the study authors attributed the dropout figure in part to this (Beard *et al.* 2011). The study, to modify attentional bias away from negative expressions and towards positive ones, utilised a simple paradigm. Participants were instructed to look at two faces on a screen, which would promptly disappear leaving a probe behind either one. The probe would be either the letter 'E' or 'F'. Participants then had to indicate, via a keyboard, which of the two letters was displayed. Reaction times speeded up as the experimental condition participants realised that the letter always appeared under the positive expression. A battery of tests followed to test for social anxiety against controls. Results showed an effect that indicated that introducing a cognitive bias, scanning for positive expressions, affected affect, lowering of social anxiety. However, the results were modest, due, the authors suggested, to the nature of the task. The task was simple, repetitive and, it could be argued, dull.

Embedding the CBM task within a gaming scenario may have improved results. VR simulations are capable of such embedding. Further improvements to a CBM tool might be had by the high Psi that VR can generate. A hypothesis could be given: the more realistically rendered and animated the faces in the above task are, the lower the social anxiety score. However, using current techniques, participants would be fully aware of the manipulations and, as such, implicit processes may not be targeted.

One approach that addressed this can be found within the psychophysics of perception. Visual masking may be of help in disrupting the non-conscious processes that are believed to underpin some forms of affect disorder (Wells 1997, Clark 2001). Visual masking describes a phenomenon within psychological research whereby the rapid presentation of one image can be used to disrupt the conscious awareness of a spatially adjacent or concurrent image (see section 2.3 below). There is some evidence, however, that the masked image is processed, albeit in a non-conscious way (see section 2.3.5 below).

Further possibilities for the use of VEs in psychotherapy present themselves when we consider NATs. According to a cognitive appraisal model of emotion (Fox 2008), NATs are seen as primary cognitions, judgements of situations that trigger affective responses (Wells 1997). Within this model, a person with an automatic bias towards negative judgements would be more likely to judge a given situation or themselves negatively. It is reasoned that such processes form an important part of many, if not all, types of affect disorder. For example, a social anxiety sufferer might be more inclined to view certain social situations, and elements within them, as threatening as well as ruminating negatively on their 'performance' subsequent to such encounters (Clark 2001). This, in turn, would lead to an automatic lowering of expected performance for future social encounter, lowering self-esteem alongside attendant increased negative affect. This vicious circle is completed, it is suggested, as the negative feelings foster an increased likelihood to view such future encounters in a threatening light (*ibid*). Two underlying assumptions are that cognitions have a causal relationship to affect and that affect itself can have a causal effect on cognition.

As noted, CBT purports to unearth and challenge NATs as well as the unconscious attitudes (core beliefs) which NATs are seen as being derived from (Neenan & Dryden 2004). Whilst CBT is currently seen as the therapy of choice from an evidence-based perspective (Layard *et al.* 2006) like all therapies, there is a limit to its efficacy, possibly running at about a 50% success rate (*ibid*). CBT may not necessarily target non-conscious processes, remaining instead on a purely conscious level. Visual masking may be able to reach the non-conscious processing level on which NATs reside. If positively valenced, masked images could be presented in such a way as to increase the plausibility of those images, i.e. their Psi, then it follows that an automatic disrupter of negative NATs may be learned. Here, affect would be the primary causal agent (Dolan 2002, Zajonc 1980) creating a change in cognition. Embedding visually masked images within a VE might be one way to increase their plausibility. The images would need to be rendered as actual objects in order to increase their plausibility; if, for example, the image of a face with a positive expression were to be used, then the face would need to be rendered as a believable, non-planar object.

Finally, using standard subliminal stimuli, as opposed to visually masked stimuli within a VE, may not be effective for disrupting NATS. There is some controversy surrounding the ability of subliminal visual stimuli to be 'read' at all by the brain (Moore 1992). In order to render them effectively invisible, the subliminal stimuli must be displayed so rapidly that poverty of stimulus becomes a factor. This may explain why techniques such as subliminal advertising do not appear to deliver observable effects (Weir 1984). Masked stimuli are also displayed rapidly. However, they are displayed for longer periods than standard subliminal stimuli. They are therefore more likely to be (implicitly) perceived for longer, leading to measurable effects. In fact, the lengths of time often used to mask a target in priming experiments (see section 2.3.5) are more than enough for a target to be visibly identified if no mask is present. During World War 2, a psychologist, Samuel Renshaw, working for the American Air Force devised a system to train pilots to recognise aircraft rapidly. The system used briefly presented images in a tachistoscope (Vicory 1968). It was claimed that most trainees could achieve a high success rate using images shown for as little as 1/75 of a second, below that used by many researchers for masked priming.

It follows that, for a target to be exposed long enough to impart any meaningful, perceptual information, masking of some form may need to be present. Regarding the kinds of virtual therapy intended to disrupt NATs without patient awareness: if such perceptual information were not available then it would be hard to envisage how any stimulus could convey either PI, Psi or begin to disrupt NATs at all. Therefore, in order for such a virtual therapy to be effective, some technique needs to be employed whereby a stimulus can remain perceivable long enough to impart the necessary information but not engage conscious processes. One immediate possibility would be to use continuous flash suppression (CFS). With CFS, images are rapidly and repeatedly presented to one eye in order to suppress the explicit perception of an image presented to the other (Bahrami *et al.* 2007), see 2.3.3 below. However, there are drawbacks to this technique, not least that it would prove difficult to build a convincing VE around it: stereoscopic vision would not be possible using CFS as it stands, for instance. Another, more promising approach, considered next, would be to use visual masking which would allow for the construction of a VE along with the inherent advantages discussed above.

2.3 Visual Masking

2.3.1 Introduction

This section will explore the phenomena of visual masking in some depth, both through a general taxonomy and then with a deeper examination of parameters and their application to masked priming. This in order to understand not only how visual masking works, as far as current knowledge allows, but also how it can be used within a VE to create non-consciously perceivable stimuli. The complex relationship between visual masking and attention will also be considered before returning to masked priming from the perspective of studies using facial expressions. Overall, this section will show current research gaps that will need to be addressed in order for visual masking to be made to work in a VE, and what form it should take. Important parameter issues will be discussed such as how depth disparity internal to virtual objects could affect the masking effect.

Masking is a perceptual effect whereby the presentation of one stimulus, the mask, affects the conscious perception of another, the target (Breitmeyer & Öğmen 2006). Masking phenomena have been reported for various perceptual modalities but a large body of research has arisen around visual masking. Since scientific reports of visual masking began, about a century and a half ago, variations on the form have burgeoned (Ansorge *et al.* 2007, Enns & Di Lollo 2000). Visual masking is studied using many varieties of words, light fields, visual noise, graphic figures and imagery as stimuli (Breitmeyer & Öğmen 2006). The neurobiological correlates of visual masking have also received much attention (Kammer 2007, Macknik & Martinez-Conde 2007, Whalen *et al.* 1998, Rolls *et al.* 1999). As noted at the end of section 2.2.3, visual masking holds the promise of being able to work on the level on non-conscious or unconscious processes and as such may have a disruptive effect on NATs or similar processes that are thought to stabilise affect disorders.

Some commentators object to the use of the term 'unconscious' when building explanatory models for visual masking data (Holender & Duscherer 2004). The issue centres on threshold models (Dijksterhuis 2005), which generally described an objective and a subjective threshold of perception. The objective threshold related to the physiological properties of the sensory system: a stimulus had either sufficient energy to register on a sensor or it did not. The subjective threshold held that, once past the objective threshold, a stimulus-derived signal had to be of a certain form to enter consciousness. If a stimulus passes the objective threshold but its derived signal does not pass the subjective threshold the implication is that subliminal perception has taken place. Some commentators (Holender& Duscherer 2004, Holender 1986) argued that it could not be reliably shown whether a signal had passed the subjective threshold or not and so subliminal perception could not be empirically demonstrated. Therefore, a definitive subjective threshold does not exist. For example, a person may be presented with a rapidly displayed image, which they claim not to have seen but that may have a priming effect (see section 2.3.3 below). Could it be that the image did enter their consciousness but they simply forgot (Dijksterhuis 2005)? This would be an example of Holender & Duscherer's (2004) 'single-process conscious perception'. A difficulty with this model is that it eliminates awareness from consciousness. A subjective awareness threshold would not only be a limit to subliminal processing but would be a limit to conscious processes as well. If this is removed then we are forced, as Holender & Duscherer are, to say that every sensory signal enters consciousness.

However, we are clearly not aware of our endocrine regulation processes, for instance: few if any of these signals enter the hippocampus, temporal lobes or other memory-associated areas but remain within the brainstem/mid brain areas (Ulrich-Lai & Herman 2009, Brunicardi *et al.* 1995). Therefore, Holender & Duscherer are forced to disassociate consciousness from awareness. Their argument becomes reduced to the level of a semantic game where afferent stimuli signals enter consciousness but not necessarily awareness. Inadvertently, they also introduce a new subjective pseudo-threshold, between consciousness and awareness.

Aside from the above reductio ad absurdum, Dijksterhuis *et al.* (2005) provide a counter to anti-threshold arguments. They appeal to signal detection theories which have shown that stable individual differences, contextual effects and other factors do predict whether a stimulus signal will reach conscious or not. It seems that the subjective threshold can be modelled but that it is a complex construct, not just a simple limit.

Perhaps a more serious charge that could be levelled against threshold theories is that they muddle subjective and objective elements together without really explaining how they could interact in a physical sense. Further examination of this falls out of the scope of this thesis. The terms ‘explicit’ and ‘implicit’ can be used to sidestep this debate. In this thesis the terms ‘conscious’ and ‘unconscious’ will be used interchangeably with ‘explicit’ and ‘implicit’ respectively.

Finally, an in-depth exploration of the explanatory models of visual masking in general will be largely avoided as this too is deemed to be beyond the scope of this thesis. Some theoretical exposition, though, can prove helpful when presenting hypotheses or interpreting results and will therefore be included.

2.3.2 Visual Masking Taxonomy

Visual masking has been studied in its own right and has also been used as a tool to study other research domains such as perception, attention and consciousness (Ansorge *et al.* 2007). This distinction is not always clear in practice though. If visual masking does arise because of the neural structure of our visual apparatus then one would necessarily expect insights gleaned from its study to impact upon other areas of vision research and beyond. This is indeed the case (*ibid*). The subject area of visual masking, large in its own right, has blurred with other areas. Following this, inevitable vocabulary confusion has arisen. Take, for example, priming studies. Visual masking is sometimes used to facilitate implicit priming. Here the terms 'prime', 'mask' and 'target' describe three elements of the overall stimulus. The prime is that figure, word or image that will not be experienced consciously. The 'mask' will be that element which blocks the conscious perception of the prime. The target will be a subsequent probe figure, word or image to which the participant is expected to consciously respond, possibly biased by the implicit perception of the prime. However, in visual masking studies per se, the vocabulary is slightly different, which may be a source of confusion. The term 'mask' remains the same for both fields but in pure visual masking research the term 'target' refers to that figure, word or image that is blocked from explicit perception by the mask. It is therefore equivalent to the prime from masked priming studies. To reiterate, in masked priming the term 'target' is used for that which is explicitly perceived whilst for visual masking the term 'target' is used for that which is made explicitly imperceptible.

This thesis will use the word ‘target’ in the sense that it is used within visual masking research i.e. that which is explicitly imperceptible. When describing a masked priming paradigm the term ‘prime’ will be used to describe the same element. The term ‘mask’, having a usage common to both, is unproblematic and so will be used as is. However, when describing masked priming studies the term ‘target’ will be replaced with ‘target probe’ or merely ‘probe’, the latter being in common use within vision research.

As another note on terminology: the phrase ‘visual masking’ itself carries some confusion. Within computer graphics, the term has been used to describe a surface rendering effect. Here, unwanted surface artefacts on a virtual object, such as banding, can be disguised using noise texturing techniques (Ferwerer *et al.* 1997, Sayer *et al.* 2004). Although this also involves obscuring percepts, it should not be confused with the term ‘visual masking’ used in this work.

‘Pattern masking’ describes visual masking which uses graphic figures, imagery or visual noise to form masks or targets. Often the target and mask are of similar form and indeed their similarity or otherwise has been the subject of much study (Breitmeyer & Öğmen 2006). Generally, the case is made for masks that resemble their targets producing stronger effects. However, the negative compatibility effect (NCE) occurs when a stronger inhibition of target visibility arises in conjunction with target and mask dissimilarity at certain parameter values such as longer stimulus durations (Schlaghecken *et al.* 2007). NCEs may be related to attention and would seem to suggest a target-form analogue to certain other spatial attentional capture effects (Posner & Cohen 1984), see section 2.3.4 below.

The mechanics of visual pattern masking are straightforward. The mask is a figure or image which is either spatially adjacent, overlapping or fully obscures the target in the visual field. As well as spatial parameters, temporal factors can be altered such as the mask being presented before, after or concurrent with the target: known as forward masking, backwards masking and common onset masking respectively.

Finer grains of both spatial and temporal delineation can be added. A continuous spatial variable can be derived from the relative positioning of the mask and target where gaps between target and mask can be decreased to zero, then further to overlap and finally on to one figure or image completely obscuring the other. One form of visual masking that is widely studied, lateral masking, uses the first extreme of this continuum, the target and mask lying adjacent to each other. Masked priming studies, on the other hand, have tended to use masks that fully obscure the primes.

A continuous temporal scale can be produced by varying the onset of target and mask, known as stimulus onset asynchrony (SOA). Other time variables include the length of exposure of each element, the gap between the disappearance of one element and the appearance of the other and also the time to the next stimulus pair. This latter temporal gap is interesting as it introduces the possibility of subsequent targets acting as masks for preceding masks. This has been exploited with a technique known as rapid serial visual presentation (RSVP) whereby a string of stimuli are presented in rapid succession (Shapiro *et al.* 1994). This produces an effect known as sequential blanking. Linguistic masking commonly uses this approach to study implicit semantic and lexical processes, as well as attentional blindness (Giesbrecht & Di Lollo 1998, Shapiro *et al.* 1994). A form of sequential blanking had also been observed much earlier in masking research, using arrangements of lit apertures in black spinning disks

(Breitmeyer & Öğmen 2006, p15). That said, under some forms of lateral masking introducing a third element, this time a second mask, can lead to the recovery of target visibility (*ibid*, p255).

The term 'inter-stimulus gap' (ISG) can lead to some confusion, sometimes being used to refer to a temporal variable, at other times it is used to refer to a spatial relationship. In this thesis I will adopt a terminology used by other researchers wherein the 'inter-stimulus interval' (ISI) refers to a temporal variable with ISG referring to a spatial one.

Varying the time and spatial arrangement has led to a plethora of pattern masking paradigms. As noted, one such is lateral masking. Here the mask and target are presented adjacent to each other with a positive ISG. Temporal variance has given rise to the emergence of two well-studied versions of lateral masking: paracontrast and metacontrast masking (Breitmeyer & Öğmen 2006). For the former, the mask precedes the target in time whilst a metacontrast mask follows the target. We can therefore see that paracontrast is a type of forward masking with metacontrast being a form of backwards masking. A common form of mask and target for both of these is shown in figure 2.1, below.

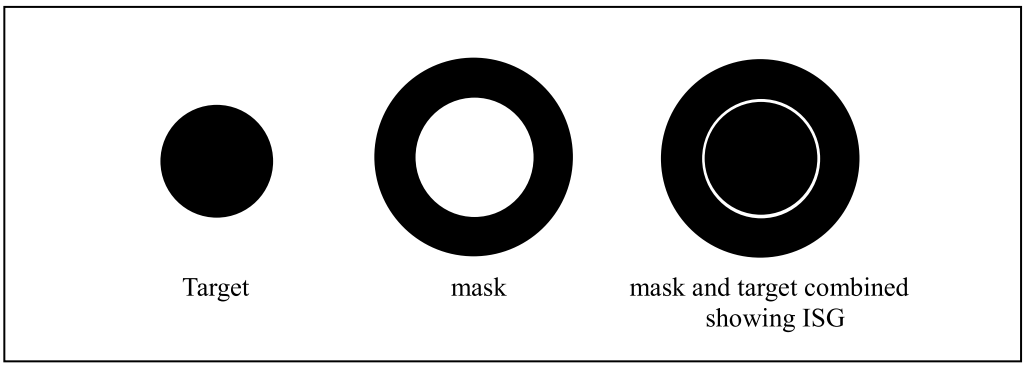


Figure 2.1. Common target and mask forms for paracontrast and metacontrast masking (from Breitmeyer & Öğmen 2006).

In order to achieve a paracontrast effect, for example, the masking annulus could be presented on a display screen followed by the target disk with an SOA of anywhere between 600ms and common onset (0ms). Optimal results would usually be obtained with an SOA of 200ms to 0ms. However, paracontrast is often a weaker effect than metacontrast. Metacontrast effects will often be strongest with an SOA of 60ms

The picture is further complicated by other visual factors such as target/mask contrast, colour, contours and emitted light energy. For instance, when contrast judgements are a factor then paracontrast SOA will be optimal between 200 and 100 ms whereas contour discrimination will yield a better SOAs between 50ms and 0ms (Breitmeyer & Öğmen 2006, p43). For metacontrast, target contrast is suppressed at around 10ms whilst contour suppression can take place at around 40ms SOA (*ibid* p45).

Depending on the experimental paradigm used, lateral masking will often produce what is termed a Type B schematic, see figure 2.2, below.

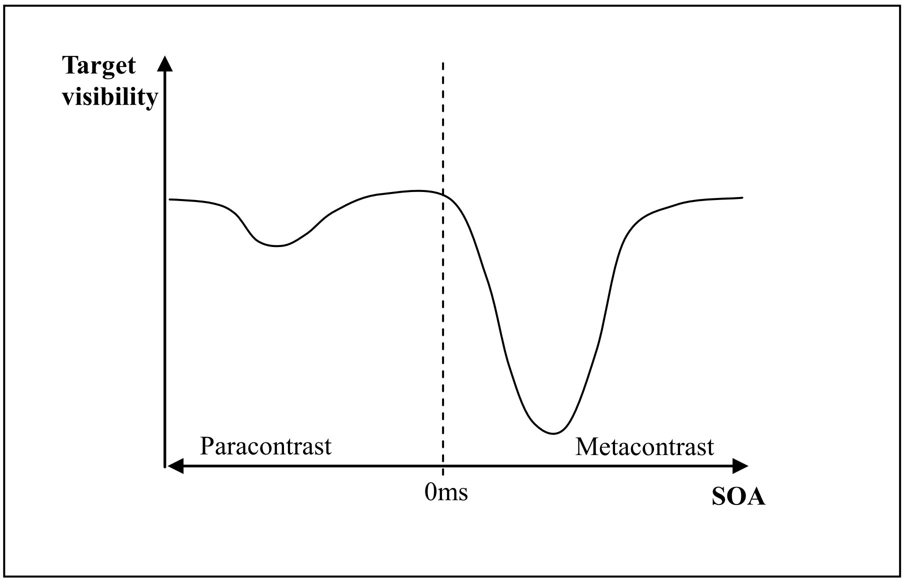


Figure 2.2. Type B profile (from Breitmeyer & Öğmen 2006)

Several features are of note here. Firstly, paracontrast effects are less capable of yielding target invisibility than metacontrast effects (Breitmeyer 1984, Breitmeyer *et al.* 2004b). In addition, at common onset (0ms) target visibility is as high as at the extremes of SOA where masking effects break down. Another common profile is Type A, see figure 2.3 below. Here each side of the common onset point monotonically mirrors the other. This 0ms SOA point is optimal for masking.

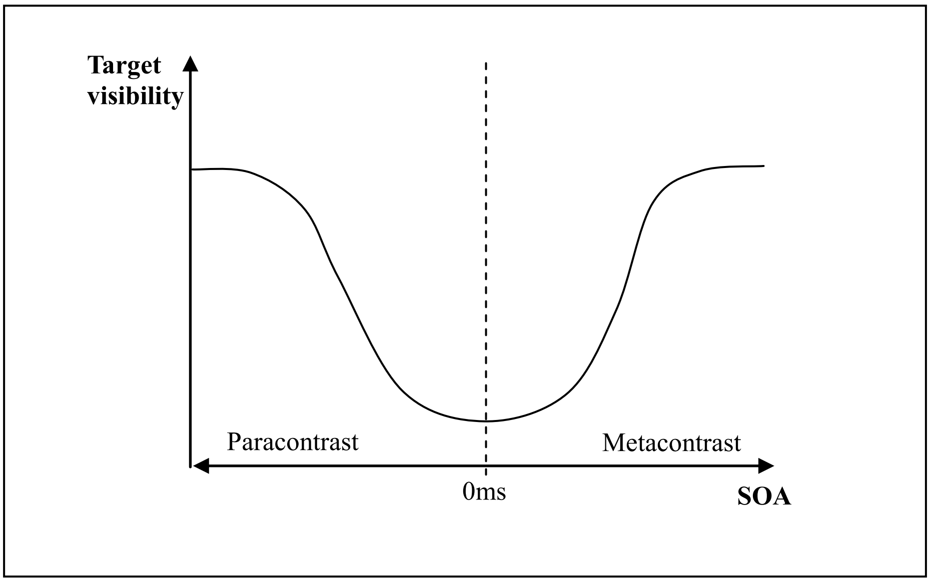


Figure 2.3. Type A profile (from Breitmeyer & Öğmen 2006).

Kahneman's 'Onset-Onset' law states that SOA is the key factor in producing a metacontrast effect (Kahneman 1967). Other factors, such as stimulus duration, whilst important, are not considered as critical. To an extent, this has influenced later masking research, whether using metacontrast or not. However, the emphasis on SOA or temporal aspects in general is not without its critics (Herzog 2007, Hermens & Ernst 2007). In particular, the stimulus termination asynchrony (STA) appears to be important also (Macknik & Martinez-Conde 2007) and, with masked priming, the stimulus form/complexity and duration are important (possibly interacting, see section 2.3.5 below).

Other forms of pattern masking include non-lateral masking by structure and masking by noise. The former uses graphical figures as masks that often fully overlap the target. For instance, if the target is a letter then the mask may be composed of an overlapping, jumbled series of that same letter using the same typeface, font and weight. This would produce graphic similarities between target and mask that could strengthen the masking effect or inhibit it, as per the NCE, depending on the parameter values used in that instance.

Masking by noise is simply a paradigm that uses a mask comprised of graphic noise. Targets can be a coherent, graphic figure such as a common shape or letter. Thus, the target will be dissimilar to the mask in form, reversing the points made about masking by structure, above. A related paradigm, where both mask and target are composed of graphic noise, can be found in the study of depth effects within visual masking. Here, random dot stereograms (RDS) are used to generate both target and mask with the illusion of a depth disparity between them (Lehmkuhle & Fox 1980). The stereoscopic image therefore allows depth to become yet another spatial variable to be manipulated. RDSs can control for monocular perspective cues in a similar way to stereoscopic images that separate eye movement convergence from focus accommodation (Wickens 1990). Increasing RDS depth cues tends to result in the inhibition of masking (Lehmkuhle & Fox 1980), known as 'binocular unmasking' (De la Rosa *et al.* 2008, Moraglia & Schneider 1990) although Type A effects have been shown with cyclopean backwards masking (Breitmeyer & Öğmen 2006, p58).

2.3.3 Parameters and Priming

Depth as a parameter has been studied using other arrangements of stimuli. Gratings such as Gabor patch figures (Moraglia & Schneider 1990) occur in this research, as do planar image masks (De La Rosa *et al.* 2008). Returning to lateral masking, another way in which depth has been studied within the field of visual masking has been to locate the target within the figure of a particular planar mask, each mask design suggesting a degree of depth using graphic, linear perspective techniques (Weisstein *et al.* 1982). For example, the mask could be a two dimensional, orthogonal, perspective drawing of a cube and the target one of the lines that make up the cube design. In this variety of structural masking, both target and mask are planar and coplanar. The impression of depth is created entirely by the mask design. By varying the mask design a greater or lesser impression of depth can be created, or eradicated entirely: the so-called 'object superiority effect' (OSE) (Weisstein *et al.* 1982). Participants are asked to indicate which line of the visible mask was the preceding target. Again, the results show a tendency towards depth, or apparent depth, inhibiting the masking process although in this case a form of *monocular* unmasking appears to be taking place.

One factor which all of the above depth cue paradigms have in common is that the masks and targets are all planar, even if coplanar or non-coplanar. There appears to be a lack of research using masks or targets with internal depth disparity within a stereoscopic display. Virtual objects within a stereoscopic environment would seem ideal candidates to test whether or not masking would be inhibited within a non-planar context.

Planar context can be defined by the masking pattern that a lateral target is embedded into. Planar contexts can sometimes act as a noise source, disrupting the object superiority effect (Weisstein *et al.* 1982). For instance, a mask may be complex but not display a robust, object-like appearance, based on the rules of linear perspective. Such a mask would then be expected to deliver decreased target visibility. This should be less of an issue for non-lateral pattern masking as the target is not seen as part of the masking form although the related 'feature migration' and 'shine through' effects may have an impact. Feature migration (Enns 2000) describes an effect whereby a fully obscured target may nevertheless transfer some of its characteristics onto the visible mask percept. However, it follows that the closer a mask and target are in visual form the less of an issue this will be: would–be, transferred features may already be present in the mask rendering feature migration less important. Similarly, 'shine through' (Herzog 2007), an effect where mask and target features become blended, should also be rendered less significant an issue with increasing target and mask similarity. It can also be argued that decreasing evidence of feature migration and shine-through should be proportional to increased visual complexity of the mask, particularly for non-lateral, pattern masking. If cognitive processes are involved in masking then there may be a limit to this complexity: a highly complex noise mask may not be as engaging as a complex geometric design for example (Flake 1998, p135), thereby lowering masking potential.

Masking-by-light is achieved by using contour-less flashes of light emanating from flat visual fields such as those produced by photographers' light boxes. Using this paradigm, it is also known that increasing the visual complexity of the target has some effect (Breitmeyer & Öğmen 2006, p26): increasing target complexity decreases recognition or identification performance (Hermens & Ernst 2007, Hermens *et al.*

2008), even in cases where masking itself is weak. The strength of the effect is inversely proportional to the ISI. From this, we can conclude that the more visually complex the target, the longer the ISI needed to extinguish recognition inhibition effects. What is not clear, however, is whether this is merely a sensory effect or if cognitive processes are involved.

Light energy levels emitted by stimuli are also considered important for masking effects. The emitted energy can be defined as the product of intensity and stimulus duration. During the early years of masking research, this luminance parameter was perhaps seen as more important than the respective wavelengths of light emanating from masks and targets and background. It has been shown that, for metacontrast masking in particular, background energy levels do affect masking. Lowering background intensity tends to be associated with lower SOA for peak masking (Purcell *et al.* 1974). Similarly, holding background energy constant, lowering either target or mask intensity will lower the necessary SOA although mask/target ratio also becomes a factor here. At ratios of less than one, the response profile tends to follow a Type B pattern, above this a gradual change to a monotonic, Type A profile is observed for backwards masking paradigms (Wiesstein 1972 from Breitmeyer & Öğmen 2006, p48). It has been shown that stimulus duration is an important factor in the energy product (Macnick & Livingstone 1998). The ratio of mask duration over target duration ($\Delta M/\Delta T$) is directly proportional to the mask/target energy ratio. As $\Delta M/\Delta T$ increases peak effective SOA decreases. We can therefore see a link between the parameters SOA and stimulus duration.

Chromaticity of mask, target and ground has been studied. This area can be usefully resolved into two sections: The sensory effects of stimulus wavelength impacting upon retinal rods and cones and also the cortical effects of metamer variance (Wyszecki & Stiles 2000). Again, using a masking-by-light form, studies have shown that rod and cone mechanisms interact (Breitmeyer & Öğmen 2006, p60): rod responses can be suppressed by both rod activating masks and cone activating masks, even when the target and masks flashed are designed to isolate either the rods or cones. For example, reversing the wavelength of target and mask flashes from green and red to red and green respectively lowered SOA considerably shifting peak effectiveness from metacontrast to a paracontrast form (Foster 1976). Here the short wavelength green stimulus was designed to isolate rod response, the longer wavelength red flash activated cones. A similar relationship has been shown to exist between the three wavelength categories of cones. As target and mask chromaticity diverge, metacontrast masking effects lower (Bevan, Jonides & Collyer 1970).

Colour also seems to be a post-retinal masking parameter. For instance, direct parameter specification (DPS) describes a situation where stimuli features appear to directly specify the response, without any intervening awareness cognition (Neumann 1990, Skalska *et al.* 2006). DPS is applicable to the topic of masked priming where responses are thought to be driven by specific characteristics of the prime. In the case of colour, Schmidt (2000) found that colour congruency between priming target and mask was important. Red and green primes and masks were used, interchangeably, on a mask colour identification task. Colour congruency significantly affected performance favourably, shortening reaction times, compared to colour incongruity. Further work appears to suggest that the masking effects of colour may be less dependent on percept/metemeric processing than on lower sensory/wavelength processing, possibly in V1 (Lamme 2004, Macnick & Livingstone 1998).

Colour priming has been shown to be wavelength dependant rather than percept dependant (Breitmeyer *et al.* 2004b). Desaturated blue, green and white disks were used as primes, along with desaturated blue and green masks. The white disk was intended to be a neutral prime. The main task was a standard, colour recognition reaction time test. The expectation was a standard colour priming result in that mask/target congruity would facilitate reaction speed whilst incongruity would impinge on it. This is what was found. However, the white disk did not act as a neutral prime would be expected. There white prime showed an incongruity effect towards the blue mask but a congruency effect towards the green. In other words, it was acting like a green prime. A further experiment tested for perceptual confusion of *visible* blue and green disks with white ones. The rational here being that, if white and green disks were consciously confused more than blue and white then that confusion was happening at a higher percept level. This would rule out sensory/wavelength dependent processes underpinning colour masking. What was found was the opposite: white primes were confused more with blue than green. This stands in opposition to the apparent white/green confusion for masked primes shown by the first experiment. The authors therefore concluded that DPS by stimulus wavelength was taking place during masked colour priming. Upon taking photometric measurements of the display screen it was discovered that white was weighted with a higher phosphor green component than either red or blue. This, it was suggested, could account for the green wavelength bias for white primes discovered.

Despite a dependence on wavelength rather than metemeric percepts, some researchers have suggested that neural correlates of visual masking are generally located beyond V1 and V2 (Tse *et al.* 2005). Although higher areas are thought to feedback percept dependant signals into earlier processes (Lamme *et al.* 2000, Schwabe *et al.* 2006, Bar

2003), they seem to have little impact on masking. Taken to the extreme, Macknik & Martinez-Conde (2007) state that top-down/feedback processes merely modulate attention towards and away from stimuli; masking effects themselves can be entirely explained by recourse to a feedforward model alone. They explain backwards masking using this model by suggesting that target 'after discharge' is disrupted by the subsequent mask and not the initial target signal. When a target disappears, a neural signal is generated marking its end point in time. It is suggested that after discharge occurs constantly as objects move in and out of visual fields due to eye saccades. It is thought that this signal is important for the creation of the perceptual consistency of objects (Macknik & Livingstone 1998). Its disruption may enable backwards masking to work as this may be analogous to the erasing of the sensory memory of an object.

Breitmeyer *et al.* (2004b) undertook to test prime/mask colour congruency effects using a colour reaction task. Results showed that under optimal *paracontrast* masking conditions, priming by colour was suppressed. Metacontrast colour masking showed no such suppression. In the same study, they found no suppression for priming by form either. That form acted as a DPS priming characteristic under both optimal paracontrast and metacontrast masking conditions suggests that form and colour primes are processed in differing, if not different, ways. This concurs with the view that form (contour) binding occurs in different places to colour, both in striate cells and beyond (Whitney 2009, Martin 2006). Here, differing neural correlates are assumed to represent differing processes although both chains may share some common neural areas. In particular, both processes would be expected to converge in the process of global feature binding, on both an abstract and cellular level (Whitney 2009, Grill-Spector 2003).

That masked colour priming exists at all appears to be at odds with another view in visual neuroscience. Higher cortical processes have sometimes been resolved into two streams: dorsal and ventral (Martin 2006). Although crude, this model may have some merits, perhaps attested to by its longevity (*ibid*). The ventral stream, running roughly from the striate cortex to the temporal poles, has been labelled the ‘what’ stream of object perception and is considered to consist of conscious processing; the dorsal stream has been labelled the ‘where’ stream and has been viewed as a non-conscious chain. However, as Schmidt (2000) points out, colour processing mainly lies within the ventral stream and so its processes should be subject to conscious scrutiny. Therefore masked colour priming should not be possible. Try as we might we cannot see the primes but their colour can affect our behaviour. This presents a problem for the conscious ventral/unconscious dorsal stream model and perhaps exposes it as being an overly simplified approach. We do not appear to be conscious of our brain’s colour processing, we just experience the qualia of colour. Visual illusions such as colour contrast effects (Lotto & Purves 2000) demonstrate our lack of conscious access to these perceptual processes. Introspection certainly does not allow access in the way it appears to for other cognitions such as decision-making.

The DPS of form has been demonstrated in masked priming experiments (Ansorge *et al.* 1998, Klotz & Wolfe 1995). When target and mask share congruent form, response times to a mask-shape identification task are higher than when an incongruent pair is presented. There is evidence that form is processed either by later V1 cells or by higher areas in the inferior-temporal cortex. For instance, Breitmeyer *et al.* (2004a) rule out early V1 processes in the masking by form. They ran a priming-by-form paradigm, which used target ‘wholeness’ as a parameter. Target primes consisted of either edges, corners or whole shapes and a choice reaction task was used to test for congruity/incongruity between target and mask, revealing priming processes. Results showed that priming was best achieved by whole form targets, followed by shape corners. Edge contours had no significant priming effect. This suggests form priming takes place beyond contour binding, an early visual process. Later V1 processes could be implicated by the 'corner' based object percept, perhaps at the level of Gestaltian object perception, e.g. the Kanizsa triangle.

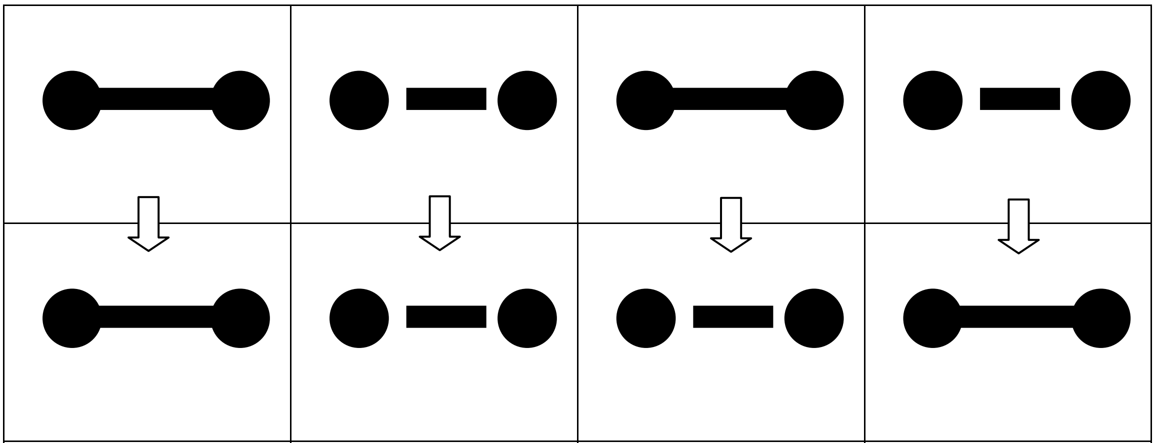
Motion induced blindness can be used to test levels of non-aware object processing. Mitroff & Scholl (2005) showed that object percepts could be updated to accommodate changes introduced in moments where conscious perception is absent. These updated percepts are subsequently available to conscious inspection. Motion induced blindness (Bonneh *et al.* 2001) is related to masking in that one stimulus can be used to interrupt the conscious experience of another. Fields of objects in congruent motion can act in a similar way to a mask. When a singleton is displayed on such a field, it can drift in an out of our conscious awareness. Mitroff & Scholl (2005) used a rotating grid of crosses as a motion field. A 'dumbbell' design was the singleton object consisting of two disks connected by a straight line. In half of the trial, the line fully connected the disks, in the other half it did not. Loss of conscious awareness was indicated via a key press at which point the central line of the dumbbell would shrink on some trials or grow on the others. This gave four conditions, see table 2.1, below:

a. no change connected;

b. no change unconnected;

c. change connected to unconnected;

d. change unconnected to connected,



no change, connected	no change, unconnected	change	change

Table 2.1. Stimuli conditions (derived from Mitroff & Scholl 2005)

A further key press indicated when the dumbbell disks returned to full awareness *simultaneously*. Results showed that the starting condition did not matter in that it was not important whether the disks were initially connected by the central line. What mattered was the finishing, post-blindness, state. The disks of the connected dumbbells appeared simultaneously back into conscious awareness roughly twice as often as the unconnected disks. The results also show that when a change ends in a connected (whole) object, the disks enter awareness simultaneously again around twice as often as when the percept is of two unconnected disks. The conclusion reached is that this latter discrepancy indicated that a non-conscious awareness of an object breaking apart is taking place: separate objects, such as unconnected disks, would not be expected to re-enter consciousness simultaneously but disks connected as parts to a larger whole would.

It might be expected that masked priming by form would share similar neural processes as to priming by colour. However, there appears to be a disassociation between form and colour masking. Although both colour and form priming are probably subject to V1 action, there are apparent differences. Surface qualities, such as colour, are seen as being subject more strongly to earlier, sensory dependent events. Contour qualities, however, seem to be subject to higher, percept dependant processes. Overall, visual masking appears to be subject to several differing streams depending on the stimuli and paradigm used. However, the neural correlate picture is far from clear.

It is unclear whether these masking streams suffer extinction effects. Some perceptual effects are subject to attenuation when in prolonged use; neural systems can cease to become excited by a stimulus, habituating to it, resulting in the lowering or suspension of an effect (Rains 2002, p423). Is visual masking subject to similar constraints? A single exposure lasting around 15ms for a target followed by a 200ms mask would seem unlikely to suffer such limitations but what about repeated exposure? There appears to be very little in the way of direct data on this topic other than the common practice of allowing participants breaks when engaging in extended trials. However, such ethical practice is common to much vision research and concerns eyestrain prevention as much as anything else. Studies do not generally report a tailing off of masking effects towards the end of prolonged trials but there may be a case for a meta-analysis of the statistical data.

One study took a different approach and looked at extinction of effect by prolonging the invisibility of a prime. Barbot & Kouider (2011) used CFS to induce extended periods of invisibility in a prime consisting of a facial image. They claim that under such circumstances prime effects did indeed diminish. The authors use CFS as opposed to standard binocular rivalry due to the latter's unreliable inter-ocular perception oscillation. When different images are presented to each eye, one perception will usually dominate although oscillations between the two can occur. The problem is that the oscillations are not exogenously controllable under standard binocular rivalry conditions. Instead of a steady state image presented to both eyes, one receives a cyclic flash that draws conscious attention away from the other eye-field. This paradigm is related to visual masking in that this masked stimulus, usually an image, can be processed out of conscious awareness.

Using CFS Barbot & Kouider were able to achieve primes up to one second in length, the effects of which were compared to those produced by shorter prime lengths of 60 ms. A 'famous/not famous' reaction time task was employed for visible, monocular probe faces. Prior to each visible face, a priming face had been presented. Masking flash/prime pairs were randomised left and right for each trial. The visible probes were also randomly presented to either left or right eye, controlling for ocular dominance confounds. Prime/probe pairs were either the same face or different. The study showed the expected priming process using short primes. However, results from long prime trials appeared to reveal a lowering of the priming effect. Looking more closely at their study, however, shows that the picture is not entirely clear. For example, in three out of the four experiments that used invisible primes, neither the long nor the short primes produced a significant result for non-famous people, where prime and probe were presented to the same eye. This may indicate differing processes happening for non-famous and famous (or recognised?) faces that could be confounding results.

Does the above study have anything to say about the visual masking paradigm used in this thesis? The answer has to be not directly. The dichoptic suppression induced by CFS is not visual masking and the results are not transferable, not least because of the temporal differences discussed above. However, the study results may support an argument for shortening the blocks of exposure when using visually masked primes within a VE, to prevent the extinction of any effects. This may be especially important if an RSVP paradigm were used. More work would be needed to address this in the future as the psychotherapeutic goal of long-term change may necessitate repeated exposure to the VE.

2.3.4 Attention and Visual Masking

It is known that attention can be captured in several ways including sudden spatial or temporal changes or the introduction of salient stimuli. Temporal capture can include abrupt onset (Posner & Cohen 1984) or offset (Miller 1989). A basic form of saliency can be found with singletons used in feature search paradigms (Treisman & Gelade 1980). The relationship between attention and visual masking is a complex one.

Searching for a target figure amongst a field of distracters, tends to involve lower location time where single feature searches are involved (e.g. a green letter ‘E’ amongst many black ‘E’s). If the amount of distracters is increased, this has little effect on the location time. However, increasing the number of features needing attention (e.g. a green letter ‘E’ amongst distracting black ‘E’s and green ‘F’s) can greatly decrease performance. Feature Integration Theory (FIT) claims that features such as a figure’s colours or contours are processed in parallel, prior to receiving attention (Treisman & Gelade 1980). For single feature searches, it is claimed, the figures themselves can be ignored allowing the feature alone to be attended to subsequently. For multi-feature searches, spatial attention must be allocated to each figure, which lowers task performance.

The existence of a negative priming effect, however, appears to provide evidence against the FIT claim that spatial attention is needed for figure perception (Tipper

1985). A figure can be negatively primed which leads to it being ignored rather than attended to. A subsequent task may then require that the figure be attended to. Research shows that under these conditions task performance can suffer (Lavie 1995). This may be indicating that figures are feature bound prior to attentional processing, in contradiction to FIT. One resolution is that task demand may specify late or early attentional processing, exemplified by FIT or negative priming respectively (*ibid*). This would suggest similar processes to DPS are at play. An alternative resolution would be to suggest that attentional processing arises from both top-down, percept-dependant signals and bottom-up, stimulus-dependant ones. Guided Search Theory would be a good exemplar of this (Wolfe *et al.* 1989).

Evidence for the impact of top-down processing on visual masking comes from studies where attention has been spatially divided: when attention is spatially divided then masking appears to become more effective. For instance, backwards pattern masking can be enhanced when a target letter is placed within a twelve letter array compared to a single letter array (Spencer & Shuntich 1970). For a typical masking paradigm, the target location is indicated by the mask location, even if the target itself remains unseen. Therefore, attention can be directed to the space that the target figure occupies. With a multi-figure display, this information is not available and so any attention must be spatially divided between the various parts of the stimulus. Divided attention can be seen as a candidate cause for the effect that visual masking appears to be stronger within multi-element, visual fields. A further hypothesis is that attention appears to be modulating the visual masking effect, which would imply some form of top-down feedback, contrary to Macknik & Martinez-Conde's (2007) model. However, it does not necessarily follow that attention is necessary for all visual masking. A position emergent from the research is that in order to strengthen visual masking, attention needs to be attenuated (Kammer 2007). This is similar to the way in which attenuation of figural context leads to a strengthening of masking effects (see section 2.3.3 above). Therefore, the claim is that although attention will modulate masking it is not necessary for it and can be detrimental. Increasing attention can only disrupt masking so far though. Depending upon the paradigm used, visual masking can be an extremely robust effect that no amount of directed attention can extinguish entirely. In the light of this, Macknik & Martinez-Conde's (2007) feedforward model seems plausible.

Nevertheless, it should be pointed out that the modulation of visual masking by attention is far from straightforward. One would expect that target and mask similarity would decrease attentional load and so be associated with stronger masking effects: if less work needs to be done noting all of the discriminating features which separated target from mask it follows that there would be a lower cognitive burden. This effect has been shown (Bahrami *et al.* 2007), supporting such a hypothesis. However, under certain parameter values, the NCE effect is observed (see section 2.3.2 above). Here, increasing target and mask dissimilarity increases the masking effect. NCE appears to be similar in some respects to the attention alerting effect (Posner & Boies 1971). Here, a timed response to a visual probe is recorded, the probe having been preceded by an attentional location cue. Intuition would posit that even introducing a time lag between cue and probe would still produce reactions faster than no cue. This pattern actually occurs for short time lags but, surprisingly, beyond a certain threshold the results reverse: having no cue seems to be better than having a long gap between cue and probe. Attention appears to wane if a lag is overly long, decreasing behavioural performance. For both the attention-alerting and NCE effects, increasing the attentional workload beyond a certain threshold appears to reverse the effects, even if one increase is temporal (cue/probe lag) and the other is spatial (figural similarity).

Feature migration (see section 2.3.3 above) may also be an indicator of attentional modulation. The similarities between these two effects and that of 'illusory conjunctions' is striking. FIT notes that when two stimuli are flashed upon a display in rapid succession then an illusory conjunction can occur whereby features from one element can become transferred onto another. For example, if a green 'E' precedes a red 'F' the conscious percept may be of a red 'E' (Triesman & Schmidt 1982). According to FIT, the procedure just described disrupts the focusing of attention necessary for the correct construction of visual percepts. The brain, in effect, makes a best guess as to what was seen, based upon the (poor) stimulus input and stored contextual constructs. This percept construction parallels feature migration in that aspects of one figure are transferred to another, thus forming a non-representational percept. Therefore, attention appears to struggle to discern one figural object from another given short temporal separation and duration of stimuli.

Attentional blink (AB) occurs when one target is being attended to at the expense of another (Shapiro *et al.* 1994). It does not *necessarily* imply that no attention is taking place (Giesbrecht & Di Lollo 1998) but that a processing bottleneck is occurring whereby *either* the initial stimulus processing is still ongoing or the attention system is being reset in preparation for a new input. This latter stage of AB is sometimes referred to as the 'refractory period'. It can be argued, for instance, that the ocular dominance experienced through dichoptic viewing can be thought of as a type of extended AB; the CFS effect can also be seen as a series of ABs.

Giesbrecht & Di Lollo (1998) used an RSVP paradigm to investigate the masking effect of AB and the so called 'second target deficit'. A random series of single, numerical digits were presented, one replacing another, centrally on a display.

Interspersed within were two target letters, the second of which formed the basis of a visibility response task. In the integration condition, the targets letters were masked by the simultaneous imposition of a subsequent digit, forming an aggregate figure. This describes a common-onset, overlapped, backwards pattern-masking paradigm. A second, interruption, condition had standard RSVP delayed masks trailing the targets. The results showed that the interrupt condition displayed AB masking with a second target deficit, resulting in a type B profile. The integration condition did not show a masking effect. The confound of high-level processes interfering in the integration condition was controlled for in a second experiment which used random dot masks. Following this, we can conclude that AB second target deficit is associated with interrupt masking and not integration.

However, Breitmeyer *et al.* (1999) see AB masking as an early process emanating from peripheral visual areas. This would place AB masking earlier in the visual path than Macknik & Martinez-Conde's (2007) masking but still within a feedforward model. These two forms of masking, peripheral and central, are considered to be analogous to what Turvey (1973) termed 'integration' and 'interrupt' masking respectively. The former is thought to be stimulus dependent (as in the colour primes noted above, see section 2.3.3), the latter percept dependant (as with priming by form). There is some question as to whether or not attention can group figures due to a task instruction (implying top-down cognition) or whether Gestaltian, spatial features are always used to group. Ramachandran & Cobb (1995) investigated grouping effects by using a simple target mask configuration, see figure 2.4 below.

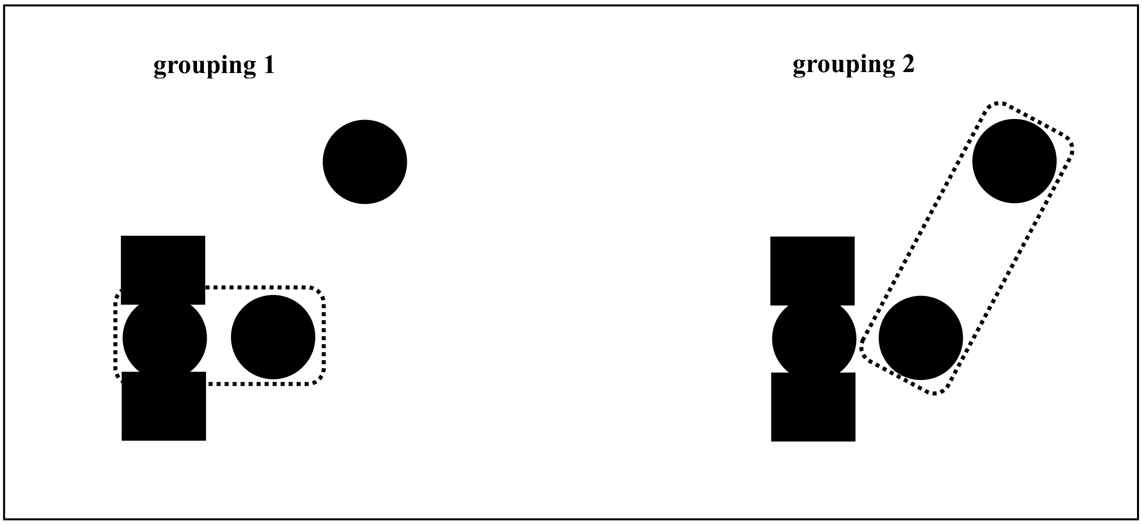


Figure 2.4. Target, mask and grouping figures, 1st main experiment (from Ramachandran & Cobb 1995).

Their study was intended to show how voluntary attention could modulate the masking process. Putting attention in the visual chain at the point of figural, object binding would put Ramachandran & Cobb's (1995) reasoning in line with FIT. It should be noted, though, that FIT posits an automatic spatial attention necessary to feature binding which one would assume to be a more fundamental process.

In the Ramachandran & Cobb (1995) study, a central target disk was flanked vertically by two masking rectangles. Adjacent were two other target disks, one close by the other further away. Note that the central disks overlap their flanking rectangles, which takes this experiment away somewhat from standard metacontrast, although it is still a backwards masking paradigm. In one condition, instructions were given to attentionally group the two lower disks (grouping 1). Instructions for a second condition asked that the two right-hand disks be attentionally grouped (grouping 2). It was found that masking of the flanked disk was stronger for grouping 2. It is possible that attentionally binding the flanked disk with its neighbour reduced masking as only part of this new, aggregate figure would be effectively masked by the rectangles.

A second experiment in the study looked at how the masking process could be turned on and off at will, using voluntary attention, see figure 2.5 below. Again, the rectangles acted as a backwards mask for the disks. All of the disks were rapidly displayed, followed by all of the rectangles. When the vertical column was attended to, the central disk disappeared but reappeared when the horizontal row was attended to. Therefore, this voluntary attentional switch seemed to act to suppress or allow masking. Whether or not such attention plays an active role in the masking process or merely acts as an awareness filter or similar is unclear from these results but they are interesting nonetheless, as they do show another type of attention having an effect on masking.

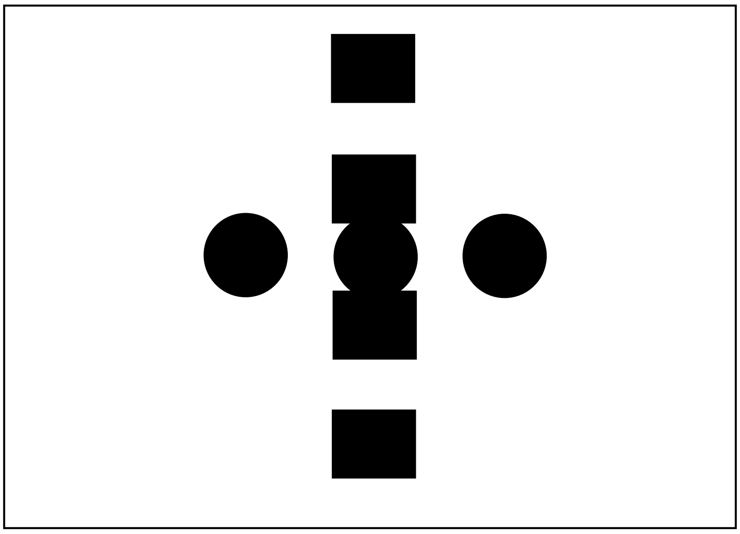


Figure 2.5. Target, mask and grouping figures, 2nd main experiment (from Ramachandran & Cobb 1995).

A further question also remains: Do these experiments actually show the attentional binding of figural objects into aggregates at the level of categorised percepts or is spatial attention the binding agent? Breitmeyer and Öğmen (2006) contend that these effects are more likely to be due to spatial attention grouping than disruption by a higher cognitive figural conception. This is supported by the Giesbrecht & Di Lollo (1998) study noted above: AB masking was not present when two target stimuli were superimposed at zero SOA. No attentional grouping instructions were given, instead automatic spatial attention processes appeared to be mustered resulting in a strong percept of the aggregate figure. Although this does not completely rule out some form of object concept grouping, FIT would further support the spatial attention, grouping hypothesis by the assertion that spatial attention was a necessary component.

The above research is mainly weighted to supporting the idea that attention needs to be inhibited in order for visual masking or masked priming to take place. However, the picture is further complicated by recent work that appears to show the opposite. Some research has found attention to be necessary to congruency priming effects. A study (Naccache *et al.* 2002) has demonstrated the requirement of a cue temporally preceding prime/probe pairs of lexical characters in order to achieve congruency effects. When probe onset could not be predicted, i.e. there was no cue present, reaction times were lower for trials where prime and probe were congruent compared to cued, congruent conditions. This appears to be evidence that temporal attention is needed to create congruency priming. Another study found a similar effect but this time for spatial attention. When directing attention to prime location on a decision task, again involving lexical characters, an effect was found to be present (see Lachter *et al.* 2004). As above, when cues were absent the effect was decreased. Taken together these studies suggest that, for certain forms of masked prime to be successful, a specific form of attention may be needed. This point will be revisited below when considering priming with facial imagery.

2.3.5 Masked Priming and Facial Expressions

DPS research supports the hypothesis that masked primes can affect behaviour

(Neumann & Klotz 1994). Ro *et al.* (2009) found that both target colour and form could influence participant behaviour. However, it may be the case that the more complex a form that a prime takes, the less effective it will be. For instance, as noted above (section 2.3.3), increasing target complexity decreases target recognition, even when masking is not strong. Care must be taken over the target form. There may be limits to target complexity, despite visual information being processed within milliseconds. Rapid, semantic visual processing does not appear to be available for either long words or sentences (Dijksterhuis *et al.* 2005). One category of prime, though, appears to run against this trend: facial imagery. This even applies to faces with complex affective expressions and may be due to the possibility of specialised facial processing within the brain (Grill-Spector 2003, Kanwisher *et al.* 1997). Faces do appear to be highly salient stimuli where human beings are concerned (Viggiano & Marzi 2010).

Complexity here can be loosely defined by how mathematically compressible an image is, for instance, when stored as a computer file. Generally a compression algorithm would be expected to deliver a larger file for a facial image than for a similarly sized graphic figure such as the annulus or disk from figure 2.1, for example (Flake 1998, p134). Although facial images may be complex when compared to simple graphic figures, it may be the case that only a simplified set of hardwired, salient features that are responded to (Johnson *et al.* 1991). In reply, though, we can argue that affective expressions can be quite complex, if not to say subtle, in their geometry. One would expect that the processing of such stimuli would impose a high cognitive load, therefore negating rapid masking effects. This does not appear to be the case, not least as evidenced by the results of the experiments presented in this thesis.

A 'facial superiority' effect (FSE) has been reported in masking studies (Van Santen, J.P.H & Jonides, J. 1978). Here, stronger masking effects tend to be obtained using facial image targets that have been inverted or have scrambled features. This can be likened to the OSE but may be the result of processing in different parts of the visual path. Masking inhibition found with the OSE possibly originates in early V1, as it would have to rely on liner contour processes. Some higher feedback *may* be present as an inverse relationship has been found between V1 activity and that of higher areas (Murry *et al.* 2002) but there is some debate as to whether or not these results are merely due to image statistics, and therefore stimulus driven, as opposed to feedback driven (Dumoulin & Hess 2006).

The FSE, on the other hand, is almost certainly the result of later processes, possibly in the fusiform face area (FFA) amongst others, beyond V1. Whilst the FFA may not deal only with facial imagery (Gauthier & Tarr 1997), nor be the brain's exclusive 'face site' (Grill-Spector 2003), it appears to be heavily involved with such work (Kanwisher *et al.* 1997). That it arises from higher processing may make the FSE more flexible than the OSE. For instance, FSE may be more easily disrupted by voluntary attention or other means such as display complexity.

At the end of the previous section that dealt with attention and visual masking, it was shown that certain forms of prime appear to need some form of attention in order to produce an effect. Yet other research has shown that attention can be detrimental to visual masking. Taken this simply, these results would suggest that masked priming cannot work. However, it clearly does as witnessed by the plethora of DPS research (see section 2.3.3 above). Unpicking this dichotomy is beyond the scope of this thesis but it is suffice to say that attention may not be required for all types of prime.

Finkbeiner & Palermo (2009) found evidence that faces in particular do not need attention to be processed. They designed a series of experiments where visible probes needed to be classified as quickly as possible. For one experiment, photographs of faces were used for both the prime and the probe. The backwards mask was a scrambled version of a facial image. The task required the rapid gender identification of each presented probe, primed with either a congruent gender image or incongruent. A second set of primes and probes used images of animals and vegetables. During the trials, primes were always presented at the top of the display and probes at the bottom. A spatial cue was shown for either prime or probe. In this way congruency priming effects could be compared across image types and cuing conditions. Image confounds were controlled for by using a set of line drawings for non-facial stimuli in a second experiment. Results showed that for facial stimuli prime type (congruent/incongruent) did *not* interact with cue type (prime or probe) but each factor demonstrated a main effect. An interaction effect between prime type and cue type *was* found for the non-facial stimuli however.

Finkbeiner & Palermo (2009) argue that these results provide evidence that spatial attention is not needed for priming with facial images. They also claim that they are *not* stating that spatial attention can have no effect on facial priming, just that it is not necessary. I would suggest that we cannot make such a distinction from their study alone. This distinction may be the case but the above study certainly seems to show that spatial attention makes no difference to facial priming. Further, it may be that faces can be processed without recourse to attention and that they are therefore a special case of a prime. However, the cueing format of the above experiment leaves open the question as to the type of attention involved. There may have been a temporal factor at work in the cuing as well as location indication. If this is so then face primes may be immune to inhibition from variety of attention types. For instance, we can posit that the FSE need not be detrimental to masked priming. If facial processing can proceed in the absence of attentional awareness then we can alter that attention, perhaps by spatial means, in order to inhibit the FSE.

It needs to be born in mind, though, that faces may not be a special case for all of cognitive processes with which they are associated. For instance, it has been documented that overt, 'within class' object recognition can be poor for prosopagnosia sufferers (Damasio *et al.* 1982, De Haan 1991). This may contra-indicate that faces are processed by isolated visual streams. Additionally, Pessoa *et al.* (2002) looked at fMRI activity in the occipito-temporal visual cortex, including what would be the FFA, and the amygadae using facial stimuli. They found that activity during a gender discrimination task was higher for attended faces than for unattended faces on an orientation discrimination task across those cortical areas. The faces carried emotive expressions for the experimental condition and neutral expressions as a control. The study is interesting in that it purported to show that attention was needed in order to process facial affect. However, two points need to be raised. Firstly, a methodological objection can be raised to the work in that the attended face task (gender) was different to the unattended task (orientation). In effect, the study did not compare like for like and task type could have been a confounding factor. It is hard to see how the experiment could have been designed otherwise but the point still stands. The second point is that the faces were visible and therefore accessible to conscious awareness. This may dictate the type of attention under study and may not apply to masked stimuli. Masked emotion may not require attention as part of its processing: Jurena *et al.* (2010) found amygdala activation in response to masked happy facial expressions, even greater than that found for sad faces.

Individual differences in attention may also play a role in how expressions are processed. Highly anxious people can exhibit heightened attention to their own expressions, seeing them in a distorted way that non-anxious people do not (Clark 2001): the heightened attention itself may be playing a part in the distorted cognition. Further, individual differences more generally may be a significant factor in face and expression processing (Conway *et al.* 2007). This needs to be born in mind with research of the type described in this thesis. To see why, it is helpful to study a standard model of face processing. Bruce & Young's (1986) model of face perception consists of two main stages. The first extracts the underlying facial features by calculating view-independent geometry and factoring out expression and so on. The resultant 'structural encoding' data is then forward fed into the next stage involving a series of lesser processes. These consist of Face Recognition Units (FRUs), Person Identity Nodes (PINs) and name retrieval. Attributes such as gender, as well as facial expression, are considered to be streamed off earlier, during structural encoding. Haxby *et al.'s* (2000) neuro-anatomical model supports Bruce & Young's system by setting out the biological underpinnings of the recognition and perception of faces in the brain, the two models mapping well together (Calder & Young 2005). The neuro-anatomical 'core systems', mainly areas within the temporal/occipital lobe junction, correlate with Bruce & Young's first stage. Haxby et al.'s (2000) 'extended system' maps onto the subsequent stage as well as linking to other functions like expression processing. Further work has taken place that attempt to model the expression-processing stream using Principal Component Analysis (PCA) (Calder *et al.* 2001). PCA does seem to show promise and may provide insight into the way facial expressions are processed in the brain (*ibid*).

There are some drawbacks to the Bruce & Young model. It may be too rigid in its linearity and cannot, for example, explain easily a case presented by De Haan *et al.* (1991): PH had several adult acquired cognitive deficits as a result of a head injury. Amongst these were a form prosopagnosia. However, PH was able to access occupational information linked to faces he did not appear to overtly recognise (photographs of famous politicians' and entertainers' faces). Somehow, he was able to bypass overt recognition, presumably stored in a Bruce & Young FRU, and access identity information (occupation) from a PIN. An MRI scan of PH indicated a bilateral abnormality in the inferior region of the occipital and temporal lobe junction, possibly the FFA. Despite such issues, it is still a useful model for orienting facial processing streams (Humphreys *et al.* 1993) and survives albeit with some modification (Burton *et al.* 1990).

As noted previously, Bruce & Young separate out expression processing early on in their face perception process model. From this, we might expect that expressions have their own set of separate cognitive processes that in turn may be vulnerable to producing their own specific deficits. Indeed, this appears to be the case. One general deficit is known as facial expression blindness (FEB). It is a well-documented phenomenon and has been observed as a recurrent feature in autistic spectrum disorders for instance (Sigman *et al.* 2006, Capps *et al.* 2006, Gross 2004, Clark *et al.* 2008, Grossman & Tager-Flusberg 2008). The link to autism has lead to arguments that FEB may be a theory of mind (TOM) deficit (Baron-Cohen *et al.* 1997) or even an attentional disorder (Begeer *et al.* 2006). In both of these cases, the underlying perceptual substrates, Bruce & Young's first stage, could be intact, the deficit arising from processes outside of the face/expression perception chain (Ozonoff *et al.* 1990). We would therefore expect FRUs and PINs to still perform normally. FEB deficits may manifest themselves for the reading of specific expressions and this may be due, in part, to the complexity of that expression (Gross 2008). Varying levels of ability to read an expression, or expressions in general, may account for some of the variance found in studies that use facial expressions as stimuli. This suggests that controlling for individual differences would therefore be important when faced with such a confound. It also underpins the methodological approaches found within this thesis such as the prevalent use of within-participants designs and randomly assigning participants to various groups.

Intuitively, we might posit that a particular facial expression is related to a specific emotion: the expression gives a conspecific viewer an insight into the affective state driving that expression. All neuro-typical people, it is assumed, possesses an understanding of this facial language. A problem with such a position is that, due to individual differences, a given facial expression may not be read, or even be readable, as conveying the actual affective state of the expresser. The dynamic, fleeting nature of expressions can lead to confusion (Ekman & Rosenberg 2005, Ambadar *et al.* 2005), especially when multiple expressions are seen as being merged (which may be as much a product of cultural categorisation as anything else). It may indeed be the case that expressions reflect the nature of the underlying affect state which itself may be dynamic and confused. However, if they do not (and there is no unequivocal evidence to say that they do) then we cannot say that there is a clear link from affect to a distinct and specific facial expression. Further, other affective states besides emotion may be even more difficult to convey. Mood states are taken to be more muted than specific emotions (Fox 2008, p26) and are perhaps harder to gauge from facial expression alone. Reading the affective state of another seems to be an important social skill, and not just for our species, but we do not always get it right.

The picture is further complicated by the way in which affect and expressions are categorised. A debate still currently runs between the categorisation of affect in either dimensional terms or as discreet entities (Bachorowski & Owren 2008). Component process models (Schere 2005) may be a way out of this but nonetheless the debate stands in contrast to the categorisation of facial expressions that still tend to be treated as discreet forms, at least from a Western perspective. Affect categorisation will be raised again in the following section of this review.

There may be some support from neuroscience for using a discreet category approach towards both emotions and expressions. FEB does not have to mean a general expression reading deficit. Adolphs *et al.* (1994) looked at the case of SM., a patient who had extensive damage to both amygdalae but who's adjacent brain areas such as the hippocampi were intact. SM had a relatively normal IQ and could recognise faces. However, SM could not read fearful expressions. Using previous controls, the study had shown that although a particular expression could be regarded as a member of more than one affect category, expressions generally tended to fall into separate groups. The authors demonstrated that SM treated expressions as distinct. They concluded that amygdalae are needed to recognise fearful expressions but that they also aided in the reading of blended expressions. Amygadalea do not appear to aid in the recognition of faces. The study also supports Bruce & Young's model as far as the separate processing of facial expression and identification are concerned.

Affective priming using visual masking often uses facial expression to elicit an emotive response in participants (Killgore & Yurgelun-Todd 2004, Whalen *et al.* 1998, Ohman 2002, Wong & Root 2003). Given the previous points just raised, it may seem surprising that masked expression priming can have any effect. After all, if all of the cognitive resources marshalled by conscious awareness can have difficulty reading an expression what hope would presumably simpler, non-conscious processes have? Expression priming does seem to work, however, although the previous points may account, in part, for variance found within experimental results: even though such studies also support the division of expression into discreet categories (Ekman & Rosenberg 2005), a strongly categorical approach may not necessarily be helpful in capturing the complexities of everyday facial dynamics (Valstar & Pantic 2012).

One study, Winkielman and Berridge (2004), looked at how visually masked, emotionally valent, static facial images can affect behavioural responses. Participants experienced either a masked smiling, neutral or angry face, all masked by a neutral face. They were subsequently asked to drink a beverage offered and rate it for monetary value. Those who indicated that they were thirsty prior to the trial, and experienced a masked smiling face, tended to drink more and award a higher monetary value than those in the other conditions. As Winkielman and Berridge point out this indicates that thirst was acting as a modulating factor on the non-conscious affective priming of a behavioural response.

Another study, Dimberg *et al.* (2000), again used masked, static emotive expressions but this time looked at facial muscle reactions via electromyography (EMG) (Cacioppo *et al.* 1986). It was found that appropriately matched facial muscles were activated as if the participant were mimicking the masked stimuli. However, unlike the pervious study, no modulating factor was described or appeared to be necessary.

Taking the two studies together may suggest that some form of modulation may be necessary to elicit a complex behavioural response, beyond the subtle muscle movement picked up by EMG. Certainly, participant phobia can act as a modulating factor for endocrinal-driven skin conductance responses, and emotional valence ratings, to such masked stimuli (Ohman & Soares 1994).

We know that people respond in similar neural ways to avatar facial expressions and real expressions (Krumhuber *et al.* 2005, Britton *et al.* 2008) although some differences are noted. For example, amygdala activity was measured during facial emotion recognition tasks that used both avatar and human faces. A greater difference in activity between the avatar and human faces was found within female amygdalae than male (Moser *et al.* 2006). This still allows for the use of avatars in masking experiments, however, as a 'robust amygdala activation was apparent in response to both human and avatar emotional faces' (*ibid*). This suggests that, to some extent, the brain treats avatar based emotional stimuli in the same way as human based ones.

There is much published work available showing expression form and dynamics, for example smile muscle motion (Ekman & Rosenberg 2005). However, it may not be a trivial task to construct an avatar smile that represents a particular type, such as a Duchene smile (Messinger *et al.* 1999). Simple classification of expressions such as Facial Action Coding (Ekman 2009) may not be adequate. Some studies involve affective reactions induced by VEs, particularly anxiety (Meehan *et al.* 2002, James *et al.* 2003, Pertaud *et al.* 2001, Wilhelm *et al.* 2005). Several studies show responses to avatars' facial expression (Bailenson *et al.* 2006, Weyers *et al.* 2006) and even TOM type judgements of avatars' emotional state (Rizzo *et al.* 2001, Yun *et al.* 2009). Fewer directly address emotion induction using dynamic avatars facial expression (Causse *et al.* 2007). Currently, there appear to be no avatar/virtual character studies that use visual masking.

2.4 Mood, Cognitive Bias and Implicit Self-esteem

2.4.1 Introduction

This section will explore the possibility and usefulness of mood induction via masked facial expressions and provides the theoretical background to justify such an approach. It will begin by setting out the scheme of affect categorisation underpinning this thesis. Following this will be an examination of the link between affect, specifically mood, and information processing. The effect of mood cognitions on both processing style and information content will be considered. Mood induction will be explored, beginning with contagion phenomena, which will then lead into a discussion on implicit self-esteem. This latter area will be seen through the lens of information processing and provides an explanation as to why mood induction via masked facial expressions could be of benefit to the psychological treatment of affect disorders. This section is of particular importance for the final two experiments presented in this thesis. The kinds of measures used by these experiments will be shown to be justified in terms of the low intensity of mood changes, as opposed to the higher intensity of emotional changes, utilising self-reporting and cognitive measures over physiological methods.

2.4.2 Affect, mood and emotion

Affect research has a history comparable in length to visual masking, stretching back into the Nineteenth Century (Darwin 1899). However, within that time affect has proven difficult to define, judging by the plethora of models (Lewis *et al.* 2008). The terms ‘affect’, ‘mood’ and ‘emotion’ can be used inconsistently, within research and elsewhere. For example, anxiety can be cast in mood-like terms whereas fear can be seen as an emotion (Rachman 2004, p3). The presence or absence of a target is sometimes posited as a distinguishing measure: anxiety is regarded as having no target whereas a target will be present when fear is extant. It may be that generalised anxiety disorder (GAD) has no specific target but the same cannot be said for social anxiety disorder (SAD) (DSM-IV-TR 2000). Specific sub-categories of SAD, such as public speaking anxiety, have well defined and predictable targets (Rachnam 2004, p148). We could argue that these are still general target forms. However, this is also the case for phobias, which are regarded as fears and therefore emotions. It seems unclear as to whether anxiety is a mood or an emotion. We could also ask is a phobic fear a mood or an emotion? If someone has a phobia, they do not necessarily have to experience it all of the time. Presumably, there will be times when an ophidiophobiac is not feeling any emotion related to the snakes. Does the category change upon encountering a snake i.e. when the phobic starts to process real-time, perceptual information? Perhaps a phobia is a background, affective process, like a mood, that only becomes an emotion upon encountering a specific target. Further, not all phobias are as clearly defined as specific animal phobias; agoraphobia, for instance, has a less clear target. Here, a sufferer may experience a constant background affect that could be described as a mood.

In order to avoid this semantic drift, the terms 'affect', 'mood' and 'emotion' will be used in line with what appears to be an emerging consensus (Fox 2008 p26). Two dimensions, intensity and temporal dynamics, underpin this approach. Emotions are relatively short lived and intense whereas moods are of a longer duration but are relatively less intense. Here, intensity can be defined by behavioural, physiological and subjective components (Scherer 2005). The subjective element is often termed a 'feeling' (Scherer 2005). Emotion can therefore be seen as being related more to state than trait measurements, whereas mood can be both a state and trait process. Mood is seen as a background proclivity for a certain emotion to occur and therefore explains how emotion can be sometimes viewed as a trait. The word 'affect' is used in its common form as an all encompassing term to describe these like phenomena. Also accepted is the aggregate nature of affect, expressed by Scherer's Component Process Model (Scherer 2005).

2.4.3 Mood and cognition

We know that affect can affect the way in which information is processed (Bohner *et al.* 1992, Dolan 2002). Information can be processed differently when in a positive mood from when in a negative mood. For instance, Gasper and Clore (2002) found that people in a happy mood tend to classify images based on global aspects more than people experiencing a negative mood. Those experiencing the latter had a tendency to focus on detail. Levine and Pizarro (2004) provide a list of motivational states, identified from research, which they link to both affect and information processing strategy. A positive valence, for instance, is associated with a motivational state arising from a lack of immediate problems that need solving. This also relates to a processing strategy that can an increasingly rely on heuristics and stored, general knowledge.

Adaptation has been put forward to explain this phenomenon. Broadly, it has been assumed that when in a positive state most, or all, of a person's needs have been, or are being, met whereas a negative state indicates that they are not (Maslow 1943, Buss 2004, p312, Panksepp 1998, p164). A reliance on internal, simplistic constructs is sufficient to engage with their environment when people are happy. This chimes with the finding that redrawing an unusual face from memory multiple times results in a 'standardised face' when in a positive mood (Bartlet 1932). Here the standard face can be seen as a simple, stored construct. Also, Bohner *et al.*'s (1992) found that, when in a happy mood, there is a tendency for people to attend less to either the context or content of arguments when compared to those in an unhappy mood. Levine & Pizarro (2004) found people experiencing a negative mood to be attentive to information derived from the environment, compared to controls, and with details being particularly salient. The argument here is that a negative mood is more likely to result in a detailed needs appraisal. The realisation of a needs deficit may then lead to behaviour more likely to attenuate the deficit (Mealey & Theis 1995).

Adaptive arguments, despite the criticism that they can lead to un-testable hypotheses, do at least offer some form of framework for understanding the links between affect, cognition and behaviour. However, there are other problems besides methodology. Clearly, a simple adaptive framework, running from affect to cognition to behaviour, cannot be the whole 'story' of the relationship between mood, processing style and actions. To begin with, an organism's appraisal of its needs may not always involve affect: we would not see hunger as a bad mood although the experience of hunger can be a precursor to food seeking behaviour. Some people *can* experience a negative mood associated with hunger, which subsequently dissipates when food is eaten. However, this is not necessarily a primary component of hunger even if the promise of mood alleviation gives an added impetus to find food.

Another difficulty with the argument is that no account is given for the body of empirical work that shows a causal link running from cognition to affect, effectively running counter to the above. CBM is a currently expanding field of research dealing with such a link. In recent years, CBM has become an important experimental tool within both cognitions research and clinical psychology (Koster *et al.* 2009). Research in the latter has arisen due to the potential of CBM to modify biases that may maintain affect disorders (Browning *et al.* 2010). At the core of many CBM procedures is a task designed to bring about type of bias priming, repetition then being used to strengthen this effect. CBM began life as a technique used in research to investigate hypotheses centred on information processing bias models of affect psychopathology (Koster *et al.* 2009, Hertel & Mathews 2011).

There are signs that CBM itself has an impact on SAD (Hertel & Mathews 2011, Beard *et al.* 2011, Koster *et al.* 2009, Dandeneau & Baldwin 2004, Dandeneau *et al.* 2007, see also, Mathews & MacLeod 2002). A trend in the research is the use of facial imagery to affect, for instance, the attentional biases associated with SAD. CBM lends itself easily to such a paradigm. A CBM task may involve attending to a pair of faces on a screen (Beard *et al.* 2011). One face may display a positive expression, the other negative. The faces are displayed for a short period, and then disappear, to reveal a letter under one of them. Using a mouse or keyboard, the participant is asked to indicate whether the letter is either an 'E' or an 'F'. In the experimental condition the letter will always, or predominantly, be revealed in the position of the positive expression. Therefore, the participant quickly learns that the positive expression signals the position of the information they need to respond correctly to the task. With this design, the strength of the effect can be measured by a reduction in response times during the task. This will be likely to occur as participants become proficient at ignoring negative expressions and attending to positive ones during task repetition. Avoiding negative expressions, for example in a crowd or audience, therefore becomes habitual. Biased attention to negative expressions (‘scanning’) in such situations can be one indicator of SAD (Rachman 2004, p34). The above task was used as part of a randomised control trial study into the treatment of SAD and is typical of the field.

The Beard *et al.* (2011) study also involved a second experiment, which tackled interpretation bias, also thought to be prevalent in anxiety disorders. This second task used linguistic stimuli instead of faces. Briefly, a word appeared on screen that was then replaced by an ambiguous sentence. The affect valence of each word was either positive or negative. The participant had to indicate whether the two were related by pressing keys, one key for 'yes', another for 'no'. They received feedback to their responses after each word/sentence pair. When the participant indicated an association between the positive words and the ambiguous sentences, the experimenter indicated that the choice was correct. This was also the case where a response was given indicating no association between negative words and the sentence. Both tasks taken together, and repeated, showed a lessening of SAD, measured by blind ratings of an impromptu, post-treatment speech as well as various inventories. The authors note that the interpretative task gave slightly stronger results although the best results were obtained by combing the two. Overall, the study demonstrated a causal link from cognition to affect.

The possibility of causal loops between affect and cognition arises: appraisals lead to certain affect states, which in turn bias processing styles, leading to particular cognitive outcomes and so on. CBM research adds to prior findings where cognitive style correlated with particular affective outcomes such as affect disorders including anxiety (Mathews & MacLeod 2002) and depression (Beck 1967). An interesting point to note, though, is that whilst CBT encourages a patient to focus explicitly on dysfunctional thoughts (i.e. biased processing) which may have remained implicit or unnoticed, CBM techniques, in contrast, encourage the opposite: a habitual ignoring of implicit processes formed by repeated distraction from them. However, it is possible that CBM works at a different process level to CBT and would not therefore act as a competing processes should they be used together in a wider therapeutic program.

Damasio's 'Somatic Marker Hypothesis' (Damasio 1994) moves embodied affect into a central position in decision making and judgemental processes. Damasio presents a neurological model that underpins his position that successful life decisions cannot be made without an affective component. This implies that moods, for instance, could alter judgements. There is evidence for this: Forgas (1999) found that, as situations become more complex, mood effects on judgements become more prevalent. Happy, confident individuals might be expected to take more risks in complex situations, for example. Forgas' 'Affect Infusion Model' is developed out of such findings (*ibid*). Here, happiness might be considered as a state mood, bordering on an emotion, with confidence seen as a trait mood with associated cognitions including high self-esteem.

Although perhaps a longer-term trait, confidence can be eroded by a single episode. One, negatively valenced event may affect future information processing leading to vulnerabilities for anxiety and depression. Some such events appear to not need prior judgement as to whether or not they are aversive. This seems the case where PTSD arises. The co-morbidity of PTSD and depression is common (Hammen & Watkins

2008, p21). It has been suggested that the negative processing style which follows the trauma sets up further negative mood states which can lead to further processing bias and so on (Mathews & MacLeod 2005). Such processes may contribute to the stability of affect disorders generally. A question remains, though, as to why do traumatic events result in PTSD for only some people? Could this be an indicator of prior processing vulnerabilities in those individuals who go on to develop PTSD and be a reason why moods can affect subsequent information processing? This position is at odds, though, with Zajonc's claim that 'preferences need no inferences' (Zajonc 1980). Although perhaps overly simplistic here, Zajonc does point to the primacy of affect in some processing chains such as the 'mere exposure' effect. It may simply be that these causal loops sometimes begin with affect (Zojonc 1980, Chen& Bargh 1999), others with cognition (Mathews & MacLeod 2005).

Primacy of affect can also be found in Dual Representation Theory (Brewin 2001). Here, traumatic memories are categorised into two broad types: basic sensory data and memories embedded within a narrative context. It is suggested that the traumatic memories of PTSD sufferers are stored as sensory data, the normal process of episodic embedding having broken down. PTSD flashbacks would consist of the triggered or unsolicited retrieval of these sensory memories. If correct, this would demonstrate an event's affect value having a direct effect on memory encoding, perhaps analogous to the DPS sometimes found with masked priming. Here the affect should be seen as an emotion rather than a mood because of its intensity.

Mood has been found to influence memory encoding generally. For example, it has been shown that learning is more effective when the material and mood state are congruent (Bower *et al.* 1981). The findings for the influence of mood on retrieval are not as clear-cut though. Fox (2008, p209) describes a typical retrieval experiment whereby lists of positively and negatively valenced words are learnt whilst in a neutral mood. Later, a valenced mood state is induced and retrieval measured. However, Fox states, results have varied with some studies reporting mood congruent effects yet others report no mood effects at all. When the moods at both encoding and retrieval are congruent (i.e. either both are positive or both are negative) then a mood dependant memory effect can be seen. Items learned in a particular mood state are more easily recalled later in the same mood state. That this is reproducible in a controlled experiment exposes a causal line from affect to information processes: mood can bias both memory encoding and retrieval. Further, it has been claimed that differing mood valence can lead to differing processing styles and that these biases, in turn, result in different memory content being present (Fiedler 1990).

As well as biasing memory content, it has also been suggested that moods can also bias the content of environmental information processing (Mealey & Theis 1995). It has been found that people in a positive mood state will look for prospective signs within a landscape as part of their exploratory patterns. Those in a negative mood tended to look for refuge possibilities within their environment.

Moods might also act as informational content in their own right. Both experiments in Schwarz and Clore's (1983) classic misattribution paper found that the environment had an effect on mood. However, their results showed that the environment only lessened the influence of negative moods, on a life satisfaction rating. They claim to have shown that moods, along with environment, act informatively. They further explain that negative moods in particular act as motivational data to make people seek explanations for those moods. A problem with this, though, is that it is unclear what is actually acting as an input into the judgement processes. Schwarz and Clore would perhaps claim that the subjective feeling of the mood is acting as information. After all, Schwarz and Clore criticise Wyer & Carlston's (1979) version of an affect-as-information model for a lack of 'current experience'. Based on research into arousal and dissonance (beyond the scope off this thesis), Scwartz (2010) claimed that a subjective component, termed the 'online experience', should play a crucial part in an affect-as-information model. However, the outcome is an objective (life satisfaction) score. At what point does this conversation from subjective to objective data take place and how? Affect-as-information models do appear to struggle with subjective awareness.

We could regard a mood as information for input into appraisal processes, with no subjective element. However, this would stand odds with not only our individual experiences but also the component process models that currently offer progress in the understanding of affect in its various forms. An extreme, affect-as-information position might counter this by claiming that the presence of the subjective indicates a redundancy built into component process models. However, affect-as-information models either ignore or dismiss subjective elements or, as in the case of Schwarz and Clore (1983 & 2003, Schwarz 2010) muddle subjective and objective components together in a similar way to the threshold models of subliminal perception (see section 2.3.1 above). It might be suggested that moods do have a subjective component but that this element is not used as information. Here, though, we would be returning to Schwarz and Clore criticism of Wyer & Carlston's model.

Another difficulty with Schwarz & Clore's model is the assumption that affect, when acting as information, is explicit, due to the necessity of subjective experience. Masked affect priming studies show that this need not be the case (Ohman 2002, Winkielman & Berridge 2004). Information processing models are prevalent in all areas of cognitive psychology: from low-level sensory processes to high level judgements and conscious awareness. Such is the pervasiveness of these explanatory frameworks that it is easy to see them as becoming synonymous with the science of psychology itself. Although extremely useful, this approach may need some review if models of affect processing are to incorporate subjective elements successfully.

2.4.4 Mood Induction

Affect contagion, and its sub-category mood contagion, appear to be related to mood induction. Indeed, it may be the case that the same or similar neural processes are present in both cases. The term ‘mood induction’ has been used to describe laboratory based techniques which deliberately attempt induce often subtle shifts in background affect (Gilet 2008). ‘Affect contagion’ and ‘mood contagion’ arguably also describe processes beyond the laboratory as well as within: affect spreads from individuals or groups to other individuals or groups, analogous to the spread of infection through a population. It is regarded as a largely automatic process: Neumann & Strack (2000) define mood contagion as the ‘automatic transfer of mood between persons’.

Affect contagion can be associated with facial stimuli. Hatfield *et al.* (1994) found that when we are smiled at warmly we may be induced us to generate a smile of our own in reply. The strength of emotion evoked can be proportional to the strength of the expression (Wild, *et al.* 2001). There may be a combination of factors at play during ‘infection’: firstly, the experience of positive affect may simply be a reaction to being smiled at warmly, possibly processed in part by the FFA, as expressed by Bruce and

Young's separate expression processing stream. Secondly and less directly, is the mimicking of observed expressions followed by a tendency to experience an affect appropriate to the mimicked expression. This latter effect is the core prediction of the facial feedback hypothesis (FFH) (McIntosh 1996). It is important to note that the mimicry alone cannot provide a complete description of such 'contagion' as affect changes involve more components (Gallagher 2007). Additionally, there is some evidence that individual differences are a factor in facial feedback sensitivity, in particular emotional empathy score (Andréasson & Dimberg 2008).

As noted previously, the mimicking of *masked,* static faces has also been observed (Dimberg *et al.* 2000). Interestingly, Tamietto & de Gelder (2009) found that appropriate facial expression were elicited in response to backwards masked targets depicting affect valent body positions. One of the stimuli showed an open-armed gesture, the other a defensive position. Facial EMG was used to measure muscle reaction. Results showed more zygomaticus major activity associated with the open gesture with the corrugator supercilli displaying more activity in the defensive condition. This indicates the possibility of cross-modal processes for the FFH, mood contagion or induction.

A wide variety of mood induction techniques exist, a large portion of which do not use facial stimuli. Standard mood induction techniques (Gilet 2008) include music (Southerland *et al.* 1982) moving images (Rottenberg *et al.* 2007), Velten mood induction (Velten 1968), autobiographical recall (Brewer *et al.* 1980) or mixtures off the above (Baños *et al.* 2009, Mayer *et al.* 1995).

Unfortunately, some approaches are susceptible to demand characteristics where a person may behave as if they have undergone a mood change but in fact have not. The early form of mood induction, the Velten (1968) procedure, is one such (Gilet 2008). With this procedure, participants are given a list of statements with particular emotional valence and asked to imagine and try to experience the feelings associated with them. The original study included two conditions that purported to control for demand characteristics by asking other participants merely to act as they felt others from the experimental condition would. A difference found between the before and after mood changes for each group was seen as evidence that the experimental condition was free from demand characteristics. However, it has been argued that this view was erroneous and that people in the experimental condition should have easily been able to guess as to the nature of the task and the results required (Kenealy 1986). In effect, all that the original study showed was that people who imagined themselves to feel in a certain mood filled out inventories differently to people who merely acted that way. This is a problem for all mood induction techniques where people are overtly asked to self-manipulate their own emotional state, for example with autobiographical recall (Gilet 2008).

Regarding the EMMA Project (mentioned in section 1.1.1 above), various techniques from mood induction science were introduced to VEs. The early stated aim of the project was to better understand presence by exploring any relationship found between emotion and presence (Alcañiz *et al.* 2003), emotion and mood being somewhat conflated here. It was found that participants' subjective presence within the VE was somehow increased by the mood induction techniques (Baños *et al.* 2006). Unfortunately, there appears to be no reason given as to how using these mood induction techniques within a VE would be free from the problems associated with using them outside of one. For example, participants could still be able to guess what was required of them in order to produce a particular result and so introduce a demand characteristic variable.

Other techniques, whilst not so readily affected by demand characteristics, may not be appropriate for raising implicit self-esteem (Farnham *et al.* 1999). For instance, induction via music could be one such technique: It is unclear how listening to music could affect a person's social self-image, in particular music without any verbal content. The importance of raising implicit self-esteem is discussed in the separate section 2.4.5 below.

Mood induction, free from demand characteristics but which can raise implicit self-esteem, may be possible by the use of masked facial expressions. We know that standardised techniques for mood induction using explicit facial images have been attempted: Schneider *et al.* (1992) have done just that. However, their process suffered again from vulnerability to demand characteristics. Participants were encouraged to try to feel a mood appropriate to either the happy or sad face that they were viewing. As above, this leaves open the possibility that subsequent inventory mood scores were reflecting what was expected rather than what was actually felt. In Schneider *et al.'s* favour, though, they found results to be relatively stable across a month and that variability within results per individual was low. This does indicate that some change has taken place although it may still be cognitive in part at least.

Habel *et al.* (2005) used Schneider *et al.*'s technique to induce happy and sad moods in healthy male volunteers. As expected, fMRI imaging showed each of the two moods to be associated with activity in different neural substrates. Although some of the brain areas identified have been linked to affect processing elsewhere, the amygdale for example, it is still unclear what part of the findings, if any, indicate cognitive processes and what affective. Participant awareness of the expected mood may still have influenced the results.

Some studies have attempted to look for mood induction effects that do not involve explicit cognitive control. For instance, Dyck *et al.* (2011) conducted an fMRI study to examine the role of amygdale in either automatic or cognitively controlled mood induction. They used two stimuli conditions: affect valent faces alone and affect valent faces plus mood congruent music. Dyck *et al.* (2011) note that previous researchers had suggested that the combination of expressive faces with congruent music would automatically induce mood without the need for explicit cognitive control of the kind indicated by demand expectations. Although both conditions were given instructions to observe the faces and attempt to feel the 'emotion' depicted, the condition containing the music was expected to evoke automatic as well as controlled affect. This allowed comparison of conditions for automatic and cognitive components but which eliminated instructional differences. Ignoring control conditions the experimental design produced two independent factors (mood type and induction type), each with two levels (happy/sad and visual/audiovisual respectively). The region of interest, the amygdale, showed differences across both factors. Across both induction types the left amygdala was activated, as expected, and was associated with cognitive strategy possibly arising from the explicit instructions in both conditions. However, in the audio-visual condition, the right amygdale was also active but no increase was found in the left (compared to the right). The authors associated this activity with automatic mood induction arising from the combination of music and facial expression. For the mood factor, the left amygdala tended to show more activity for negatively valenced stimuli than the right across both induction types.

The Dyke *et al.* (2011) study showed that it might be possible to induce moods outside of explicit cognitive control. The combination of faces and congruent music was designed to achieve this. However, even if instructions to evoke the required affect actively were absent, participants would still have been able to guess the expected mood valence from viewing the facial images alone. Therefore their technique could not be easily modified to counter expectation confounds.

2.4.5 Implicit self-esteem

What could be the benefits to therapeutic intervention of masked facial expressions? Masked facial expressions may be able to induce positive mood within the domain of social processes whilst avoiding the expectation confounds associated with other forms of mood induction. Given repeated exposure, it may be possible to affect a positive change in implicit self-esteem. Increasing a person's implicit self-esteem may be of psychotherapeutic benefit (Dijksterhuis 2004). The constructs implicit and explicit self-esteem are thought to share some properties, both are affected by social processing for instance, but they also differ on several points (Greenwald & Farnham 2000): implicit self-esteem appears to be free from such constraints as self-deception (Gur & Sackeim 1979) or impression management (Tedeschi *et al.* 1971).

It is believed that raising individuals' self-esteem is not, in itself, a panacea for multiple social ills (Baumeister *et al.* 2005). Students subject to self-esteem raising programs do not necessarily get better grades, for instance (*ibid*). However, a significant, inverse post-treatment correlation between self-esteem and depression has been demonstrated (Gardner & Oei 1981) as well as a similar, inverse correlation to anxiety (Taylor & Del Pilar 1992). A direct, causal link between self-esteem and affect-disorder treatment efficacy is not assumed by this study; 'self-esteem' is a complex term, involving many components (Hill & Buss 2006, Baldwin & Baccus 2003) and a link, if any, could be a complex one. However, viewed through the lens of information processing, implicit self-esteem might act as an informational input to affect processes that, in-turn, can feed back into self-esteem. Caution should be exercised, though, as such models are not without issue and can suffer similar charges to those presented previously regarding mood as information (see section 2.4.3 above).

It is implicit self-esteem, automatic and usually non-conscious, that equates to what Hill & Buss (2006) see as the on-going, affect evaluation of internal self-representation and is therefore regarded as a process rather than something static i.e. 'State Self-esteem' as opposed to 'Trait Self-esteem'. It may be possible to increase feelings of social inclusion, via VEs, so increasing implicit self-esteem following the 'Sociometer' model (Leary & Baumeister 2000). Here, social judgements, of the self in relation to others, act as information for input into a self-regard computation process. Such a process may become gradually 'internalised, i.e. made automatic via frequent exposure. Thus, implicit self-esteem could provide a vehicle for an affective shift with the promise of a longer-term therapeutic effect.

As noted, low self-esteem is listed as an indicator of several varieties of affect disorder including dysthemia and melancholic depression (DSM-IV-TR 2000). With atypical depression, a low sense of social acceptance is an indicator. Self-esteem is linked to social acceptance (Farnham *et al.* 1999, Leary & Baumeister 2000). Low self-esteem and dysphoric mood states have a high co-morbidity although they can have a complex relationship (Farnham *et al.* 1999, Schmitz *et al.* 2003, Bentall 2003 p247). The connection also extends into positive affect: Wright & Mischel (1982) found that happy people rate their self-esteem higher than those in a dysphoric mood state (although this paper references explicit self-esteem).

It has been demonstrated that masked text can be used to increase implicit self-esteem (Dijksterhuis 2005). Participants were serially presented with the word 'I' ('Ik' in Dutch) displayed rapidly to prevent conscious awareness. On each occasion, this was followed by a consciously perceptible positive or neutral term. After several trials, it was found that scores had increased across several measures including an implicit association test (Greenwald & Farnham 2000). From this, it was inferred that implicit self-esteem had increased.

We also know that changes in state self-esteem can be engendered by static, planar images alone (Baccus *et al.* 2004). Pairing personal data with images of smiling people making eye contact with the viewer can increase implicit self-esteem, at least if modulated via interactive, attentional mechanisms. This can even be extended to train participants to be more vigilant and attentive for such socially affirmative imagery (Dandeneau & Baldwin 2004, Dandeneau *et al.* 2007).

Masked priming may be able to provide an even more useful route to non-conscious processes than the explicit tasks used in other CBM work such as Baccus *et al.* (2004), Dandaneau *et al.* (2007) and Beard *et al.* (2011). Masking works at a perceptually non-conscious level. We may be able to use it to interrupt any negative affect processing required for attention to non-affirmative imagery, the kind of processing thought to underlie some affect disorders such as social anxiety (Clark 2001). Interrupting the non-conscious processes that confirm low self-regard (Leary & Baumeister 2000), we leave a person's implicit self-esteem open to change in a more positive direction. This in turn could lead to a virtuous circle where positive cognitions feed reciprocally into positive affect. In this way, new processing can itself then become stable. Embedding masked priming within a VE would bring all of the advantages previously outlined (see section 2.2.3.1 above) and may prove to be a very powerful tool in the treatment of affect disorders via changes to implicit self-esteem.

2.5 Literature review Conclusions

That therapies using VEs and VR can be effective treatments for some psychiatric conditions is established. Phobias, PTSD, along with various forms of anxiety and depression have all been shown to be susceptible to improvement using this approach. Like all therapeutic tools though, virtual systems have their limits. They may not be sufficient in themselves to treat the above conditions and indeed may not have any impact at all for some users. Therefore, there is a case for looking at how to improve virtual therapies, to not only increase their potential effects but also widen their scope to include more users and even different conditions to those currently treated. A potential approach for the improvement of virtual therapies would be to look beyond their use as simulators of the world, looking instead for techniques and effects that have a therapeutic potential but are not perhaps possible in the real world.

CBT claims a corrective effect on cognitive processes, not normally accessible to explicit manipulation, and from which affective benefits follow. CBT does have a good record of treatment success, comparable to psychopharmacology but with a better long-term prognosis. However, it may not really be correcting any implicit processes at all, hence its failure rate that, although perhaps lower than other 'talking cures', is still capable of improvement. Virtual systems do appear to possess the potential to access implicit processes that may not be possible by other means. This presents the possibility that such systems could directly manipulate those processes in a way not available to other types of stimuli such as the predominantly linguistic based techniques encountered in CBT.

CBT is grounded in the proposition that changes in cognition can change affect. We know that affective primacy is also possible and, as such, can give rise to changes in the opposite conceptual direction. We also know that the valence of background affect can give rise to particular perspective biases in cognitions: for example, if a person is in a positive mood then their thinking, it is argued, is also positive in character, such as readily attributing success to their own effort or abilities. If it were possible to change implicit affect, via VEs or VR, then we might be able to engender a more positive implicit attitude, for example a greater sense of implicit self-esteem. With such an affective background, a patient's negative cognitions may become more malleable, enabling the challenges put to them by a CBT therapist to take effect.

Visual masking research seems to demonstrate a way in which implicit processes might be engaged with. Masked facial expressions have shown promise as being capable of reaching those implicit, affective processes. Embedding these within a VE brings with it the added advantages of providing an engaging and distracting stimulus as well as the potential amplification, transfer and extension of affect change due to Psi. In order to create such a stimulus, technical limitations may need to be overcome such as depth disparity unmasking effects. In addition, parameters would need to be investigated for possible modulation effects. At its core, though, the research would need to be focused on investigating the affective effects of the VE.

The above can be distilled down to form specific pointers towards research: The non-simulation properties of VEs may be of use to disrupt negative, non-conscious, affective processes. There may also be an interruption of those cognitions, such as those involved in creating and upholding implicit self-esteem, that support and are supported by those affect processes. Visual masking promises one way in which we might exploit such properties to reach those implicit processes.

It may be possible to bring about changes in affect cognitions using masked primes within a VE. Masked facial expressions may facilitate this, with the assumption that a valence shift in affect such as mood being congruent to the expression used. Specifically, exposure to masked, positive facial expressions may engender a change towards a positive state affect. Taking into consideration points drawn from the topic review, the masking paradigm chosen was non-latteral, backwards pattern masking with a matching surface texture on both target and mask. Because of the assumption of the importance of Psi, non-planar target faces were used, as these would appear to be more realistic, i.e. plausible, than planar objects even when processed implicitly. Section 3.4 in the Methodology presents a more detailed discussion of the initial parameters for the masking form used. Beyond this, part of the aim of the experimental work was to elucidate further the parameters needed to achieve the wider goals of investigating the possibility of using this kind of masked priming as a therapeutic aid.

At this point, it remains to be shown that visual masking will work within a VE and whether would be possible to use such a method to change implicit, affective or cognitive processes. It is possible that there will be a direct effect via affect contagion or an indirect effect via cognitive modification. Aside from testing the initial choice of masking paradigm and extending parameter knowledge, including possible effect modulators, the experiments must also establish if the valence of masked stimuli can have an effect on participant affective cognitions.

3 Methodology

This chapter will introduce various materials, measures and other methodological resources that are common to all six experiments. Commonalities include the hardware platform, the software platform and the VEs constructed from these. In addition, the type of visual masking used and participant issues will also be discussed.

Methodological points, specific to a particular experiment or experiments, will be discussed within that experimental chapter or in the chapter of the first experiment within which the points arise.

3.1 The Hardware Platform

All of the experiments in this work were performed using the same core computing and display hardware, with additional machinery added as necessary. The equipment was a desktop-based set-up and in each case faced a blank wall to minimise distraction with an additional background being provided by an A1 sheet of black cardboard.

The virtual worlds were run on a 4GB, 2.66 GHz quad core PC with a graphics card comprising of the Nvidia GeForce 260 chip set. The stereoscopic display was produced using a high frame-rate monitor, the Samsung SyncMaster 2233RZ, along with the Nvidia 3D Vision active LCD shutter system incorporating an infrared hardware driver.

The shutter glasses were light meter tested and both lenses presented an approximately

50% drop in light levels when active, as would be expected with an active shutter system. Although 50% may appear to be a large amount, users experience counteractive pupil dilation in response to these lower light levels. Additionally, monitor brightness can be increased to compensate still further for the light drop.

The stereoscopic display works well, with very little image cross-talk. The monitor has a 120hz refresh rate allows for a stereoscopic display of up to 60 frames per second, with a response time of 3ms (grey to white to grey) to minimise cross-talk. The stereoscopic display worked well for keeping participant attention trained on the task as well as arguably increasing some sense of presence for participants (Ijsselsteijn *et al.* 2001, Freeman *et al.* 1999, Ijsselsteijn *et al.* 1998).

Ideally, the timing of shutter glasses should approach a 50% duty cycle whereby each lens should alternate between being either open or closed 50% of the time, in an opposite state to the other. There should be little or no time when both are either open or closed concurrently. In addition, signal response latency should be in the microseconds. Tests, independent from the manufacturer (Woods & Helliwell 2012), suggest that the infrared signals from driver unit to glasses indicate some timing crossover but this is in the order of microseconds only.

As the Nvidia 3D Vision system is proprietary, as opposed to open source, no detailed specification is available. However, based on the above, it is assumed that each eye receives close to the ideal with any cross-talk being in a factor of microseconds. As the perceptual effects used in these experiments are sensitive to timings within milliseconds it is a reasonable assumption that they will not be disrupted by any very rapid, microsecond discrepancies.

Due to the high frame rates demanded by visual masking (see section 2.3), it was not possible to use a HMD or CAVE-like facility. Unfortunately, this almost certainly had an effect on the participants' sense of immersion and affected their experience of PI (section 2.2.3.1), although perhaps having less of an effect on Psi. This presents an opportunity for future research with faster immersive technology such as a high frame-rate HMD.

Finally, to expand on a point touched on in section 2.3.2, the images produced by shutter glass systems, and indeed most stereoscopic displays, separate eye convergence from accommodation on the focal plane: focus remains fixed with stereoscopic imagery whereas with actual depths encountered in real life it does not. Aside from being another difference between the real world and VE simulations, convergence and focus separation may be a factor in the discomfort sometimes reported when viewing such images. However, the stereoscopic system used in the current work appeared to be comfortable for participants. Out of approximately 140 participants, only two expressed any discomfort and only one of those failed to complete.

3.2 The Software Platform

The software platform used was the Unreal Engine3, in the form of the Unreal Tournament engine and, later on, the Unreal Development Kit. Being a gaming platform and not a virtual reality platform per se, it has the necessary optimisation to produce and control frame rates in the order of milliseconds that the masking requires.
The Unreal Engine3 can be limited to a specific output that would then be doubled by the 3D Vision system. For instance, 60 frames per second would be reproduced as 120

frames per second stereoscopically. Therefore, if a target object were presented for one frame, in a 60 frame per second scenario, each eye would receive that specific frame, adjusted with the appropriate perspective, for approximately 8.33 milliseconds, followed by a closed lens of the same duration. During this closed lens period, the other eye would perceive the same frame, again adjusted for perspective, for the same duration of 8.33 milliseconds. Overall, each frame would be perceived for approximately 16.66 milliseconds, split between each eye.

3.3 The Virtual Environments

3.3.1 Form and Content

The virtual worlds themselves were realistically rendered (see section 2.2.3.1) with dynamic shadows, baked global illumination and photorealistic textures. Their geometry was kept uniform and structurally simple to avoid possible navigation bias (Loomis 2002) and generally consisted of a series of courtyards through which the participants moved, each terminating in a pair of doors through which the next series of courtyards could be accessed, see figure 3.1 below. The courtyards were open to the sky so as not to introduce any confounding factors associated with claustrophobic reactions (which participants may not have been aware of). The only exception to this was the world created for one of the conditions in the fifth experiment. Here, an open world afforded views across grass towards water with mountains beyond. This also included some virtual foliage and masonry. The worlds were also kept as free as possible from distracters that could increase attentional load, again with the exception of the fifth experiment.

Overall, the various worlds were optimised during construction to place the minimal load possible on the graphic hardware whilst still supporting the experimental goals. In this way, it was possible to achieve 60 frames per second visual masking. The Unreal system performed well, allowing optimisation of the procedural logic behind each world that further aided the balance between low graphic load and experimental goals. For example, several worlds used a player 'teleportation' procedure: when a participant reached a certain point in a world, they would be placed back into the original starting courtyard, all courtyards being identical in look. In this way, worlds that appeared to contain many courtyards could be modelled using only a few, thereby reducing memory loading.

Figure 3.1. Images of various VEs used.

Generally, one or both of the doors in the VEs acted as planar masks, rendering target objects behind them explicitly imperceptible. The viewer perception was of one or both doors appearing to blink, in the doorways at the end of each courtyard. Normally, the participants had to navigate through the targets in order to progress even if they were consciously unaware that they were doing so.

The keyboard arrow keys were used for participant movement within all of the VEs. Depressing the up and down arrow keys would cause forwards and backwards motion respectively, at a fixed maximum speed. Releasing the keys would cause the motion to stop. The left and right arrow keys were used to revolve left or right respectively. Therefore, aside from forwards and backwards camera motion, participants were allowed to change the camera heading and so access all areas open to them within a particular VE. Roll and pitch changes were not available to participants, camera movement being constantly upright and parallel to the ground plane. It was found that this simple arrangement was quickly grasped, even by participants with little or no experience of VEs. Camera speed was capped to allow a number of target exposures when traversing a room, depending on the experiment. Hesitation and changes in direction by the participant often meant the traverse times were longer.

The VE walls acted as constraints to movement, keeping participant camera views within the confines of the courtyards and short corridors between them. The only exception to this was in Experiment 5 (see Chapter 8) where some side-walls were invisible in one of the environments, to allow for views across the virtual world. Nevertheless, the main courtyards within which the participants could move were large enough so as to not have an impact on movement: they rarely collided with the walls, invisible or not, when practiced in navigating the VEs. The times when collisions mostly occurred were to be found when participants navigated the short corridors between the courtyards. The walls defining these corridors were visible in all of the VEs within all of the experiments.

A note on the terminology of virtual characters: for the purposes of this thesis, an avatar is regarded as a human-controlled, virtual character and an agent as a computer-controlled, virtual character. The latter can imply some form of algorithm intelligence but not necessarily, as they can be static or merely repeating a looped animation for instance. A hybrid virtual character is regarded as a mixture of the above two whereas a virtual character is seen as any of the above three.

Some of the masked objects used in the experiments were faces carrying affective expressions. The validity of the various expressions used was established in small pilots. The faces were presented in a non-masked, i.e. visible, way and people were asked both a forced choice question and an open question regarding the expressions. The forced choice question was 'do you consider the face to have a positive, negative or neutral expression?' The open question was simply to ask 'what expression does the face have?' Only those expressions that produced the same, clear answers repeatedly were used in the experiments.

3.3.2 The Development and Construction Process

The creation of the stimuli used in this thesis began with simple video prototypes. The intention was to learn lessons in a straightforward, non-interactive environment and then transfer knowledge gained to an interactive one at a later stage. However, the rates at which standard computer video is displayed at, as opposed to any underlying frame rate supposedly recorded in a piece of footage metadata, are usually from 15 to 25 frames per second. At these speeds, the least time a target image could be on screen for is around 40 milliseconds. It was noted that this was too slow to mask the targets, even if they were not fully distinguishable. Even when the frame rate was pushed to 30 frames per second, the targets would be displayed for about 33.33 milliseconds, which again was too slow to mask the target fully. The paradigm adopted for these early trials was mainly backward masking (where the mask follows the target) although forward masking was tried but found to be even less effective.

There was also an issue with the LCD display used in that, although the manufactures claimed a display frequency of around 75Hz the refresh time was such that some image crossover must have occurred from frame to frame, enough to affect the rapidity of target image display. Ordinarily this would not matter but display timing is crucial for masking effects. It was felt that this route would not be able to simulate the effects of, for example, a tachistoscope used in various early priming studies, and so was abandoned.

It was decided that the rapid frame rates possible from a game engine would be advantageous for the creation of the necessary stimuli. Although this meant developing in a complex, interactive 3D environment from the beginning it was felt that this would be the best way of easily attaining the frame rates required. This allowed the development of the stimuli to proceed from a stereoscopic perspective also. It had been an intention to eventually aim for some form of immersion/attention holding by using a stereoscopic system. The monitor sourced for this stage of the work was capable of high frame rates (120Hz) and importantly had a low refresh time to avoid the problem of image-cross over.

The game engine chosen was the Unreal3 engine. This was for a number of reasons. Firstly, it is a highly capable engine, used to produce many 'AAA' games. As such is optimised for the rapid display of game elements. Secondly, the author was familiar with this engine and so could move to development rapidly without learning such a complex piece of software from scratch. Lastly, the full engine is available free of charge as both an accompaniment to the various games created from it and, latterly, as a downloadable and updated version: the Unreal Developers Kit. The Unreal3 engine also works well with the stereoscopic system sourced for this work.

A frame rate of 60 was used initially as it was well within the capabilities of the engine, was typical of the rates expected in games and matched prior research. This frame rate would mean the minimum a time a target object could be displayed for would be approximately 16.66 milliseconds, a figure close to those used for masked priming experiments using static, 2D stimuli. Backwards visual masking would therefore be feasible using such a system.

Initially a custom material was created, using Unreal's material creation system. The material was capable of rendering the object it covered both non-visible and visible based on a timed cycle. It was intended that this material type be used for both targets and masks, with the timings adjusted as appropriately. However, this method proved unreliable, in terms of both timing and control, and was subsequently abandoned. Following this, 'Kismet', Unreal's node based scripting system, was elected to be used alongside a combination of UnrealScript and Unreal's console commands where necessary. This proved effective as different timings could be expressed as various, developer-adjusted variables. Therefore visibility times for both target and mask could be adjusted to find optimal values, as could the non-visible times. The outcome of experimenting with various settings resulted in timings that approximated the various times quoted in the masking and priming literature: between 10 and 20 milliseconds per target and between 250 and 500 milliseconds for the masks. As 60 frames per second (equating to 120 frames per second in stereo vision) meant a single frame could be displayed at roughly 16.66 milliseconds it was decided that targets should appear for a single frame, which should be adequate for masking purposes.

Regarding the development of the targets, the initial aim was to achieve a more life-like form than that of previous masking studies, which used 2D, static and often greyscale images. The static nature of the targets would be addressed later, in experiment 4, see Chapter 7. However, from the begining a 3D, ie. non-planar, form was considered for the targets, the belief being that this would contribute to Psi albeit in an implicit way. It was decided to keep the masks as planar objects to lessen the processing load on the engine and graphics card, lowering the possibility of mistimings leading to targets becoming visible. Non-planar targets with planar masks became the standard format upon which later the experimental stimuli would be built. It was felt that later work, beyond the thesis, could study alternative paradigms such as both mask and target being non-planar.

At this stage, the target objects consisted of various non-planar primitives such as cubes and spheres. These targets were tested on a variety of planes relative to the mask in an empty VE. It was found that placing the target just behind the mask, with the frontal polygons of the non-planar target almost touching the back of the mask, was optimal. The distance between the target and mask at the closest point was 1 Unreal unit, which is ‘officially’ equated to about 2cm in real world terms, although this is arguable as VEs can scaled according to the display equipment used and so present differing illusions of space. The masking doors were taken from the Unreal assets provided with the editor and as such retained their texture. This texture was applied to the masked objects as well and, as such, was arbitrary to them. It was found that when targets and masks had markedly different textures then the masking effect failed and the targets became visible.

A standard digital camera was used to capture images of the targets that were effectively masked to the viewer. This was to ensure that the engine was actually rendering the target frames. It was also found to be possible to manoeuvre the VE 'camera', i.e. the viewer's point of view, to a point just behind the masking door and view the unmasked target. From this position it was possible to see the target being rendered, experienced as a brief flickering of the object: this was found to be enough to identify the object and confirm generally that the masking was working. It was also found at this point that unmasked, non-planar targets were more readily identifiable using a stereoscopic display compared to a standard monoscopic display yet the masking worked just as effectively in both. Thus, it could be speculated that the implicit identification of an object would be more effective with stereo vision.

Up until this point, the author had acted as a tester, experiencing the masking effect and adjusting the timings as necessary. Following this initial stage of development, a simple VE was built, consisting of a flat area of ground, an archway, a masking door and target objects whose various forms were household objects and faces. The system was then piloted with several acquaintances of the author agreeing to act as viewers. They reported a similar experience to the author in that the targets were not explicitly visible, provided the target and mask textures matched. The same unmasking effect was noticed when target and masks carried disparate textures, such as when each was predominantly coloured in a complimentary hue to the other.

It was at this stage of development that expressions on the faces of target objects could be tested to see if a consensus was there regarding the affective content. Variations of several expressions were shown and those that were consistently judged to be representative of a certain expression were retained for use in the experiments. For the smiling expressions a series of videos were shot of people looking at a rapidly revealed, amusing image: a cat with an orange peel hat. This method was able to capture fleeting but real 'Duchene' smiles whose configurations formed the basis for a series of smiling, virtual character faces. For other expressions such as angry-looking faces, images of expressions were sourced and their configurations averaged by hand to form the final expressive faces. All faces were ultimately formed in Maya, a 3D modelling package, before being imported into the Unreal system. This was package was also used to create the other targets used in the experiments such as household objects.

The final stage of development was to construct a larger VE whose geometry contained several rooms and that would serve as a prototype for the final experimental VEs. As it was intended that door choice would be a behavioural metric, a series of rooms connected via pairs of doors was constructed, one door leading to one room, the other door to another and so on. Photorealistic textures were used on the surfaces to increase the illusion of place. At this point it was noticed that, if Kismet became too loaded with tasks the engine might not render every target frame, possibly due to culling routines within the engine code itself designed to prevent deterioration of performance. It was therefore important to bear this in mind when constructing the procedural logic that underlay a VE.

3.4 Visual Masking within the Virtual Environments

3.4.1 Basic Issues

Visual masking theory, alongside some of the technical aspects of the phenomena, has been covered extensively elsewhere (see section 2.3). Only those aspects as they specifically apply to the series of experiments contained in this thesis will be considered here.

Although some visual masking research has looked at external depth disparity between target and mask, to the author's knowledge there has been none that used stimulus images with an internal depth disparity, and no work directly within VEs. Many initial stimulus builds were needed before a useful masking effect could be constructed. The final paradigm settled upon, and common to all experiments, was to use non-planar targets with planar masks. One of the general aims of the thesis was to examine the effects of differing target forms, especially facial expressions. To this end it was felt to be expedient to make the target carry the internal depth disparity; this could then be used to model realistic forms on the assumption that these would be responded to more readily by participants, even if implicitly.

Regarding the construction of the masking effect, either or both the target and mask could have been non-planar forms. It was known from the literature that increasing depth disparity between target and mask in a stereoscopic field could cause binocular unmasking. To avoid this we would need to sit both on the same depth plane. This makes sense if both target and mask are of the same form e.g. the same three-dimensional mesh. However, this needs some consideration if one is planar and the other mesh has volume. Supposing that the mask were planar and the target had internal depth disparity. We would need to establish whether the target should cut through that plane, sit in front of it or behind it. A series of non-planar objects were created, intended for later use as targets in the experiments (see Chapters 4 to 9). These were either faces or domestic objects and could be scaled up to fit into the doorways, just behind the masking doors. All target meshes had a polygon count of below 4000 triangles, most consisted of several hundred. They were built to a rule of thumb that was to make their depth no more than one-half that of the mask width as presented perpendicular from the participant's initial line of view. This rule was arbitrary, as the perceived depth of the objects changes on approach etc., but acted as a starting point. Initial trials with the system described suggested that it was possible to use non-planar objects sat behind a planar mask and create a backwards masking effect.

More of an issue than depth was the subject of surface texture of the target and mask. Previous work had shown that a similarity between target and mask was helpful to the masking effect (see section 2.3.2 and 2.3.3). During construction of the masking VEs, the strongest effects were found when mask and target were rendered with the same surface texture.

One issue of concern initially was that the textures would be displayed differently on a planar mask compared to a non-planar target. Specifically, that placing a surface texture on a non-planar object would distort the texture geometry and so effectively produce the impression of a different texture to that on the planar mask. The textures for these experiments were created as two-dimensional images. When placing them onto planar objects no distortion will occur, unless the images are subsequently scaled differently in each dimension. The masks used in the experiments used the same scaling in both dimensions so no texture image distortion occurred. However, when such a texture is rendered onto a non-planar object, it is liable to distortion. This is due to the non-planar coordinate geometry used by the rendering engine to map a texture onto the surface of the object. Many differing forms of projection can be used to apply this mapping. All will distort the image in some way, compared to a planar map, if the mapping is not viewpoint dependant. For the target objects in these experiments, a standard, non-viewpoint dependant, cylindrical mapping was used. It was found that, the slight distortion in texturing image geometry that this mapping produced appeared to have no effect on the masking effect.

As texture image colours are not affected by the mapping, they remained the same between target and mask. It is conjectured that the frequency of image detail was not significantly affected by the mapping to create the perception of there being a different texture on the target.

Another issue of concern arose during the design and construction of the VEs: the relatively free movement that participants would enjoy within the VE might undermine the masking effect. As noted, participants were able to move forwards or backwards, stop at any time or change direction. This meant that not only did the size of the target and mask alter relative to the participant field of view but that they did so in an irregular way. Additionally, participants could also alter their heading which meant that their approach angle to the target and mask varied as did the proportions of the target and mask (from the participant's point of view). However, it appeared that the masking effect developed was sufficiently robust to withstand these variables.

3.4.2 Why Use Backwards Pattern Masking?

Backwards pattern masking was the format used for several of the priming studies that acted as starting points for this thesis (see section 2.3.5) and as such constituted a good starting point. The images used as targets and masks constitute patterns as opposed to simple geometric figures. One reason why backwards masking has gained much use could lie in its ability to mask more effectively than forward masking, as demonstrated by metacontrast and paracontrast (see section 2.3.2). Certainly, this is the form that appears to be used most often in masked priming literature.

It was decided that, if the VE ran at 60 frames per second, the target could be displayed for one frame, approx 16ms, per masking cycle. This could then be followed by the mask, shown for around 250-500ms. Finally, an ISI of around 250-500ms would complete the cycle. This cycle could then be then repeated continuously as needed in a particular experiment.

Using the hardware and software combination above, it was possible to construct a VE that contained a backwards masked, non-planar target. As noted, the paradigm recurrent throughout the experiments used either or both doors as planar masks, with the target objects sat just behind them. A digital, SLR camera was used to photograph the screen to show that the target objects were being rendered, despite appearing imperceptible. It was also possible to manoeuvre the viewpoint of the VE camera to a point just behind a masking door and observe the effect of a target directly, rendering it explicitly viewable.

3.5 Participants

3.5.1 Population

The participants for these series of experiments were all adults recruited from a number of sources including the UCL Psychology Department subject pool, poster advertisements and word-of-mouth. This enabled a diverse group to participate, beyond the student population. This in turn strengthened the argument that any results found were generalisable to the wider population.

The diverse participant spread included occupational backgrounds outside of the student population. The population also showed something of a global nature in that many people from beyond the UK participated. Although not explicitly tested for, there appeared to be a spread of educational backgrounds. Most were interested in the research but others were less so. Gender was approximately balanced.

The main similarities within the population were that they were all self-selecting and most were domiciled in London, UK, at least temporarily.

Each received either five or six pounds for expenses, depending on the experiment. Monetary reward was not used as a variable to distinguish between conditions.

3.5.2 Screening

Participants were screened for a number of factors. The main one was for conditions wherein epilepsy or similar could be triggered by flashing imagery. This would be more of a problem for experiments using CFS, a technique not used in this work.

Colour-blind participants (of any category) were also screened out. This was particularly important for Experiments 3 and 5 where colour played an integral part in the experimental design.

Participants also had to have a good grasp of the English language in order to correctly understand and carry out task instructions. This was also an important ethical point, as all participants needed have a clear understanding of understanding of what participation in a particular experiment meant and involved.

It was a condition of involvement that no participant performed in more than one experiment. Each was asked not to discuss their experiences with anyone else that they knew would be attending until after all had completed their participation.

During the preliminary activity before each experiment, participants were asked to confirm that they could get a sense of the depth into the screen when looking through the stereoscopic glasses. This effectively acted as a self-screening measure for stereo perception.

3.5.3 Initial Questionnaire

This questionnaire was filled out by each participant before beginning an experiment but contained no personal identification data, see Appendix 1. It was intended to identify any possible, collective-characteristic confounds and provided the basis for any necessary statistical adjustment needed.

Participant ‘handedness’ was recorded. This would allow results to be adjusted to take account of any navigation bias. Such a bias could have acted as a confound, in particular where participants had to choose between moving through a left or right hand door. However, the overwhelming majority of participants were right-handed and the results from left-handed participants did not show any marked difference from the right-handed group. It was therefore concluded that handedness did not constitute a confounding variable for any of the experiments. This concurs with research that shows navigation bias is not related to handedness (Souman *et al.*2009).

Aside from standard demographic information, the questionnaire collected data on the participants’ experiences of VEs and related phenomena. This included time spent passively watching stereoscopic events, for example seeing a 3D film at the cinema or on a home system. In addition, time spent interacting with VEs such as playing 3D computer games, as well as time spent on computer games in general was recorded. Also noted was any time spent within an immersive VE.

Many participants had some experience with computer games of various descriptions in the month prior to participating although very few stated that they spent a large degree of time playing. None of the main participants had any experience of an immersive VE in the month prior to an experiment with the exception of two computer science students.

3.5.4 Ethics

Masking experiments can entail some form of omission in the experimental description initially given to participants. This does not necessarily mean deceiving them as to the nature of the experiment but, rather, not fully informing them of all of the details. Outcomes could have been prejudiced through demand characteristics based on top-down cognitive interference. It was therefore reasonable not to inform participants, where appropriate, as to the nature of the masked targets until their involvement was over. Even then, it might have been judicious to wait until all of the trials were over as word may have spread as to the nature of the experiments. However, this could have been seen to be in conflict with ethical considerations and so was not done.

In terms of safety, throughout decades of study of masking phenomena there have been no reports of adverse or detrimental effects to participants. There is no reason to suppose that the introduction of masking into VEs would be any different. VEs also have a safe track record, with the possible exceptions of motion sickness and mild eye strain. The latter is more of a problem in fully immersive systems and can be lessened by not asking participants to spend too long on a particular VE task.

The experiments in this thesis used deception by omission and some had affect-loaded material. They were all passed as ethically acceptable by the University College London ethics process.

3.5.5 Possible Participant Confounds

Experimenters can sometimes impart information unwittingly to participants, known as the ‘Little Hans Effect’. The name derives from a seemingly numerate horse in early 20th Century Germany. It is thought that Clever Hans was actually drawing information from his owner’s behaviour that nevertheless gave the illusion of the horse being able to perform feats of arithmetic (Pfungst 1911).

For the current experiments, measures were taken to ensure that a similar effect did not occur. Each time a participant performed a task, the experimenter would turn to face away from the screen. The experiments relied on in-world cues to let the participant know when a particular trial was over. These cues were explained to participants at the beginning of the tasks.

Demand characteristics can arise from participants knowing whether they are in the control condition or an experimental treatment. Within-participant designs can suffer from this particularly as participants can directly compare conditions. One of the benefits of testing masked objects is that participants are explicitly unaware of the condition in which they are currently performing. It could be argued that their non-conscious processes do know and therefore will respond accordingly. This can be countered by pointing out that this may actually be the point of the experiment.

The within-participants design used for the first four of the experiments has several advantages. They can control for individual differences and necessitate fewer participants to realise a statistically significant result, to name two. However, there are some disadvantages inherent in such a design. Order effects can emerge, chiefly training effects. Participants can become adept at particular tasks and if, for example, all control version tasks are consistently performed before the experimental versions then an effect may be falsely indicated (a type 1 error). To solve this, each of the four, within-participant experiments employed techniques to randomise the order of conditions; either by interspersing conditions randomly throughout a particular VE or randomising task order (see Chapters 4 to 7). The two between-participant experiments would not suffer from training effects but other order confounds might be an issue.

Therefore, for these two experiments, participants were randomly assigned to various conditions on attendance (see Chapters 8 & 9).

3.6 Statistics

This thesis relies heavily on statistical methods commonly used in the Social Sciences. Both descriptive and inferential statistics are used to support or refute the hypotheses. The two main inferential methods used were the Students t tests and ANOVA variants. The t tests use both paired sample and independent samples, on Experiments 1 and 6 respectively. Repeated measures ANOVA was used for the within-participants experiments 2, 3 and 4. Experiment 5 used a mixed-model ANOVA variant. The standard significance cut-off point of 0.05 was used within this thesis.

3.7 Concluding Remarks

We conclude this Methodology Chapter with a table matching the research questions with the following program of experiments.

Research question a. Initially, the problems to solve will be technical ones: can visual masking be made to work within a VE? There are aspects of the masking effect that may present difficulties here such as inhibition by stereoscopic depth disparity.	Experiment 1 & pilots
Research question b. The next question will be to ask whether it is possible to elicit a response to emotionally valent, masked stimuli in a VE. Will participants respond in a measurable way to visually masked, emotive expressions? At this stage, some initial modulation factor exploration may be useful.	Experiment 2
Research question c. If effects are found, the following work should be again of a technical nature: what parameters, if any, could be manipulated in order to modulate any effects or increase the reliability or consistency of those effects. Here masked, expression dynamics could be introduced within a VE. Will these be understood as carrying the valence intended? Other parameters should include the surface qualities of masks and targets such as colour saturation or texture legibility tested against affective salience.	Experiments 3 & 4
Research question d. The final problem will be how to use visual masking in a VE to change a person's affective state, specifically altering their mood in a positive direction, directly. Alternatively, can cognition be changed implicitly, with the expectation that this will, in itself, lead to a change in affect.	Experiments 5 & 6

Table 3.1. The research questions mapped to the program of experiments.

4 Experiment One – Visual Masking in a Virtual Environment

4.1 Abstract

Visual masking techniques produce strong effects but the stimuli used are often co-planar and with no internal depth disparity. Prior research has shown that external depth disparity, the distance between target and mask, tends to inhibit any masking effect. This study investigated not only whether a virtual object's internal depth disparity inhibits masking but also whether it is possible to achieve visual masking within a VE in the presence of either depth disparity type. Non-planar, virtual objects were used as targets and planar objects were used as masks within a stereoscopic VE. Forced choice tests were administered to measure any masking effect. An effect was found that showed it was possible to construct target objects with internal depth disparity in a VE. In addition, the effect was tolerant of some external depth disparity.

4.2 Introduction

The first aim of this study was a technical one, which was to establish whether the masking effect could work within a VE. This was measured by the imperceptibility of the target objects. The second aim was to see if these could influence behaviour despite being unseen. This utilised a behavioural measure in the form of a choice preference task.

Pilot trials were carried out to establish the parameters needed to create the conditions for backwards, pattern masking to take place, using virtual non- planar targets and planar masks. Either or both the target and mask could have been non-planar in form. We know from the literature that increasing depth disparity between target and mask leads to binocular unmasking so to avoid this we would need to sit both on the same depth plane (De la Rosa *et al.* 2008, Moraglia & Schneider 1990). As noted in Chapter 3, the format used experimentally involved virtual objects whose depth was approximately no more than one-half of the mask width, perpendicular to the participant's initial line of sight. The masking effect appeared to work when the target object sat just behind the mask, relative to the participant starting point.

Some non-conscious, visual processing may take place. Therefore, the participant may be aware that there is a masked object aligned to a certain door. However, the target objects chosen are modern, complex and innocuous. These should therefore not be able to be rapidly processed non-consciously by our vision system, in the way that faces and other organic forms may be, for example. This is also why participants were asked to proceed fairly briskly. It was therefore expected that, for unseen objects, the door may be correctly recorded whereas the object type may not. This would indicate the presence of limited, implicit visual processing.

4.3 Method

This experiment used a 2 condition (masked object vs. no masked object) within-participants design, with a binary value preference measure (left or right door) to test for between-condition effects. After a practice period, including one practice trial, participants completed 28 trials in total with 14 repetitions per condition.

4.3.1 Participants

Twenty-four people were recruited for the experiment, drawn from the UCL psychology subject pool, posters and word of mouth. Two participant data sets were discarded for non-compliance leaving twenty-two data sets for the final analysis. These represented thirteen, right-handed females, eight right-handed males and one left-handed male.

4.3.2 Materials

The hardware and software platforms used were those described in the Methodology set out in Chapter 3.

Following pilot trials, it was decided that the target should be displayed for timings around 17ms, followed by the mask shown for 500ms and finally an inter-stimulus gap of 500ms. The cycle then repeats continuously. These timings are close to those used by Winkielman & Berridge (2004) and so constituted a good starting point. As the frame render time of Unreal engine being used was approximately 16.67ms (60 frames per second) the nearest multiple of frames allowing an object to display for over 500ms would be 30 frames. This meant that both the mask and ISG would be displayed for about 500.1ms. This was divided into 250.05ms of display per eye (see section 3.2 of Chapter 3).

The VE was realistically rendered with dynamic shadows and photorealistic textures. The geometry was kept simple to avoid a possible navigation bias. Each room consisted of stone-like walls and a floor, open to the sky. The VE consisted of a series of virtual rooms at the end of which were two doors, see figure 4.1 below. The doors oscillated rapidly between being visible and non-visible, producing a ‘blinking’ or ‘flashing’ effect. The doors were planar masking objects, initially parallel to the virtual camera plane i.e. the starting point of the participants in the VE. One of the doors, either the left or the right, was masking a target. The masking door varied per room.

The non-planar targets were a variety of recognisable, innocuous, yet complex objects such as domestic or mechanical items, see figure 4.1 below. For this study, faces or other organic forms were avoided as object types. This was to avoid using any objects that might carry a population-wide affect bias. Each of the objects had an approximately equal internal depth disparity in the axis perpendicular to the virtual door plane. The objects all shared the same surface material: that of the masking door, as figure 4.1 shows. This was found to be important in pilot trials. The front-most polygons of each object, from the participant’s view sat as close to the depth plane of the doors as possible, without protruding through that plane.

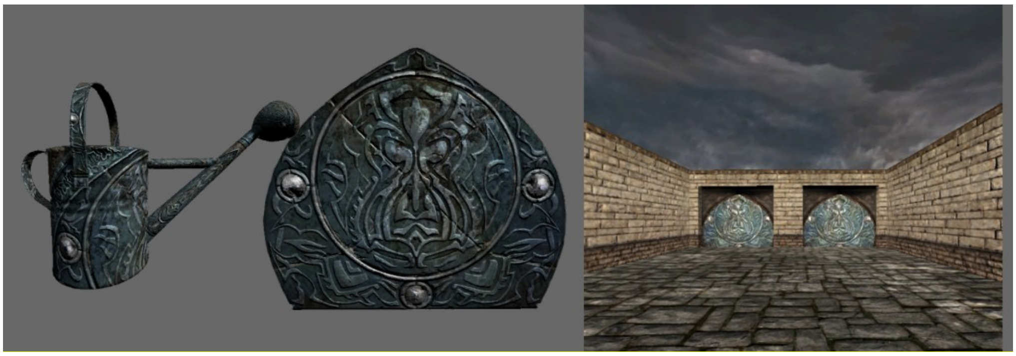


Figure 4.1. On the left, a masked object and a masking door and on the right, the VE.

Several virtual objects were used: a chair, a teapot, a watering can, a tap, a lamp, a telephone, a cup, a pan, a bicycle and a television. This ensured that confounds were not introduced by using one kind of object, such as individual participants having differing reactions to one object type. Appendix 3 presents images of these objects in the form of a reference sheet given to participants for use during the experiment. On the sheet, the objects are rendered in neutral grey tone, using a simple Phong shader to delineate form. The objects on the reference sheet were also displayed at a different orientation to how they were oriented as virtual targets within the VE. This was to make sure that the results would not be based solely on shape recognition but on object recognition, thereby indicating the type of processing involved.

Both experimental and control conditions consisted of fifteen rooms. The experimental condition consisted of the VE with masked objects, the control condition used the same VE but without any objects. All other factors were the same for both conditions in the VEs.

Participants were also given a response questionnaire to answer during the experiment, see Appendix 2. Within each room, the questionnaire asked if any object was visible and what type of object it was, referring to the object reference sheet (see Appendix 3). Even if participants could not explicitly perceive any object, and therefore its type, they were asked to make a guess as to the type.

4.3.3 Measures

Data were gathered from two main sources: a response questionnaire and door choice figures (see below). The main independent variable arose from the randomised, balanced positions of the target objects: either behind the left or right door. As part of their task, participants were asked to pass through the door that they thought contained the masked object, see section 4.3.4 below. Which door the participant then chose to pass through, either left or right, formed the main dependant variable: the door choice figures. This was designed to test for any implicit positional awareness of the target objects. The visibility and object-type data gathered from the questionnaires were intended to indicate the extent of any explicit object recognition.

The data from the first of the fifteen rooms were to be discarded for all participants as they were deemed to be acclimatising themselves to the VE at this point and so may produce atypical behaviour. This approach then had to be repeated for both conditions to balance the results. Therefore, data from fourteen rooms per condition per participant were gathered for the final analysis. The study was within-subjects in design as all participants experienced both the experimental and control conditions. The order in which conditions were experienced by the participants was randomised.

The door choice figures were analysed by grouping data in two ways. Firstly, two groups of correct door choices were formed, one for the experimental condition and one for the control condition. This was labelled Set 1, see Results, section 4.5 below. For the experimental condition, the participants were scored for correct responses i.e. choosing to go through the door that masked an object. The control condition responses were scored for the same object positions as the experimental condition, even though there were no target objects. This ensured a record of participant choice, in the absence of any masking, which could be compared directly with the scores from the experimental condition.

Secondly, the door choice groups, as opposed to the *correct* door choice groupings of Set 1, were analysed. This was to ensure that any effects found were not simply due to factors other than the one that was being manipulated. After all, the control condition contained no target objects yet participants' choices were being labelled as either correct or incorrect. Therefore, door choice variance itself was measured, comparing the experimental against the control group. Given the fixed number of door choices available to participants, every time one side was chosen, the other was not. Therefore any differences found between the experimental and control group would be reproduced whether left door choice or right door choice were chosen to be analysed. The left door choices formed Set 2.

Two further, independent variables that were formed from the data were gathered via the response questionnaire, Appendix 2. Target-visibility and target-type scores from the experimental condition could be compared to ascertain figures regarding any explicit processing of objects. For object type, the reference sheet, Appendix 3, contained ten object types from which to choose. This number was thought sufficiently high to differentiate between random guessing and correct choices in this context.

4.3.4 Procedures

The order of experimental events for each participant was as follows:

1. Participants were briefed about their task in the experiment.

2. They were then introduced to the VE and shown the controls.

3. Participants were randomly assigned to either an experimental or control condition to carry out the first task. This was to navigate through the rooms of the VE, stopping in front of each set of doorways to indicate whether a masked object was visible and what that object was. Participants then passed through that doorway which they had designated as having an object within it and on to the next room to repeat the process.

4. A short break of a few minutes allowed participants to rest their eyes from the stereoscopic display.

5. Step 3 was repeated either with an experimental or control VE, complimentary to the participant's initial VE.

6. When completed, participants were debriefed.

4.4 Hypotheses

It was hypothesised that the dependant variable scores will show that:

Hypothesis 1a

The target objects will not be detected explicitly.

Hypothesis 1b

Where it is claimed that target objects are seen, they will not be identified correctly.

Hypothesis 1c

The position of target objects will be detected implicitly, as measured by door choice.

4.5 Results

4.5.1 Descriptive Statistics

Table 4.1 presents the descriptive statistics for the two analysis groupings used in the subsequent inferential analysis.

	Set 1		Set 2	
	Correct door choice total*	**Correct door choice means**	**Left door choice total***	**Left door choice means**
Experimental condition	173	7.9, sd.= 2.1	155	7.0, sd. = 1.4
Control condition	136	6.2, sd. = 1.7	142	6.5, sd. = 2.7

*out of a possible 308 (14 rooms x 22 participants)

Table 4.1. Descriptive statistics for both sets of analyses.

Further results were derived from the questionnaires regarding the identification of objects in the experimental condition, see table 4.2.

	correct identification	incorrect identification
Objects claimed seen	0	3
Objects claimed unseen	31	274

Table 4.2. Objects seen, not seen, identified and not identified, out of a possible 308.

4.5.2 Inferential Statistics

The data were analysed with SPSS using a paired-sample t test. As noted, the data were analysed using two different groupings: 'correct door choice' and 'left door choice'.

For Set 1 (correct door choice) $p = 0.007$ (two-tailed, significant). Cohen's d gives a figure of 0.2 for effect size that would be considered small by traditional standards.

For Set 2 (left door choice) p = 0.337 (two-tailed, non-significant).

4.6 Discussion

Following the development of the VEs for this experiment, it appeared to be possible to make backwards pattern masking work within a VE. This study therefore had to show more than just visual masking working. It was able to show a behavioural consequence of that masking, laying the groundwork for further study. The choice of a behavioural measure was deliberately taken to avoid criticisms aimed at experiments that used inventories as their sole or main metric (Slater & Garau 2007, Slater 2004), in particular post-task measures of presence. Questionnaires do provide important data of a type that cannot currently be obtained in any other way. Indeed, this thesis uses several. Nonetheless, issues of validity, consistency and reliability peculiar to questionnaires are raised and need to be borne in mind. A better approach can be to use more than one measure type, to cross-support evidence for underlying constructs. To this end, an observable metric in the form of a behavioural measure appealed, alongside a response questionnaire. Regarding the latter, of the two questions to be answered per room, was anything visible and what was the object type, answers to the second question were perhaps the more reliable. For the first question, participants could succumb to demand characteristics and give answers that they thought might be expected. The second question did not suffer from this potential drawback.

In the experimental condition, of those objects that the participants claimed to have seen, none were identified correctly and those objects that were identified correctly were all claimed to have been unseen. Unseen objects identified correctly totalled 31 out of a possible 308, equating to about one in ten correct identifications. This is what would have been expected by chance alone as participants had ten objects to choose from on the object reference sheet. Therefore, explicit guesses as to object type showed no effect, supporting Hypothesis 1b.

Mainly, though, participants indicated that they had not explicitly seen the targets in the experimental condition. Only three out of a possible 308 targets were claimed to have been seen, whether identified correctly or not. This result supports Hypothesis 1a.

The inferential results show a significant effect for Set 1, albeit a small one (assuming $\alpha = 0.05$). This supports Hypothesis 1c, indicating that participants showed an implicit positional awareness of the target objects. Door choice alone showed no significant effect between the two conditions, as shown by the Set 2 analysis.

4.7 Conclusion

From this study, we can conclude that the object position but not type was identified implicitly. This is consistent with the hypothesis that the modern, human-made object-types chosen would not be 'understood' by the particular neural processes responsible for detecting the position of a masked object.

We can also conclude more generally that visual masking can work within a VE, as participants were not able to perceive consciously the target objects. Additionally, it was shown that visual masking within a VE could produce measurable behavioural effects, despite a lack of explicit visual perceptibility. This gave a reasonable basis for continuing the work.

Several issues arose from this study. The effect size was small. This may have been due to a number of factors such as design elements or simply that the effect was intrinsically small. The latter of these is perhaps the most key for the aims of this thesis, as a small effect would presumably be of limited use in engineering a psychotherapeutic tool. Previous research had found that masked priming effects could be modulated (Winkielman & Berridge 2004). It might therefore be possible to find a modulating factor for visual masking within a VE to increase effects.

As for elements of the design, some areas could have been improved upon. For instance, participants might have been asked to walk through the door which they felt was *not* blocked by a masked target. This would have made intuitively more sense when navigating through the VEs. Another approach to improving any effect would be to investigate further the parameters of visual masking within a VE. Experiments 3 & 4 set out to do just this by looking at surface qualities and dynamics of the virtual objects used in masking.

The type of target has been shown to be important in prior masked priming research. Certain kinds of response have been elicited by using affectively valent facial expressions as priming targets, see section 2.3.5 of the Literature Review. The target objects chosen for the next experiment were of this form.

5 Experiment 2 – Choice Behaviour with Non-planar, Masked Facial Expressions as Targets

5.1 Abstract

It was investigated whether masked, non-planar facial expressions could have an effect on participant choice behaviour within a VE. The VE set-up was similar to Experiment 1, using the same basic virtual world. The participants were free to move through the world choosing which doorway to pass through to enter the next room. However, each doorway was masking a non-planar, static face with one of three expressions: smiling, angry-looking or neutral. Each pair of doors contained different expressions on either side. There was also a modulating factor introduced: a conflict or non-conflict condition created by varying instructions to participants. The results showed a significant difference between expressions but the modulating factor produced no significant interaction with facial expression nor a main effect.

5.2 Introduction

Previous research has shown a masked priming effect using facial expression targets. Some studies did not use a modulating factor (Dimberg *et al.* 2000). Others have found a modulating factor to have an effect (Winkielman & Berridge 2004), see section 2.3.5 above. It was felt therefore that a modulating factor offered the possibility of increasing the small effect size found in Experiment 1.

Much prior research has generally shown reactions to masked facial expressions (Killgore & Yurgelun-Todd 2004, Whalen *et al.* 1998, Ohman 2002, Wong & Root 2003, Jurena *et al.* 2010). The stimuli though, have tended to be planar images, often monochrome for control purposes, presented on various display screens. It would be hard to argue that such stimuli contained much in the way of immersive properties and certainly could not generate PI or Psi. As such, questions of validity and transferability arise: are researchers actually measuring responses to facial expressions or images of facial expressions? If the latter, then can such findings be applicable to the world outside of the laboratory? These questions are important because if we are to engineer a tool for therapeutic use, it must be able to affect change in the day-to-day life experience of recipients and not just produce effects in a controlled, laboratory context. Stimuli which can make an experimental experience closer in form to a real-life counterpart can support this aim. Non-planar facial expressions within a stereoscopic VE are a step towards such stimuli.

This experiment required participants to navigate a series of empty, identical virtual rooms. Upon entering each room, they saw two doors at the opposite end. Participants had to choose which of the two doors to navigate through to enter the next room. Each planar mesh door acted as a mask for a non-planar facial mesh. Each facial mesh supported either a smiling, angry-looking or neutral expression. The objects all shared the same surface material, that of the masking door. The objects each sat on the same depth plane to each other, their foremost polygons being immediately behind the plane of the masking doors from the initial participant point of view.

Another factor to take into account was the number of times a participant was potentially exposed to a particular target. This was dictated by the length of each room and the navigation speed, assuming a roughly linear trajectory from entry to exit of choice. We assumed each target must be 'seen' at least once to facilitate a choice based on facial expression, although in reality a single exposure would probably have very little effect on the participants. It was believed that any affect-driven effects would be the result of cumulative exposures to target stimuli. This meant that for each pair of doorways perceived the mask/target pair had to cycle synchronously with its neighbour. Any asynchrony could potentially introduce a confound in that it increases the likelihood that one of the expression will be experienced more than the other. In addition, capping participant speed ensured several exposures per room. The issue of total target exposure time was revisited in Experiment 4.

5.3 Method

This experiment used a 3 (expression: angry-looking vs. smiling vs. neural) x 2 (threat: conflict vs. non-conflict) within-participants design, with a binary value preference measure (left or right door) to test for between-treatment effects. After a practice period, participants completed 48 trials in total with 24 repetitions per threat level and 32 repetitions per expression level.

5.3.1 Participants

Twenty-four people were recruited for the experiment, drawn from the UCL psychology subject pool, posters and word of mouth. Three participant data sets were discarded for non-compliance leaving twenty-one data sets for the final analysis representing seven right-handed males, one left-handed male, with the remainder being all right-handed females.

5.3.2 Materials

The hardware and software platforms used were those described in the Methodology set out in Chapter 3.

The same basic VE used in Experiment 1 was also used for this study, see figure 5.1 below. The only change being that the non-planar masked targets were facial expressions as opposed to domestic items. Each pair of doors masked a pair of expressions, always different to each other. Various expression pair combinations, see table 5.1, were randomly distributed throughout the VEs.

Figure 5.1. A view of the VE and two example expression targets: neutral and angry.

The target facial meshes were non-planar objects displaying one of three expressions: happy, angry-looking or neutral. Prior to the experiment, it had been established that the facial expressions could be understood appropriately, at least consciously. Each of the target objects had an approximately equal internal depth disparity in the axis perpendicular to the door plane. The face texture material, shared with the planar mask doors, had no obvious linear depth cues. This helped to dampen any monocular masking inhibition from linear perspective depth cues and so strengthened the overall masking effect. The mask texturing should have some structural similarities to a face and so help increase any masking effect (Loffler *et al.* 2005).

As the participants navigated through each room, they moved through whichever doorway they preferred, and on into the next room. This was repeated through the series of target pair combinations in a series of rooms. The conflict condition (see Measures, section 5.3.3 below) consisted of twenty-four rooms, as did the non-conflict condition. Door choices were automatically recorded by the software. This allowed participants to remain focused upon the screen, which was possibly an issue for Experiment 1.

5.3.3 Measures

There were two independent factors in this experiment: Expression and Threat. Participants each experienced all levels of both factors, making this study a within-participants design. Each participant experienced the conflict and non-conflict VEs in a random order.

The Expression factor had three levels: ‘smiling’, ‘angry’ and ‘neutral’.

The Threat factor had two levels: a ‘conflict’ condition and a ‘non-conflict’ condition (see below).

The Threat factor was intended to be a modulating factor, possibly interacting with the Expression variable.

The dependant metric was derived from the number of times a participant went through a doorway containing a particular masked expression. These scores could then aggregated per participant per treatment and the means found for example, see table 5.3, or collapsed across conditions and means derived as necessary, see table 5.2 for example. These scores also formed the basis of the inferential analyses.

Each facial expression was displayed an equal number of times on each side within each VE, although in a randomised order to control for the potential introduction of confounding biases.

The Threat factor was embedded within the instructions given to participants. The conflict condition wording attempted to induce a need to avoid a possible conflict. This was done by asking participants to imagine that they had a low health score, similar to a video game. They needed to avoid any encounters within the environment (there were none) and exit as quickly as possible. The non-conflict instructions were that they merely had to move as briskly as possible through the environment to the end of the task. The instructions for each of these modulation conditions were matched as closely as possible in length and complexity.

It was intended that both sets of instructions induce a sense of urgency, the only difference being that one contained the implication of threatening encounters. The goal was for the participant to choose quickly to lessen the possibility of cognitive, 'top down' interference with the choice and so act on the simplest impulse i.e. to avoid any conflict implied by the angry face. It was thought that this behaviour would be heightened by the conflict condition, which would also perhaps dampen the differences between the happy and neutral choices.

An on-screen graphic of a participant's fictitious 'health score', was not used as it was thought that this could potentially introduce uncontrolled variance. Even if a similar graphic were used for each condition, a confound could arise from the visual distraction, depending on whether the participant was a regular video game player for example. A verbal narrative was seen as a good alternative.

A sound track was also considered, to heighten the differences between the Threat conditions, but again it was felt that this may introduce a further confound as individuals may respond differently to the same soundtrack. An alternative view was taken for Experiment 5 where soundtracks were used to contribute to the affective valence of conditions.

pair	**Left-hand side**	**Right-hand side**
1	*Happy*	*Neutral*
2	*Happy*	*Angry*
3	*Neutral*	*Angry*
4	*Neutral*	*Happy*
5	*Angry*	*Happy*
6	*Angry*	*Neutral*

Table 5.1. Matrix of expression pair combinations.

Each Threat condition VE, either 'conflict' or 'non-conflict' consisted of twenty-four rooms with each room containing one pair of doors.

Each masked expression was present sixteen times per condition, eight in left-hand doorways and eight in right-hand doorways. These expressions were matched against equal numbers of the other two expressions.

Each expression pair, see table 5.1, was randomly distributed four times throughout the VE of a given Threat condition.

Within a given Threat condition, an Expression score for a particular participant denoted the amount of times a particular expression was implicitly ‘chosen’ by moving through a doorway containing that expression. These choice scores of individual participants were therefore aggregates per Expression per Threat condition. The number of potential points a single expression could amass within each condition, for each participant, was sixteen.

5.3.4 Procedures

The order of experimental events for each participant was as follows:

1. Participants were briefed about their task in the experiment.

2. They were then introduced to the VE and given a short practice to become accustomed to the controls.

3. Participants were randomly assigned to either a conflict or a non-conflict condition and given further instructions particular to that condition.

4. Their task was to navigate through the rooms of the VE, passing through one of the two doorways at the end of each room. This continued until an onscreen indicator appeared, marking the end of that trial.

5. A short break of a few minutes allowed participants to rest their eyes from the stereoscopic display.

6. Steps 3 & 4 were repeated except that the VE used now used was that of the other Threat condition VE.

7. When completed, participants were debriefed.

5.4 Hypotheses

It was hypothesised that the dependant variable scores will show that:

Hypothesis 2a

There will be a significant interaction between the Threat variable and the Expression variable.

Hypothesis 2b

If no interaction is present then there will be a significant main effect of the Expression variable.

5.5 Results

5.5.1 Descriptive Statistics

Factor	Level	Mean
Expression	smiling	8.5, sd. = 2.0
Expression	angry	7.2, sd. = 1.5
Expression	neutral	8.4, sd. = 1.9
Conflict	conflict	8.0, sd. = 1.6
Conflict	non-conflict	8.0, sd. = 2.2

Table 5.2. Collapsed factor level means.

The above factors and levels resolve into the following treatment groups:

Treatment Means	**smiling**	**angry**	**neutral**
conflict	8.2, sd. = 1.8	7.5, sd. = 1.4	8.3, sd. = 1.5
non-conflict	8.7, sd. = 2.1	6.9, sd. = 1.7	8.4, sd. = 2.3

Table 5.3. Treatment means: the rows denote the levels of the Conflict factor, the columns denote levels of the Expression factor.

5.5.2 Inferential Statistics

Repeated measures ANOVA was used to determine whether there was a significant interaction between the Conflict factor and the Expression factor as well as highlight any main effects. There was no significant interaction between the Expression and Conflict factors. There was a significant main effect for Expression $p<0.05$, partial eta squared = 0.18.

Greenhouse-Geisser adjustmented figures were used, due to the possibility of repeated measures designs violating the sphericity assumption. However, as there was no noteworthy difference between the adjusted and non-adjusted figures, it is assumed that sphericity was not violated in any significant way.

Table 5.4, below, shows the significance of pair-wise treatment comparisons grouped according to the Threat factor levels.

treatment comparisons			**significance**
non-conflict/smile	vs.	non-conflict/angry	$p < 0.05$, sig.
non-conflict/smile	vs.	non-conflict/neutral	$p > 0.05$
non-conflict/angry	vs.	non-conflict/neutral	$p < 0.05$, sig.
conflict/smile	vs.	conflict/angry	$p > 0.05$
conflict/smile	vs.	conflict/neutral	$p > 0.05$
conflict/angry	vs.	conflict/neutral	$p > 0.05$

Table 5.4. Significances of pair-wise comparisons of treatments

None of the same-expression, pair-wise treatments for the two levels of Threat (i.e. conflict/smile vs. non-conflict/smile etc.) were significant. However, we can see from table 5.4 that within the non-conflict verses non-conflict treatments, angry faces produced statistically significant results to the other two expressions.

5.6 Discussion

Hypothesis 2a was not supported. The modulating factor, conflict threat, failed to produce an overall effect on facial expression choice scores. It is possible that the chosen method, instructional difference, was simply too ‘safe’ or controlled to produce any effect expression choice.

Hypothesis 2b was supported. Partial Eta Squared (a conservative alternative to Cohen's d) for Expression was 0.18, whether corrected for sphericity or not. This constitutes a small to medium effect by traditional standards and is similar to that found for Experiment 1. Masked facial expression had an effect on participant choice behaviour even if those choices were not modulated by task instructions.

Pair-wise comparisons can be useful, although must be treated with caution as Type 1 errors can arise from increased family-wise error rate. The non-conflict level pair-wise comparisons show the Expression level 'angry' to be significantly different to the other two expressions, see table 5.4. However, 'smiling' and 'neutral' showed no significant difference in the non-conflict condition. This suggests that participants were treating the angry face differently to the other two expressions. In addition, neutral expressions were not treated significantly differently to smiling faces in this particular VE.

None of the conflict level comparisons were significant. This may suggest that expressions were treated in the same way when participants were apprehensive of a conflict but caution is indicated: a smile may have been treated as malicious, and therefore negative, but processed differently from an angry expression even if producing the same behavioural outcome on the measure used in this experiment.

In terms of the collapsed score means, see table 5.3, it was assumed that a higher score for one expression constituted a preferential bias towards that expression and a bias away from the other two. When scores were collapsed by Conflict, the Expression level 'smiling' gave the overall highest mean, closely followed by 'neutral' with ‘angry’ the lowest. The gap between smiling and neutral means was 0.1 whereas the gap between neutral and angry means was 1.2. This concurs with the pair-wise comparisons in the non-conflict conditions at least where angry expressions appeared to be treated differently.

5.7 Conclusion

We can now see that non-planar, masked facial expressions within a VE can affect participant behaviour. This suggests that implicit processes were being driven by the affective valence of the masked objects. This concurs with prior work that suggests affective facial expressions to be particularly significant forms of stimuli for human beings (Viggiano & Marzi 2010).

The lack of modulation from the Threat factor may have been down to its form. Although all participants indicated that they did at least understand what the instructions meant, the ‘health score’ idea might have been particularly weak. Its implications may have been understood and reacted to more clearly by those used to such ideas, for example people who regularly spend large amounts of time playing 3D computer games, perhaps of a kind that simulate life threatening conflict. The majority of participants could not be said to fall into this category.

The remaining experiments in this thesis all use masked expressions in one way or another. The next two experiments will further test the parameters of visual masking within a VE with the aim of initiating the search for the most effective configurations and other modulation.

6 Experiment 3 – A Study of Visual Masking Parameters within a Virtual Environment: Target Salience, Target Surface Saturation & Target Surface Fluency

6.1 Abstract

The third experiment extended the work of Experiment 1 and 2 by looking at the effects of altering various properties of visually masked, non-planar objects within a VE. Two aspects of surface quality were investigated: texture image fluency and texture colour saturation. In addition, the saliency of the masked objects was introduced as a variable. The aim was to look for disruptions to the masking effect and to see if any interactions were present. On encountering flashing doorways, which sometimes acted as masks for various target objects, participants were required to state if they could perceive any objects and, if so, indicate their type (from a reference sheet provided). Targets that may or may not have been in the doorways included various affect neutral domestic objects and highly salient facial expressions. This experiment found a significant interaction between surface colour and texture fluency. Saliency did not interact with either of the other two variables nor produce a main effect.

6.2 Introduction

Experiment 3 used static, non-planar faces and household objects as targets, investigating the interaction between the three parameters target colour, target texture fluency and target saliency on visual masking within a VE. The question asked was, would changing these parameters render targets more or less susceptible to conscious perception in a VE?

Here, fluency refers to the ease with which an image or text can be understood by an observer (Reber & Schwarz 1999) and can be measured by the speed of cognitive processing (Reber *et al.* 1998). High perceptual fluency, for instance, can be associated with visually distinct imagery (Jacoby & Whitehouse 1989). The sharpness of a texturing image on a virtual surface is one example of a fluency factor. This was used for Experiment 3.

Masked priming studies have often used static facial expressions as their target stimuli. Facial expressions can be highly salient stimuli (Viggiano & Marzi 2010). Although a human face appears to be quite complex in comparison to the figures used in many masking experiments, there are neural systems that rapidly process such images (Humphreys *et al.* 1993). It is a reasonable argument to make, therefore, that facial expressions could be considered as high salience stimuli to the population of human beings when compared to other objects such as domestic items for which we would have no general, evolved saliency response.

It would appear, therefore, that fluency and salience might have an effect on visual masking processes and might interact. The purpose of Experiment 3 was to test this.

Saturation was added as a third test factor as this has been shown to affect implicit processes (Breitmeyer *et al.* 2004a, Schmidt 2000). All three factors have the potential to interact with each other as they are presumably capable of affecting related visual processing streams.

6.3 Method

This experiment used a 2 (saliency: high-salience vs. low-salience) x 2 (saturation: high-colour vs. no-colour) x 2 (texture fluency: high-fluency vs. low-fluency) within-participants design, with a verbal indication measure (absence/presence of a target and target identity) to test for between-treatment effects. After a practice period, participants completed 24 trials in total with 24 repetitions per saliency level, 24 per saturation level and 24 per texture fluency level.

6.3.1 Participants

Twenty people were recruited from a UCL participant database of students and non-student adults. There were nine females, two left handed and 11 males, two left handed. All complied with the experimental protocol.

6.3.2 Materials

The VEs consisted of a series of courtyard-like rooms terminating in a pair of doors, each door leading to a similar room, see figure 6.1. Behind each door was a non-planar, visually masked target, either a domestic object or facial expression. Each door acted as a planar mask for the target object in its doorway. Each target was shown for approximately 16ms, followed by the masking door for 250ms and finally an inter-stimulus gap of 250ms. Cycled repeatedly, only the masking doors were explicitly perceptible, appearing to blink. This shortened cycle was intended to increase participants' exposure to the targets, approximately doubling that encountered in the previous work.

Two different VEs were used: high saliency and low saliency. These labels corresponded to the type of targets contained within the particular VE. Six, low salience, domestic objects were used as targets: a chair, a teapot, a watering can, a cup a lamp and a pan. Six, high salience, facial expressions were used as targets: smiling, angry, sad, surprised, disgusted and afraid.

Combinations of texture were applied to the targets, masks and the wall immediately behind the target and mask, termed the ground. There were two texture factors manipulated: fluency and saturation (see Measures, below). In each case, the target and mask had the same surface, always in contrast to the ground, i.e. if the target and mask had a high-colour, fluent texture, the ground would have no-colour, low fluency texture. The combinations were randomised throughout each VE.

Two reference sheets were provided: one showing the six facial expressions used in the high salience conditions (see Measures, section 6.3.3 below), the other showing the six domestic objects for the low salience conditions. The reference sheets depicted the objects or facial expressions in neutral grey tones, arranged on a simple grid structure of two items across, three items down. The items on the reference sheets corresponded to the masked targets used throughout the experiment. The difference between the reference sheet depictions and the actual targets used in the VE was that the targets did not have neutral grey-tone surfaces but one of the two saturation levels and one of the two fluency levels, see figures 6.2 & 6.3. Appendices 4 & 5 show the object and expression sheets respectively.

Figure 6.1. Image of initial, participant view of the basic VE showing high fluency, high colour doors.

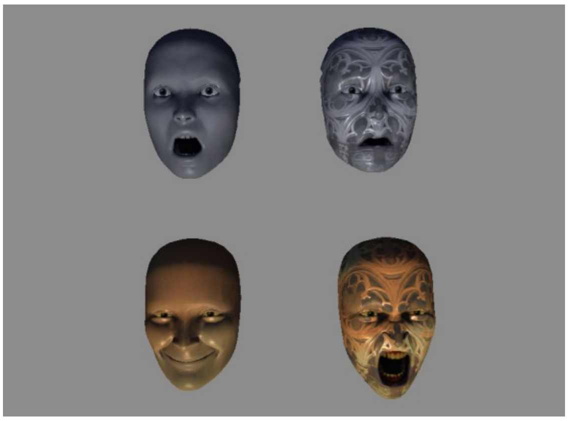


Figure 6.2. Examples of high salience targets used showing no colour (top), high colour (bottom), low fluency (left) and high fluency (right).

Figure 6.3. Examples of low salience targets used showing no colour (top), high colour (bottom), low fluency (left) and high fluency (right).

6.3.3 Measures

This experiment obtained quantitative data from participant responses within each room of the virtual world. A verbal response was chosen over a written one because the latter would involve prolonged attention away from the screen, therefore disrupting participants' engagement with the stimuli.

There were three independent variables, each with two levels. For examples, see figures 6.2 and 6.3. The variables were:

Target saliency: 'high-salience' used expressive faces;
'low-salience' used domestic objects.

Saturation: 'high-colour' used full colour textures;
'no-colour' used greyscale textures.

Texture fluency: 'high-fluency' level used a sharp, legible texture image of detailed stone tracery;
'low-fluency' used a blurred version of the same.

The experiment was divided into two trial-blocks per participant, one using high salience targets, the other low salience. Each trial-block contained twelve rooms, with two doors per room. Trial-block order was randomised for each participant. Further randomisation was introduced by having two versions each of the high and low salience VEs as well as randomising the order that each domestic object or facial expression was shown in.

Each participant was instructed to indicate whether anything could be seen within each doorway, apart from the blinking doors. If they could see the target, they would then indicate the nearest image to it on the appropriate reference sheet. Each correct answer scored one point for that particular saliency/fluency/saturation combination. Each incorrect answer scored zero. If nothing could be seen, the score was entered as zero. Unknown to the participants, every doorway contained a target. These final point scores for each of the saliency/fluency/saturation combinations constituted the dependent variable to be analysed.

6.3.4 Procedures

The order of experimental events for each participant was as follows:

1. Participants were briefed about their task in the experiment.

2. They were then introduced to the VE and given a short practice to become accustomed to the controls.

3. Participants were then assigned to either a high saliency or low saliency VE to carry out the first trial-block. This was to navigate through 12 rooms of the VE, stopping in front of each set of doorways and indicate what, if anything, was visible.

4. A short break of a few minutes allowed participants to rest their eyes from the stereoscopic display.

5. Step 3 was repeated, with a VE of differing saliency level trial-block.

6. When completed, participants were debriefed.

6.4 Hypotheses

It was hypothesised that the dependant variable scores will show that:

Hypothesis 3a

Target salience will interact significantly with saturation.

Hypothesis 3b

Target salience will interact significantly with fluency.

Hypothesis 3c

Saturation will interact significantly with fluency.

Hypothesis 3d

All factors will interact significantly with one another.

6.5 Results

6.5.1 Descriptive statistics

Table 1 shows the score means for all treatments. The means were the accumulated scores per treatment divided by the number of participants.

Salience	Fluency	Saturation	Mean score
low	low	no colour	0.35, sd. = 0.75
low	low	high colour	1.75, sd. = 1.16
low	high	no colour	0.15, sd. = 0.37
low	high	high colour	0.45, sd. = 0.51
high	low	no colour	0.40, sd. = 0.60
High	low	high colour	1.35, sd. = 0.81
High	high	no colour	0.10. sd. = 0.31
High	high	high colour	0.30, sd. = 0.47

Table 6.1 Treatment Means.

6.5.2 Inferential statistics and hypotheses evaluation

Repeated measures ANOVA was used for the analysis and the results showed an interaction: fluency vs. saturation $p = 0.01$, partial eta squared was 0.42.

Greenhouse-Geisser adjustments were applied due to the tendency of repeated-measures designs to violate the sphericity assumption. This did not affect the significance of any results.

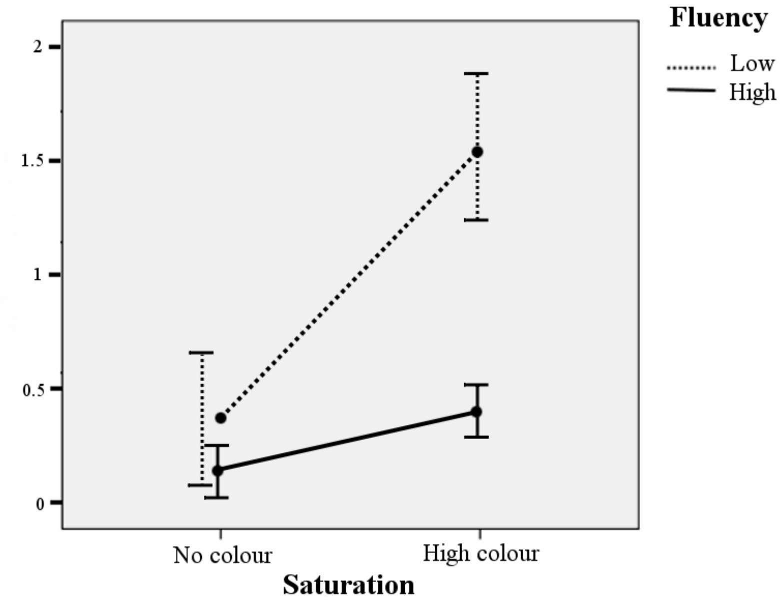


Figure 6.4. Plot of the Saturation vs. Fluency interaction. The ordinate depicts the estimated treatment means: the higher the score, the more likely that the masked object is seen.

The plot in figure 6.4 depicts the estimated treatment means of Saturation versus Fluency, controlling for Salience.

6.6 Discussion

Only Hypothesis 3c was supported, Hypotheses 3a, 3b and 3d were not supported.

The treatment mean scores were out of a possible six (eight treatments across 48 doors) and show that, generally, masking was working across all treatments. However, there was some variance: the ANOVA results show that Saturation and Fluency interacted, which figure 6.4 depicts visually. When comparing low fluency to high fluency scores the largest difference was shown by the high-colour surfaces. Similarly, when comparing scores from greyscale with high-colour, the greatest range was found within low fluency.

Although attempting to look at effects in isolation from an interaction will not result in any definite conclusions, from the interaction plot of figure 6.4, there may be some conclusions to be drawn. High fluency gives the least difference between means: the mean score difference between high colour/high fluency and low colour/high fluency is less than the difference between high colour/low fluency and low colour/low fluency. From this we could argue that the surface image fluency of virtual objects is perhaps the most important factor of the three tested when constructing a masking scenario within a VE: whether using a colour or greyscale texture, try to use a high fluency design, matching target and mask.

Saliency did not produce a significant interaction with the other parameters and neither were any main effects observed. This may suggest that expression processing was being treated implicitly in a similar way to non-expression processing by participants although this is not necessarily the case. Previous research that has shown conscious perception does not always appear to be necessary for facial expression processing to occur (Dimberg *et al.* 2000, Winkielman & Berridge 2004). Therefore, in terms of the current experiment, domestic objects and facial expressions were not necessarily processed in the same way by the participants, merely that they were not treated differently in terms of becoming explicit with the form of visual masking used.

6.7 Conclusion

The first experiment, see Chapter 4, had shown that visual masking was possible within a VE and that masked virtual objects could be non-planar. It has been shown that targets can contain a limited, internal depth disparity, necessitating an external depth disparity between target and planar mask at certain points in the geometry. However, that work did not explore fully other parameters beyond depth.

This experiment begins that process by examining how surface qualities can affect visual masking within a VE. Potentially there is a large range of forms that masked, virtual objects can take; work that modulates the parameters of those forms can further our understanding of how these effects operate and how they can be applied.

Saturation and textural legibility appear to be factors involved in the shift from implicit to explicit perception within a VE. High fluency surface textures are recommended for visual masking within a VE, the same texture being used for both target and mask.

Further work could be used to widen the investigation of colour effects by increasing treatments to include varieties of hue balance and saturation. Similarly, fluency has several sub-properties that could be examined such as image content or element frequency.

7 Experiment 4 – A Further Study of Visual Masking Parameters within a Virtual Environment: Target Dynamism

7.1 Abstract

The fourth experiment looked at whether it was possible to form masked targets from dynamic (i.e. changing) objects. Facial expressions were used as non-planar targets that changed from neutral to a full expression over several cycles of target and mask presentation. Each time a target was flashed onto the screen it would change slightly from before, similar to a frame of animation. The intention was to see if these dynamic targets were read, implicitly, as animated expressions or as a series of still images. In the VE, each dynamic, masked expression was juxtaposed with an affect neutral object to which the participants had to give an attractiveness rating. A comparison control used static masked expressions. Results showed that the dynamism and the expression variables interacted but that the static scores produced the highest variance, indicating modulation by an exposure effect. This contra-indicates that the dynamic expressions were being read as animated faces.

7.2 Introduction

To date, no VE research appears to have been carried out using targets consisting of dynamic, non-planar objects. Experiment 4 used non-planar facial expressions as targets within a VE, see figures 7.3 & 7.4. Some of the targets were dynamic i.e. on each exposure the target form changed incrementally from a neutral facial expression to a full affect display, analogous to the frames of an animation, see figure 7.1. Other targets were static i.e. they always displayed the same, full expression. 'static' and 'dynamic' formed the two levels of an independent factor: Dynamism.

The term 'dynamic' was used, not 'animated'. This is because it was not known whether the brain would read such targets as animated expressions or a series of separate expressions. If the former then it was hypothesised that they should produce a greater effect than static targets (Detenber *et al.* 1998, Ambadar *et al.* 2005, Humphreys *et al.* 1993). If read as a series of still expressions then an exposure effect could take precedence: static targets would produce greater effects than dynamic ones because the participant would be exposed to a full expression more times on approach.

Another consideration would be the form of the dynamic expression itself. If we choose a smile, for example, do we have to use the whole of a smile? Given that facial expressions have a temporal component, could we use part of the 'timeline' of an expression for greater effect? Anecdotally, the author can attest that the first half of a smile animation or footage gives a different subjective reaction from the second: the onset-to-peak section of a smile leaving the impression of a more positive experience than the peak-to-offset. Given that the overall intention of this thesis is to look towards ways of producing a positive change in state affect, it was decided it would be advantageous to consider this. Therefore, the dynamic targets cycle from neutral to a full expression only, see figure 7.1.

Figure 7.1. Dynamic expression target increments.

A further consideration is that a convincing Duchene smile may be difficult to reproduce on a facial mesh. Dynamics may be important but so is context. The same smile may be interpreted differently in different situations or according to the mood a person is in etc. However, it may be the case that a dynamic target, processed non-consciously, may circumvent this. The critical appraisal of a facial expression takes place across many parts of the brain (Humphreys *et al.* 1993), some of which are linked with conscious awareness. Rapid, non-conscious, automatic processing of expressions does take place but it might not be as critical as slower, more complex 'theory of mind' judgements (Saxe & Baron-Cohen 2006) which are, again, linked to conscious awareness.

A second factor, Expression, had two levels: ‘angry-looking’ and ‘smiling’. It was known from the earlier work that participants respond differently to masked emotional expressions than to neutral ones in a virtual world.

It is known that value judgements of a neutral stimulus can be affected by its spatial juxtaposition to other, significant stimuli (De Houwer 2007). This appears to be the case even if the significant stimulus is implicitly perceived (Dijksterhuis 2004, Winkielman & Berridge 2004). It was decided to test whether or not the juxtaposition of masked faces with visible, virtual objects had any effect on a participant’s liking for that object and whether expression dynamics could modulate this. The targets were juxtaposed with objects that were then rated for attractiveness, the premise being that implicit perception of the target would affect that score.

7.3 Method

This experiment used a 2 (dynamism: dynamic vs. static) x 2 (expression: angry-looking vs. smiling) within-participants design, with a verbal indication measure (attractiveness scores for visible, sculpture-objects) to test for between-treatment effects. After a practice period, participants completed 32 trials in total with 16 repetitions per dynamism level and 16 per expression level.

7.3.1 Participants

Twenty-two people were recruited from the UCL participant database of students and non-student adults. Two sets of data were discarded prior to analysis due to non-compliance, leaving 14 female datasets, all right-handed, and six male datasets of which five were right-handed.

7.3.2 Materials

The hardware and software platforms used were those described in the Methodology set out in Chapter 3.

The VE consisted of a series of open-roofed rooms terminating in a pair of doors, each door leading to a similar room. Within each room, a static, non-masked, virtual sculpture was rendered in front of one of the doorways, chosen at random. Each sculpture resembled an enlarged, domestic object. There were eight object types in total: a cup, a chair, a lamp, a tap (faucet), a telephone, a table, a teapot and a pan, see figure 7.2. All were rendered in neutral grey tones. These were randomly distributed throughout the series of 32 rooms, each appearing four times. The door behind each sculpture was not a mask and was always static and impassable.

Figure 7.2. Sculpture objects used in Experiment 4

Within the each adjacent doorway, a target face was rendered, the blinking door acting as a mask. The faces displayed either an angry expression or a smile, each face being rendered as either a static or dynamic. Dynamic targets changed form with each target/mask cycle, each beginning as a neutral expression. Every time it was re-rendered, the expression changed incrementally towards either fully angry or fully smiling, resembling frames of an animation, see figure 7.1. Each expression took seven increments to change from neutral to full. Each increment lasted for approximately 16ms. The following mask door appeared for 250ms, with an inter-stimulus gap of 250ms. The VE was constructed so that at the maximum speed of travel the participants would experience at least fourteen target/mask render cycles (increments) before entering one of the two doorways. The dynamic targets would therefore reach full expression at least twice per room visit.

A comparison condition used static targets comprising of the two, full facial expressions, the two faces shown on the right of figure 7.1. Therefore, each static target exposed either a smiling or an angry face at least 14 times per room visit.

Figure 7.3. Image of VE with object and smiling face target: the face is visually masked whereas the chair sculpture is visible in every frame.

Figure 7.4. Image of VE with object and angry face target. Again, the sculpture was displayed permanently whereas the face was masked.

7.3.3 Measures

Two independent variables were present, each with two levels:

Dynamism: ‘dynamic’ used multiple frame targets, see figure 7.1;

‘static’ used single frame, full expression targets.

Expression: ‘smiling’ used a smiling expression;

‘angry’ used an angry expression.

This arrangement gave 4 treatment groups: dynamic smiling, dynamic angry, static smiling and static angry.

Dependant data was formed from scores given verbally, by the participants. Each participant was asked to rate how attractive the sculptures were. Participants were asked to give a score ranging from 1 to 999, with no zeros. The lower range equated to 'very unattractive', the upper to 'very attractive' with a score of around 499 being neutral. Participants indicated they understood the instructions. This form of measure was chosen to disrupt a participant's memory of a previous score for the same object. This was to avoid the introduction of a confound wherein participants felt they had to produce consistent scoring per object. This increased the likelihood that scores were independent from each other and reflected current attractiveness of the object.

The juxtaposition of an implicit, valent target with the consciously perceptible sculptures was intended to affect the attractiveness score for the sculptures.

This within-participants experiment was divided into two trials-blocks per participant, to allow a break to prevent eyestrain. Each trial-block contained sixteen rooms with two doors per room. Each sculpture was shown randomly four times throughout the 32 rooms, adjacent to one of the four target groups.

7.3.4 Procedures

The order of experimental events for each participant was as follows:

1. Participants were briefed about their task in the experiment.

2. They were then introduced the VE and given a short practice to become accustomed to the controls.

3. Participants then carried out the first experimental task. This was to navigate through 16 rooms of the VE. After passing through each open doorway, they had been instructed to stop and verbally deliver an attractiveness score for the particular sculpture that they had just passed. They then moved on into the next room and repeat the process.

4. A short break of a few minutes allowed participants to rest their eyes from the stereoscopic display.

5. The step 3 task was repeated, within a VE containing different, randomised combinations of targets and sculptures and levels of dynamism.

6. When completed, participants were debriefed.

7.4 Hypotheses

It was hypothesised that the dependant variable scores will show that:

Hypothesis 4a

An interaction will occur between the Expression and Dynamism scores.

Hypothesis 4b

The difference between the two dynamic group scores means will be greater than difference between the two static group score means.

Hypothesis 4c

When scores are aggregated into a smiling group and an angry group, the mean of the smiling scores will be higher than the mean off the angry scores.

7.5 Results

7.5.1 Descriptive statistics

The participant score means were formed by adding the total scores per factor combination and dividing by the number of participants: table 7.1 shows these for the four treatment groups:

Dynamism	Expression	Score Mean
dynamic	smiling	437.0, sd. = 219.2
dynamic	angry	453.9, sd. = 221.4
static	smiling	489.9, sd. = 224.8
static	angry	437.5, sd. = 231.3

Table 7.1. Treatment Means

From table 7.1 we can see that the difference between the two scores means of the Dynamism levels was 16.9 and the difference between the two Static levels was 52.4. Overall, the smiling mean was 463.5 and the overall angry mean was 445.7.

7.5.2 Inferential statistics and hypotheses evaluation

The series of scores per factor combination were analysed using a repeated measures ANOVA model. An interaction between the Dynamism and Expression was found, $p = 0.02$, partial eta squared 0.03. This supports hypothesis 4a. Figure 7.5 plots this interaction, which indicates a greater response for the static factors against Expression and a larger score for the smiling expressions when Dynamism is taken into account.

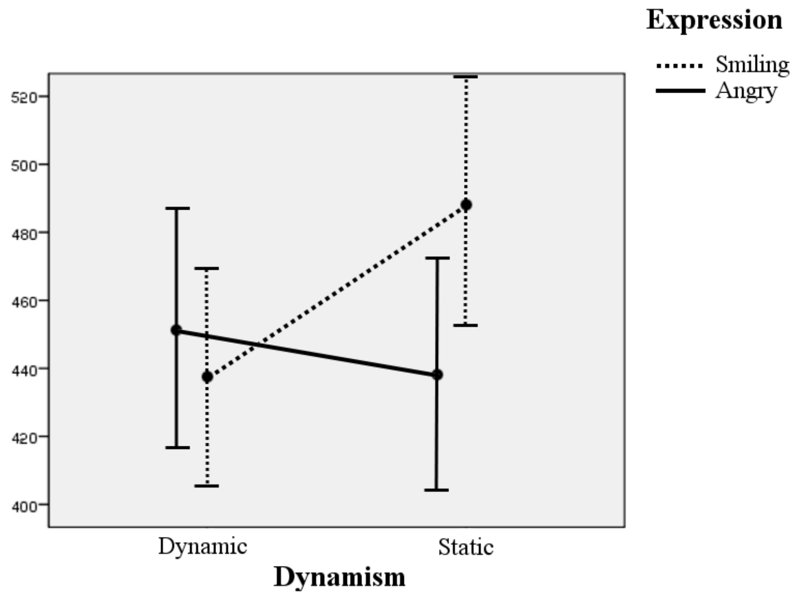


Figure 7.5. Plot of the Dynamism and Expression interaction. The ordinate depicts the dependant variable means.

Hypothesis 4b was not supported: the difference between the two dynamic group scores means was *less* than difference between the two static group score means. This indicates an exposure effect because the static faces appear to have had a greater effect than the dynamic. However, caution must be observed because of the significant interaction. Dynamism *could* be creating an exposure effect but we cannot clearly separate this out from Expression valence.

Hypothesis 4c was supported, as the overall smiling score mean was higher than the angry score mean.

As this was a within-participants design Greenhouse-Geisser sphericity adjustments were tested. These did not affect the significance of any results.

7.6 Discussion

The results from this experiment show that there was a modulation effect but it was not internal to the dynamic condition. It arose due to the presence or absence of the dynamic form of the target itself i.e. static targets produce the stronger effects. This could indicate that some form of exposure effect was over-riding any dynamic target effects: as noted, in the static conditions the participants would have been exposed to a facial expression target at least 14 times if they were moving at maximum speed. If they proceeded more slowly, then this exposure figure would increase. The dynamic conditions would, at maximum speed, expose participants to a full expression at least twice. This, again, would increase if the participant moved more slowly. However, as this was a within-participants design, with static and dynamic conditions randomly distributed throughout the VE, any individual participant's speed of movement drop below the maximum possible should be distributed approximately equally across all conditions. Therefore, it can be argued that an increase in exposure time to the static conditions can be assumed to approximate the increase in exposure time to the dynamic conditions. This means that, even taking into account speed variance between participants and conditions, it is likely that the static conditions exposed the participants up to seven times the amount of full expression targets as the dynamic. It could be argued that the sixth or even fifth frame of a dynamic sequence approximated the full expression. However, this still does not give the dynamic conditions anything like the exposure of the static.

Dynamic effects would be expected to arise if the brain were reading the changing sequence of target frames as animation, in the dynamic conditions. We would have expected to obtain a greater affective impact and, it follows, the greatest range of scores in the dynamic condition: the smiling dynamic combination would be expected to receive the highest overall scores with the angry dynamic condition receiving the lowest. This is not what was found. In fact, the opposite was actually the case indicating, as stated, some form of exposure effect.

The support for hypothesis 4c in the results shows that objects juxtaposed with smiling faces scored slightly more on average than those placed with angry faces. This is what we would expect due to evaluative conditioning. Although the difference between average scores was only 17.8, well within the standard deviation for each group, aggregating scores in this way effectively flattens out modulation from the Dynamism factor. Therefore, caution must be observed because of the significant interaction: we cannot unequivocally conclude that smiling faces score higher on average than angry although the evidence suggests that they do. This concurs with other work; Experiment 2 for instance.

7.7 Conclusion

As noted, the results show that there was a modulation effect due to levels of the Dynamic factor. Although this could indicate that some form of exposure effect was over-riding any dynamic effects, it may also be that there were no dynamic effects at all. In this case, we suspect that the brain is not experiencing the dynamic target as one, changing object but as a series of separate objects. Further work would be needed to confirm this, perhaps using fMRI to look at areas associated with the processing of facial or biological motion. If a lack of activity in these brain areas were shown then this would almost certainly be down to the constraints imposed by the masking model. In our paradigm, each 'frame' of the masked expression was presented approximately every half a second. Animation running at this speed does not show smooth motion effectively. The masking format is further hampered by there being a gap between frames: the target image does not, of necessity, stay on screen during the masking fames and the inter-stimulus gap. This could be creating too long a gap between target frames to be read as animated movement.

Overall, this study supports the idea that either the number of target exposures is important or accumulative exposure time is. Therefore, other factors, such as participant eye saccades, might also need to be taken into account, suggesting some form of eye tracking would be needed in future work. For the purposes of further experiments in this thesis, static expressions appear to be the best option for initiating affect-driven responses.

8 Experiment 5 – Does the Form of the Environment Matter for Affective Responses to Masked, Virtual Expressions?

8.1 Abstract

The intention of this experiment was to discover if an affect-laden VE could have an effect on either subjective mood assessments, affect-driven cognitions or both. Two independent, between-participant factors of the VE carried an affective valence: the environment and visually masked, facial expressions. The environment factor had two levels, one designed to be mildly pleasant and the other mildly unpleasant. The expression factor utilised two kinds of expressive, masked faces placed in doorways throughout the VE, the expressions being either angry or smiling. The intention was to measure the affective responses in a manner suggested by the component process model (Schere 2005), specifically subjective feeling and cognition. However, physiological measures were not appropriate as mood was to be measured, not emotion. Explicit subjective feelings were measured prior and post experiment using the Profile of Mood States (POMS) and the Positive and Negative Affect Schedule, Expanded (PANAS-X) respectively. Cognitive responses to the affective stimuli were measured using an emotional Stroop reaction test. This also introduced a further independent factor, word valence that was within-participants in form. Results showed that, although a direct effect on explicit feeling was not found, both environment and expression did interact significantly to affect reaction times. Word valence was found

to have no significant effects, interactive or otherwise. When environment was factored out, taking account of expression and word valence only, a similar result was obtained: explicit feelings were unaffected but expression had a significant effect on reaction times with word valence, once again having no effect. These results suggest a direction for further developing the use of visually masked, affect-laden VE stimuli to bring about mood state change: a direct, affect-contagion route is less likely to yield results than a more indirect, change in cognition. In addition, the effect of the environment factor also suggests the form of VE in which such a state change might best be effected.

8.2 Introduction

Previous experiments have looked at some of the effects of masking facial expressions within a VE. For example, Experiment 2 found a navigation bias effect away from negative facial expressions towards positive or neutral expressions and Experiment 4 found an aesthetic preference bias elicited by masked expression valence. These results may be seen as describing directed, i.e. voluntary, behavioural outcomes of affective processes. However, other affect components, such as explicit feelings or undirected reactions, have not been measured so far.

A change in explicit, subjective feelings on encountering a masked, virtual expression, may indicate the influence of a direct form of affect contagion, as exemplified in Zajonc's (1980) declaration that 'Preferences Need No Inferences'. Here the masked facial expression would be acting analogously to the DPS found in other masking studies (Neumann 1990, Skalska *et al.* 2006). Conversely, it could equally be the response of a cognitive judgement. Either way we might expect an accompanying behaviour, in a direction appropriate to the valence. For example, people in bad moods appear to observe detail more readily compared to those in a good mood (Levine & Pizarro 2004, Bohner *et al.* 1992). We might then expect those in bad moods to perform better on a reaction-time task that required attention to detail, for instance. If, on the other hand, we found no explicit feeling change yet *did* find a non-directed response then we could rule out any direct affect contagion. Instead, we might posit other affect-driven, cognitive processes were at work.

As well as the internal expression valence, affective effects could be enhanced by the setting. Research has pointed to the beneficial effects of the close proximity of naturalistic landscapes (Wells & Evans 2003, Taylor *et al.* 2001, Ulrich 1986). It is possible that environment could act as a modulator of affective effects within a VE. Therefore, the form a VE takes may have a useful interaction with any affect-laden stimuli within that VE. Whether culturally conditioned, innately wired or both, there does appear to be an affective dimension to the human relationship with environment.

When used to describe a particular landscape, the term ‘naturalistic’ can perhaps be seen as a loosely defined dimension ranging from wilderness, through rural countryside into suburbs and onto urban areas. Each, it could be argued, is successively less naturalistic. Attempts to define this measure more tightly could run into difficulties. For instance, we could assume that a landscape’s naturalistic value was inversely related to the amount of human shaping. We would need to show that suburbs depict more human working than a rural area. Whilst this may seem intuitively obvious it is not necessarily true of all instances. Areas of the UK countryside at least have been shaped and reshaped by centuries of farming whereas some suburbs may have been carved only recently from un-worked land. Further complexity is encountered when we consider that not all urban and suburban landscapes are regarded as equally pleasant by people. Nor are views consistent, with some people enjoying city life and others a more suburban existence. Additionally, of the many different types of wilderness, each may be experienced differently, depending on cultural factors or possibly innate preferences proposed, for instance, by the Savanah Hypothesis (Orians & Heerwagen 1992). Interesting though such debates are, they are beyond the scope of this thesis.

Another way to define 'naturalistic' would be by landscape content. A landscape could be regarded as being more naturalistic if human-made items were less frequent and natural objects more so. Again, though, we run into problems as we have merely shifted ambiguity away from whole landscapes and onto discreet objects. In order to obtain a quantitive value for naturalism, we would need to be able to define accurately what a natural object was. This would then allow a count per area, defining naturalism as a type of frequency. However, rather than being categorised as natural or not, objects themselves can be seen as falling on a continuum running between completely natural to completely manufactured (leaving aside semantics about whether something is unnatural just because humans constructed it: this philosophical debate can apply to landscapes as well and is also outside the scope of this thesis). Ambiguous objects might include hybrid flora that could not exist without human intervention. There may be some merit in apportioning such objects various fractional scores but it could be argued that we would merely be reducing the ambiguity down to a lower scale where it becomes less visible but still remains.

Nonetheless, categorising scenes according to content or form has been used to show various effects ranging from preference behaviour (Kaplan 1992) to hospital recovery times (Ulrich 1984). The latter may be a variation of a restoration theory effect: Nature Restoration Theory (NRT) describes how viewing natural scenes can reduce anxiety and stress (Ulrich 1986, Ulrich *et al.* 1991). Other forms of restoration theory exist such as Attention Restoration Theory (ART) (Kaplan 1995) which purports to shown that natural landscapes have a recovery effect on exhausted attentional abilities (Taylor *et al.* 2001, Wells & Evans 2003). Again, the landscapes used tend to be categorised by either content, form or both; hills with trees for example. This type of research contains a tendency to show that landscapes have a strong affective effect. The work, though, does so without the need to be too definite regarding what constitutes a natural landscape or not. It is even acknowledged that some (although relatively fewer) urban scenes may have a similar effect (Ulrich *et al.* 1991).

Virtual landscapes have also been studied to some extent in this way. Results seem to show similar effects to that of real landscapes or landscape images, with immersion level appearing to be a factor (de Kort *et al.* 2006) although how realistic the immersive experience has to be to be in any way restorative is inconclusive (de Kort & Ijsselsteijn 2006).

If a VE does need to be either non-stressful or even restorative, then a good starting point might be to look at landscapes which have been shown to achieve these and work backwards, analysing the content and form afterwards. Ulrich *et al.* (1991) provided us with one such description used in a successful recovery condition, see table 8.1 below.

Visual content	Sounds
Setting dominated by trees and other vegetation; some openness among trees; occasional light breeze in background; no people or animals.	Birds; light breeze.

Table 8.1. Natural Vegetation (after Ulrich et al. 1991).

Yet another description is provided by the Savanah Hypothesis (Orians & Heerwagen 1992). To simplify, the claim made here is that landscape preference is shaped by the environment within which humans evolved. This, it is further claimed, would be the savannah grasslands of East Africa. There are certain features given such as sheltering trees to avoid predators, good lines of sight to spot game and water. Incidentally, features such as these, it is also claimed, form the basis of our aesthetic appreciation of landscape imagery in art. Although such a hypothesis may be hard to establish scientifically it does appear to make intuitive sense and may well be good starting points for the design of a restorative VE.

As well as form and content, lighting can also play an important part in affective responses to an environment. For example, seasonal depression can become heightened during darker winter months, levels of sunshine appearing to be a contributory factor (Rastad *et al.* 2011). Therefore, another manipulatable aspect of a VE that was intended to enhance affect would be the lighting quality. A bright, sunlit day would be expected to be associated with a different affective response generally to a stormy, overcast sky that produced dull lighting (Schwarz & Clore 1983).

8.3 Method

This experiment used a 2 (environment: utopia vs. dystopia) x 2 (expression: angry-looking vs. smiling) x 3 (emotional Stroop test word valence: positive vs. neutral vs. negative) mixed model design, with a reaction time measure (emotional Stroop test) to test for effects; environment and expression were between-participants factors and word valence was a within-participants factor. After a practice period, participants completed 40 trials in total with 40 repetitions per environment level, 40 per expression level and 22 per valence level. Additionally, two inventories were used as measures of mood: one before the trials (Profile of Mood States) and one after (Positive and Negative Affect Scale, extended).

8.3.1 Participants

Nineteen people were recruited for the experiment, drawn from the UCL psychology subject pool, posters and word of mouth. Fourteen were male, two of which were left-handed, and five females, one of which was left-handed. All complied with the experimental protocol.

8.3.2 Materials

The hardware and software platform used is that given in the Methodology, see Chapter 3.

There were two VEs used for this experiment, one labelled *Utopia* the other *Dystopia*. Both had the same basic layout and the same movement freedom and constrains within their geometry. Each consisted of a series of areas, across which the participants had to move, see figures 8.1 and 8.2. The Dystopia environment was perhaps the closest in form to previous experiments, each area being courtyard-like. Each courtyard terminated in a pair of blinking doors, acting as masks, one of which the participants passed through into the next court. This environment was textured in sombre tones and lit with an overcast light. Shadows were muted. The courtyards were open to a grey sky suggesting stormy weather. A soundtrack of wind and rain accompanied Dystopia.

Figure 8.1. Images of the Utopia environment.

The Utopia environment did not look like a series of courtyards, despite participants being constrained to move within the same floor-plan as the Dystopia environment. Utopia had a series of high garden hedges perpendicular to the direction of travel. The hedges were geometrically equivalent to the front and back walls of the Dystopia rooms, also containing blinking doors through which the participant travelled. However, instead of the side walls of the Dystopian courts, Utopia had an open aspect allowing lateral views out across a pleasant green space, with lots of foliage, leading to water on either side. Distant hills were visible, rising from the water. However, participants were still constrained to move within the geometry of those sidewalls, despite their invisibility. The whole VE was designed to give the impression of traversing across a verdant island or isthmus. The lighting simulated a bright, sunny day with saturated colours and contrasting shadows. The accompanying soundtrack was of water lapping and birdsong.

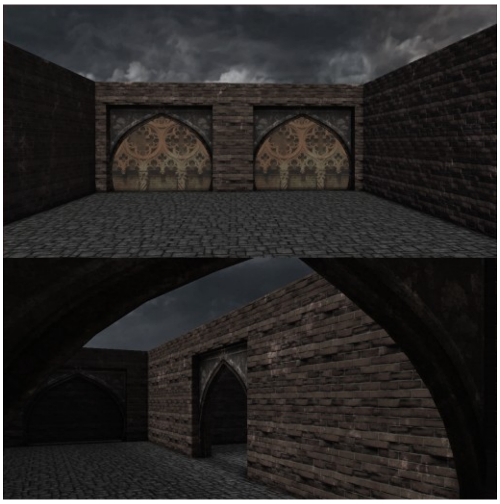


Figure 8.2. Images of the Dystopia environment.

There were twenty rooms in each VE. Utopia terminated in a pleasant, open area resembling a semi-wild park with copious amounts of flora and rocky features. This space afforded even wider views across the water to the distant hills and surrounding virtual world. It acted as a signal to the participants that they had reached the end of that world. Dystopia terminated in courtyard the same in form as the previous rooms, the stop signal being that the entire scene would glow red like a sunset. These two VEs formed one of the independent, between-participant factors: Environment.

Within the two worlds, a further, between-participants variable was expressed in the masked, virtual faces sat in each doorway. They displayed either a negative or a positive facial expression. The negative faces were composed of angry, disgusted or sad expressions, the positive of two types of smile. This was to control for the possibility that an effect may have arisen due to the peculiarities of a particular expression type and not for the general valence. For example, unlike the others, the angry expression contained an open mouth, displaying teeth. If this had been the only negative expression then it could be argued that any effects found may have been down to the presence of teeth. This would have been ambiguous in terms of affect as dental displays can also be present in positive expressions. The expressions were randomly distributed throughout their appropriate treatments: see table 8.4.

8.3.3 Measures

It was thought that a single measure would not be sufficiently reliable to indicate a change in a construct such as induced mood. To this end, two measures were be used: one to measure any implicit effects, the other, a questionnaire, to look for the explicit changes. Mood changes are thought to be characterised by lower intensity and slower onset/offset dynamics than emotion. Therefore, measures of physiological responses were deemed inappropriate for this particular study.

8.3.3.1 Mood Inventories

Medical inventories, such as the Clinical Outcomes for Routine Evaluation (CORE 2009) system used by NHS psychotherapeutic interventions, may be able to provide a subjective insight into a participant's affective experience. These do not necessarily need to be used on a clinical population but could be adapted for use within a non-clinical study. Ethical questions, though, are raised by the use of such measures. Participants would need to be made aware that any experiments are not clinical trials but use the same or similar instrumentation: any results would be treated collectively and so could not constitute a diagnosis or assessment of any kind for any one individual.

Some thought needed to be given as to how the measures would be applied. Simply 'bookending' the experiment with the same tests at the start and finish could have posed problems. Participants could have remembered some of their responses, particularly those at the beginning or end of an inventory, so-called 'primacy' and 'recency' effects. They may then have reproduced those scores later to appear consistent, introducing a demand confound. Alternatively, a demand variable could be introduced where a participant might, implicitly or explicitly, score the second inventory to show a change because of a belief that they had just experienced an experimental condition. To control for these memory effects, a test could be embedded within a very long questionnaire with answers to the 'dummy' questions being discarded before the final analysis. The inventory could be particularly loaded at the start and end with these dummy questions. However, care would be needed to ensure the dummy questions did not interfere or 'lead' on the actual data-producing questions.

A better approach might be to use two different inventories, one initially, the other following an experiment. This raises further problems of consistency though. For instance, each test may be measuring different constructs. Further, even if participants do not remember their previous entries, they may still answer particular tests in a particular biased way. This could be controlled for by using standardised inventories that have been thoroughly tested via previous research. Therefore, two standardised, validated inventories would need to be sourced which measured the same construct, or construct components, yet which would appear different to a participant.

Several inventories could be used to test for an explicit change in mood or affect. Two such are Profile of Mood States (POMS) (McNair *et al.* 1971) and the Positive and Negative Affect Schedule (PANAS) (Watson & Clark 1999). The POMS consists of several scales to measure various states such as Tension-Anxiety and Depression-Dejection. PANAS uses adjective rating to highlight state affect valence, positive and negative. The original PANAS measure was expanded also to measure specific affects: PANAS-X (Watson & Clark 1999).

Both POMS and PANAS-X were designed to measure similar affect constructs. Both can also be adapted to show changes in trait affect. Many POMS (long form) and PANAS-X sub-scales correlate well, see table 8.2. Therefore, they provide some opportunity for comparative analysis. For the purposes of this experiment, two main factors justify their use in bookending the trials. Firstly, they both aim to measure similar constructs, and in the case of some subscales, the same constructs. Secondly, they use differing word lists and so both appear dissimilar to the participant. This should control for demand characteristics involving participant memory of earlier scores, such as wishing to appear consistent. As a result, it is more likely that effects due to treatment will picked up, thus avoiding a type 2 error.

POMS sub-scale	PANAS-X sub-scale	Convergent correlation
Tension-Anxiety	Fear	0.85
Anger-Hostility	Hostility	0.91
Depression-Dejection	Sadness	0.85
Fatigue	Fatigue	0.89
Vigour	Positive Affect	0.86

Table 8.2. Comparison of POMS and PANAS-X subscales (Watson & Clark 1999).

The approach taken, therefore, was to utilise these subscales in POMS to check for intergroup discrimination before the trials and to use the PANAS-X equivalents to check for intergroup discrimination after. If the POMS results were not significant yet the PANAS-X results were then this would suggest an explicit subjective affect state change had taken place.

8.3.3.2 Emotional Stroop

Although the Emotional Stroop test may appear to be a straightforward variant of the standard Stroop test, it is not (Williams *et al.* 1996, Larsen *et al.* 2006). With a Stroop test the word content is related to the form i.e. the names of colours are used, displayed in various coloured inks. An Emotional Stroop test uses lists of affectively valenced words, whose meanings are unrelated to the ink colour, or at least should be. For example, the word ‘blue’, describing a melancholic state, might introduce a confound if displayed in blue and then later another colour, depending on the test structure.

Nevertheless, there are related elements between the tests. Both use reaction times as a measure based on recognition of the ink colour in which words are displayed. Both purport to indicate the disruption of recognition by discordant cognition. In the case of the Stroop test, when words are displayed in a different colour to the ink colour, and participants are asked to indicate the ink colour, it is argued that the automatic reading of the colour word conflicts with the attempt to respond with the correct ink colour name. The Emotional Stroop, it is argued, also indicates cognitive disruption but that this arises from the affective state of the participant. Depending on whether this state chimes with the valence of a word, the task response times will vary accordingly (Gilboa-Schechtman *et al.* 2000).

The Emotional Stroop test is a widely used behavioural measure to assess a participant's vigilance towards a particular affect valence (Fox 2008, p170), sometimes described as an attentional bias (Williams *et al.* 1996). The test by itself does not clearly distinguish trait from state affect but could indicate a difference in processing bias between groups congruent to affective state (Gilboa-Schechtman *et al.* 2000). In recent years, a problem has become apparent with the lexical characteristics of the words commonly used in the test (Larsen *et al.* 2006). For instance, emotive words may have different linguistic properties to neutral words, such as length etc., and care must be taken when choosing word lists (for instance, see Bradley & Lang 1999). This experiment used lists of lexically balanced words obtained via the author's correspondence with Professor R.J. Larsen, Washington University in St. Louis, USA:

Positive Words	**Neutral Words**	**Negative Words**	**Frequency Range**
kitten	magnet	rotten	2200-2500
cosy	twig	Lice	200-400
paradise	envelope	criminal	10000-30000
circus	pencil	coffin	3000-4000
picnic	shovel	fungus	800-1500
jewel	purse	slime	1800-2050
chocolate	satellite	infection	6000-20000
greet	viola	filth	900-1350
hero	bird	debt	13500-20000
applause	shepherd	massacre	3000-4000
cuddle	sneeze	maggot	450-600
serene	noodle	morgue	350-550
embrace	feather	allergy	2300-3300
giggle	oyster	stench	800-1350
perfume	ketchup	measles	1000-12000
mama	frog	scum	2800-3700
champagne	harmonica	paralysis	500-1500
hamster	pumpkin	obesity	900-1200
cinnamon	clarinet	cemetery	1200-1700
humility	kangaroo	mosquito	600-1300
lagoon	sleigh	manure	650-1050
spa	pea	Pus	100-1100

Table 8.3. Complete Emotional Stroop word lists (courtesy of R.J. Larsen, Washington University in St. Louis).

The words in table 8.3 were matched for frequency of use (exemplar column shown), length in letters, and orthographic neighbourhood size (Larsen *et al.* 2008). In addition, they all displayed high imagability and slow lexical decision time (Larsen *et al.* 2006) (author's correspondence).

The Emotional Stroop test was constructed using Affect4 (Spruyt *et al.* 2010) and run on a Core2 Duo laptop with 4GB RAM and a 1GB graphics card. Words from all three word-lists were presented one at a time, at random, central to a black background. The words were approximately 1cm high, in the Ariel font and in one of four coloured inks chosen at random, see figure 8.3 for examples. A response key guide was visible onscreen throughout the test, see figure 8.4. The keys were chosen for convenience of positioning on the keyboard with the colours being randomly assigned to them. As each word appeared, participants were instructed to press the key representing that colour: 'z' for red etc. Response times per key-press were recorded, forming the basis of the dependent measure.

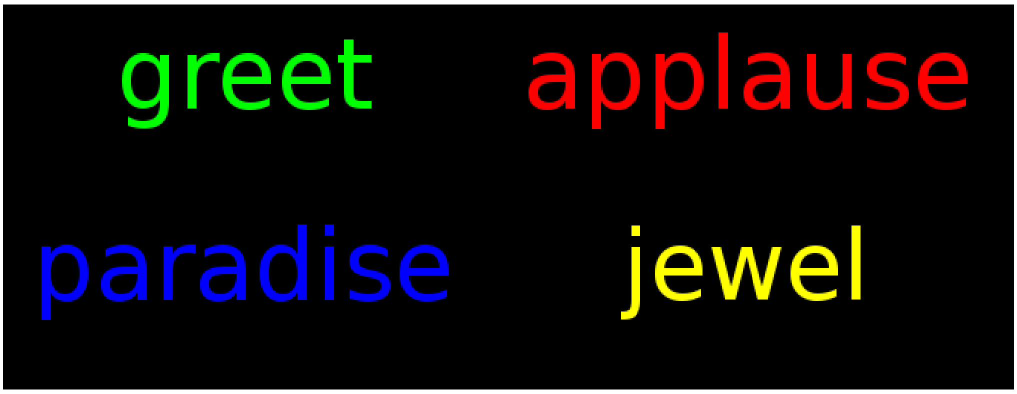


Figure 8.3. Examples of word colours for the Emotional Stroop test.

Figure 8.4. Onscreen key guide for the Emotional Stroop test.

It could be argued that the Emotional Stroop test should be run both before and after trials but this would have presented a confound in the form of a training effect, similar to the 'bookending' problem that the use of POMS and PANAS-X was designed to avoid. One group might have been able to learn the Emotional Stroop task more quickly than another, so distorting any comparison. Using a comparable task in parallel to test for training effects on the day would have needed a task which was similar enough to the Emotional Stroop task to allow a valid comparison but this very similarity could itself introduce a training confound to that which it was intended to avoid. Running a comparable task on separate days to the trial could also invalidate any comparison. The Emotional Stroop results therefore rest on the assumption that there were no affective differences between the groups prior to the trials, a reasonable assumption given the inventory results, see section 8.5 below.

8.3.3.3 Final grouping form

One issue we may encounter concerns the duration of induced mood effects. Previous studies have shown these effects to be relatively short lived at ten minutes or so (Gilet 2008). This study was not immediately concerned with elongating these effects although this would be a necessary goal to reach if we were to be able to use the results for the intended purpose of longer-term affect shift. Nonetheless, duration was still a concern for the current study in as much as it necessitated a particular change from the previous four experiments: this time, a between-participants design was necessary to capture any effects. Using a within-participants design, any affect-change duration could produce a confounding order effect: if participants experience the experimental condition prior to the control, as half the sample would be expected to do, an 'affect hangover' could contaminate the control results, if any effects had not worn off during an inter-condition interval. The final treatment groups are shown in table 8.4, below.

	Environment factor	
Expression factor	Utopia Positive	Dystopia Positive
	Utopia Negative	Dystopia Negative

Table 8.4. Experimental Groups.

Each group completed a POMS and a PANAS-X inventory prior and post experiment respectively.

Each group completed the emotional Stroop test post experiment, where they experienced the within-participants factor, word valence. Therefore, within this mixed model, repeated measures design there were three independent factors:

1. Environment, a between-participants variable of two levels labelled Utopia and Dystopia;

2. Expression, a between-participants variable of two levels labelled positive and negative;

3. Word Valence, within-participants with three levels labelled positive, negative and neutral.

There were also three dependant factors:

1. The pre-trial POMS scores for 5 sub-components form an explicit measure;

2. The post-trial PANAS-X scores for 5 sub-components form an explicit, measure;

3. The post-trial reaction times form the implicit processes measure. Scores were collapsed across colour. This would control for any differential colour effects as colours were distributed at random over the many trials.

Participants were randomly assigned to a negative or a positive expression version of one of the two environment VEs.

8.3.4 Procedures

The order of experimental events for each participant was as follows:

1. Participants were briefed about their task in the experiment.

2. Participants were then asked to complete a POMS inventory.

3. They were then introduced the VE and given a short practice to become accustomed to the controls.

4. Participants then carried out the first task, which was simply to navigate through the 20 rooms of their assigned VE. A soundtrack appropriate to that VE was run in parallel.

5. A short break of a few minutes allowed participants to rest their eyes from the stereoscopic display.

6. Step 3 was repeated, with the same VE.

7. Participants were asked to complete a PANAS-X inventory.

8. They were then introduced to the Emotional Stroop test and given a short practice.

9. The full Emotional Stroop test then followed. There were 66 trials (words) per participant.

10. Participants were debriefed.

8.4 Hypotheses

Hypothesis 5a

There will be *no* significant interactions or effects from the POMS scores.

Hypothesis 5b

There *will be* significant interactions or effects from the PANAS-X scores.

Hypothesis 5c

There will be significant interactions or effects from the Emotional Stroop scores.

Hypothesis 5d

Where no interactions are found in the Emotional Stroop scores, the word valence within-factor will show a significant main effect.

Hypothesis 5e

Where no interactions are found in the Emotional Stroop scores, the environment between-groups factor will show a main effect.

Hypothesis 5f

Where no interactions are found in the Emotional Stroop tests, the expression within-groups factor will show a significant main effect.

8.5 Results

8.5.1 Descriptive Statistics

The between factor 'environment' contained one group of twelve (dystopia) and one group of seven (utopia) participants. The between factor 'expression' contained one group of ten (smile) and one group of nine (angry).

It was decided to analyse the data using two sets of figures: Set 1 contained data from all of the four treatment groups; Set 2 contained only data from the two Dystopia treatments. In this way, any problems arising from the imbalance of the Environment factor could be highlighted. In the event, comparable results from both data sets matched.

POMS:

Environment	Expression	Affect Dimension	mean	sd.
dystopia	negative	Tension-Anxiety	6.5	4.6
dystopia	negative	Depression-Dejection	8.2	9.7
dystopia	negative	Anger-Hostility	5.7	8.9
dystopia	negative	Vigour	15.5	8.0
dystopia	negative	Fatigue	7.2	6.9
dystopia	positive	Tension-Anxiety	3.7	1.9
dystopia	positive	Depression-Dejection	5.2	4.1
dystopia	positive	Anger-Hostility	2.3	3.0
dystopia	positive	Vigour	15.7	5.4
dystopia	positive	Fatigue	5.2	4.8
utopia	negative	Tension-Anxiety	5.7	4.7
utopia	negative	Depression-Dejection	3.3	4.2
utopia	negative	Anger-Hostility	3.7	4.7
utopia	negative	Vigour	14.7	5.5
utopia	negative	Fatigue	2.3	2.1
utopia	positive	Tension-Anxiety	6.5	4.8
utopia	positive	Depression-Dejection	6.5	5.9
utopia	positive	Anger-Hostility	3.3	4.0
utopia	positive	Vigour	16.8	3.4
utopia	positive	Fatigue	7.3	3.3

Table 8.5. POMS Set 1 means.

Set 2 means for the POMS inventory is a repeat of the figures for the two Dystopia groups

PANAS-X:

Environment	Expression	Affect Dimension	mean	sd.
dystopia	negative	Fear	8.2	4.8
dystopia	negative	Sadness	9.7	5.5
dystopia	negative	Hostility	9.8	5.8
dystopia	negative	Positive Affect	28.8	13.5
dystopia	negative	Fatigue	9.2	5.8
dystopia	positive	Fear	8.7	2.7
dystopia	positive	Sadness	9.8	3.5
dystopia	positive	Hostility	7.3	1.8
dystopia	positive	Positive Affect	32.3	7.6
dystopia	positive	Fatigue	7.2	2.4
utopia	negative	Fear	8.3	3.2
utopia	negative	Sadness	6.3	2.3
utopia	negative	Hostility	8.0	1.7
utopia	negative	Positive Affect	27.0	3.6
utopia	negative	Fatigue	5.3	2.3
utopia	positive	Fear	8.3	1.7
utopia	positive	Sadness	8.5	3.3
utopia	positive	Hostility	8.3	2.9
utopia	positive	Positive Affect	35.0	4.1
utopia	positive	Fatigue	7.3	2.1

Table 8.6. PANAS-X Set 1 means.

Set 2 means for the PANAS-X inventory is a repeat of the figures for the two Dystopia groups.

A point to note for both the POMS an PANAS-X descriptive results is that the categories that correspond to positive affect consistently score more highly than all of the other categories. For the POMS questionnaire, this is the ‘vigour’ category and for PANAS-X, this is labelled ‘Positive Affect’. This would suggest that, despite individual variations, overall the participant group expressed a positive outlook via the inventory scores. If this indeed reflected a general internal attitude then we might expect this be a contributing factor to the results observed. For example, it has been shown that people in a positive mood pay less attention to details, concentrating instead on global features (Gasper and Clore 2002, see section 2.4.3 above). Therefore, this factor may result in a lessening of any effects found by this experiment.

Emotional Stroop:

On the Emotional Stroop data, the first trial response from each participant was discarded, it being consistently disproportionately large when compared to the other times. This left 65 trials per participant. If a participant pressed an incorrect key, i.e. one that did not correspond to the colour of the word being shown, then this was recorded as an error. These were discarded from the final results, the error rate being approximately 4%. Errors were removed because it was felt that whatever cognitions resulted in the errors these must have been differentiated in some way from non-error cognitions, which may have introduce a confound. This rests on the assumption that cognition is involved in error but, even if this is not the case, the removal of 4% of the data was felt to be a small enough amount to not cause a significant difference when compared to the potential of 4% erroneous data to display false results.

Unfortunately, stripping out the erroneous data left an uneven amount of scores per participant across the groups. This would have resulted in empty treatment cells for some participants, weakening the analysis. The solution was to truncate the score data down to a limit set by the participant with the most errors in each data set. For Set 1, the data were truncated to meant 48 trial scores per participant and for Set 2, this meant 60 trial scores per participant.

Environment	Expression	Word valence	mean	sd.
utopia	positive	positive	628.6	220.7
utopia	positive	neutral	587.9	170.2
utopia	positive	negative	624.5	202.6
utopia	negative	positive	697.4	288.7
utopia	negative	neutral	680.1	326.6
utopia	negative	negative	671.4	340.6
dystopia	positive	positive	759.6	399.4
dystopia	positive	neutral	853.3	481.6
dystopia	positive	negative	824.1	558.8
dystopia	negative	positive	605.6	206.2
dystopia	negative	neutral	694.0	315.5
dystopia	negative	negative	627.9	209.4

Table 8.7. The Set 1 Emotional Stroop group means, in milliseconds.

Expression	wordValence	mean	sd.
positive	positive	781.8	424.8
positive	neutral	847.8	461.8
positive	negative	814.6	548.0
negative	positive	621.1	219.2
negative	neutral	620.7	219.2
negative	negative	623.3	201.1

Table 8.8. The Set 2 Emotional Stroop group means, in milliseconds.

8.5.2 Inferential Statistics

All the data were analysed using a mixed model in SPSS.

POMS:

Main Effects & Interactions	**F**	**p**	**significance**
Environment	0.11	0.74	p>0.05
Expression	0.14	0.71	p>0.05
AffectDimension	11.37	0.00	p<0.05, sig.
Environment * Expression	1.81	0.18	p>0.05
Environment * AffectDimension	0.23	0.92	p>0.05
Expression * AffectDimension	0.34	0.85	p>0.05
Environment * Expression * AffectDimension	0.15	0.96	p>0.05

Table 8.9. Set 1 POMS interactions and main effects.

Table 8.9 shows no significant interactions or main effects, except for the affect dimension, which would be expected between the orthogonal sub-categories of a well-constructed inventory.

Main Effects & Interactions	**F**	**p**	**significance**
Expression	0.53	0.47	p>0.05
AffectDimension	6.49	0.00	P<0.05, sig.
Expression* AffectDimension	0.21	0.93	p>0.05

Table 8.10. Set 2 POMS interactions and main effects.

Table 8.10 shows no significant interactions or main effects, except for the affect dimension, as expected.

PANAS-X:

Main Effects & Interactions	F	p	significance
Environment	0.58	0.45	p>0.05
Expression	1.08	0.31	p>0.05
AffectDimension	25.68	0.00	p<0.05, sig.
Environment * Expression	1.21	0.28	p>0.05
Environment * AffectDimension	0.29	0.89	p>0.05
Expression * AffectDimension	0.59	0.67	p>0.05
Environment * Expression * AffectDimension	0.25	0.91	p>0.05

Table 8.11 Set 1 PANAS-X interactions and main effects.

Table 8.11. shows no significant interactions or main effects, except for the affect dimension.

Main Effects & Interactions	F	p	significance
Expression	0.00	0.97	p>0.05
AffectDimension	11.80	0.00	P<0.05, sig.
Expression* AffectDimension	0.40	0.80	p>0.05

Table 8.12. Set 2 PANAS-X interactions and main effects.

Table 8.12 shows no significant interactions or main effects, except for the affect dimension.

Emotional Stroop:

Main Effects & Interactions	F	p	significance
Environment	10.84	0.00	P <0.05, sig.
Expression	7.27	0.01	P <0.05, sig.
WordValence	0.51	0.60	p>0.05
Environment * Expression	21.34	0.00	P<0.05, sig.
Environment * WordValence	2.46	0.09	p>0.05
Expression * WordValence	0.17	0.85	p>0.05
Environment * Expression * WordValence	0.10	0.90	p>0.05

Table 8.13. Set 1 Emotional Stroop interactions and main effects.

Table 8.13 shows significant interactions between Environment and Expression as well as main effects for Environment and Expression: see Discussion, section 8.6 below. Word valence did not produce any significant interactions or a main effect.

Main Effects & Interactions	**F**	**p**	**significance**
Expression	38.10	0.00	P<0.05, sig.
WordValence	2.68	0.07	P>0.05
Expression* WordValence	0.52	0.60	p>0.05

Table 8.14. Set 2 Emotional Stroop interactions and main effects.

Table 8.14 shows a significant main effect for Expression. Word valence did not produce a significant interaction or main effect.

The data residuals generally described a Gaussian distribution but a positive skew was evident as can be found in data with a possible range of zero upwards.

8.6 Discussion

The first point of note is that the analyses of both Set 1 and Set 2 produced results that concurred: the inventory scores of both sets produced similar results for both POMS or PANAS-X and this was also the case for the Emotional Stroop tests. The exception to this came from of the Environment factor, which only applied to Set 1 across all of the three tests.

There were no significant between-groups differences for the POMS data, for the Expression factor on Set 1, or the Expression and Environment factors for Set 2, whether as interactions or main effects. The Affect Dimension factor was significant for both sets, which is to be expected from a well-constructed inventory. As POMS was the pre-trials test, we can see that there were no initial, explicit, affective group differences for either Environment or expression.

There were no significant between-groups differences for the PANAS-X data, for the Expression factor on Set 1, or the Expression and Environment factors for Set 2, whether as interactions or main effects. As with the POMS data, the Affect Dimension factor was significant for both sets. As PANAS-X was the post-trials test, we can see that there were no final, explicit, affective group differences for either Environment or Expression. Therefore, neither factor appeared to produce an *explicit* affective effect. If such an effect were present then PANAS-X may have not been powerful enough to detect it at that sample size, a type 2 error. However, this argument could also equally apply to the POMS test.

The Expression factor did produce significant results for both sets of data. For Set 1, this factor produced a significant interaction with Environment, as well as a main effect. We might exercise caution in giving any evidential weight to the Expression main effect because of the presence of the interaction. However, Expression also produced a significant effect for Set 2 where no interaction was present. Word Valence did not produce any main effects or interactions with the other factors. From this, we can see that Expression, modulated by Environment, had an effect on response time regardless of Word Valence. This could be interpreted as demonstrating that the semantic processing of the words did not interfere with the effect of either the masked expression effect or the environmental modulation of that effect.

Interestingly, if we look at the results of Set 2, Word Valence is almost significant ($p=0.07$), see table 8.14. We can group the descriptive results of Set 2 according to whether their Word Valence factor level contained either emotive words or neutral words, see table 8.8. There were four emotive groups that used either positive or negative Word Valence levels, and two neutral Word Valence groups. On average, the emotive-words groups' response was 710.2 ms and the neutral groups' average was 734.3 ms. This could be indicating that emotive words were being processed faster than neutral, perhaps as a result of an attentional effect. For example, Participants' eyes could wander momentarily across the screen between words but jump back more quickly for emotive words, indicating a parallel processing to 'spotlight' attentional jump.

Similarly, we could group the results by congruence: the two incongruent groups would be those where Expression and Word Valence were mis-matched, i.e. one was positive whilst the other was negative; for the two congruent groups, Expression and Word Valence matched. The two remaining groups can be labelled 'neutral' as designated by their Word Valence factor, corresponding to the non-emotive grouping, above. The congruent groups were slightly faster on average, at 702.6 ms, than their incongruent counterparts, who averaged 717.9ms per response. The two neutral groups were the slowest of all with an average of 734.3ms. These figures could again be explained by an attentional jump effect in that congruent groups generate the fastest attentional response. Therefore, excluding the environment as Set 2 does, semantic processing might be having an effect.

A further grouping that can be made is to compare positive word responses against negative. Here the results show average positive word responses were 814.7ms against negative word responses of 621.7ms. As previously noted, research has suggested that when in negative mood, participants pay more attention to details (Gasper and Clore 2002), an advantage on a task such as the Emotional Stroop. These results would seem to concur, assuming that an implicit affective state change had taken place.

In addition, as pointed out in the descriptive statistics for the two inventories (see section 8.5.1 above), an overall positive outlook was expressed by the whole participant group for this experiment. It is entirely possible, therefore, that a different group, which expressed a less positive overall mood, may have produced an even larger difference between positive and negative responses, even given differing shifts in affect due to the experimental treatments.

8.7 Conclusions

There does not appear to have been an effect of either environment or masked expression on explicit affect. This might rule out any explicit affect contagion. There may have been implicit contagion when environment was excluded although the word valence still fell short of being significant. When environment was a factor, as was the case with the Set 1 figures, semantic interference does seem to disappear or at least diminish below a level detectable by the tests used.

It is possible that environment can override any semantic task effects, giving rise to the intriguing possibility that environmental factors may be robust enough to transfer across contexts i.e. whatever the affective valence of a situation encountered subsequent to a VE session, environmental factors within that VE will interact with masked expressions to produce an effect.

Even if no semantic interference occurred, there is nevertheless strong evidence of an effect arising from the Expression factor manipulations. This implies that some form of cognitive manipulation took place, which holds out the promise of using masked expressions to modify cognitive processes. As these stimuli are social in nature, we might speculate that an indirect modification of background affect might be achieved via a change in implicit self-esteem cognitions.

9 Experiment 6 – Cognitive Bias Modification via Visually Masked Expressions in a Virtual Environment

9.1 Abstract

This experiment was designed to test cognitive bias modifications using masked expressions within a VE. The intention was to prime a bias away from negative expressions, towards positive expressions. Participants were instructed to follow a route through a virtual world in the form of a series of garden hedges, similar to a maze. At each junction, they could choose to go through one of two doorways, each of which was blocked by a masked facial expression. In one condition, only the doorway with the smiling expression would lead to the next area, passing through an angry expression would lead to a dead end and the participant would have to retrace their route back to the other doorway. A second condition reversed the expression placements. Participants were required to complete the route as quickly as possible. They were then asked to complete an adapted CBM training task that was originally designed to modify attentional bias. This task used response times to attentional bias towards smiling expressions and away from negative ones. Results showed as significant difference between group times: participants who experienced the smiling prime condition generally displayed faster reaction times when compared to those who experience the angry primes. This indicates a transfer from implicit to explicit learning.

9.2 Introduction

The 6th and final experiment tested for cognition modification effects of masked facial expressions within a VE. The cognition in question was visual attention and the aim was to introduce a scanning bias for expression type by exposing participants to masked virtual expressions. Of the two expressions, see Materials, section 9.2.3 below, it was intended that one would become associated with making successful navigation choices throughout the VE. This would act as an implicit reward aiding speed of completion of the navigation task. The other expression would act in the opposite capacity, becoming associated with dead-ends and the slowing of progress. In this way, it was intended that participants become implicitly primed to bias attention towards one expression, to the detriment of the other.

The effects of the navigation task were measured with an attentional CBM task. CBM tasks have been shown to alter social processes such as expression scanning (Beard *et al.* 2011). The use of CBM techniques holds out the promise of developing new forms of intervention for psychiatric disorders such as affect disorders like social anxiety disorder. People suffering from the latter can display scanning behaviour in social situations, looking for stimuli that they regard as threat cues. Certain facial expressions can act as such and, once found will raise anxiety levels, whether relevant to the person or not. CBM tasks can be constructed to divert attention away from these stimuli habitually and so lower social anxiety levels. However, CBM techniques, just as with any other intervention or potential intervention, are not completely efficacious and can be regarded as being in the early days of development. Even mature interventions such as CBT are seen as having limits in terms of treatment success. It can be argued that one of the reasons why an intervention may fail for a particular individual could lie with the form of the intervention being vulnerable to disruptive cognitions held by the individual them-self. They may take the form of an unspoken, unacknowledged or even unknown desire on the part of the recipient for the therapy ultimately to fail.

These kinds of psychological processes may be regarded differently according to a therapist's model. Cognitive therapists may see them as implicit, biased information processes, psychodynamic therapists would talk about unconscious resistances and ego preservation. Whatever the explanatory model, though, it would seem reasonable to assume that interventions that do not address them may achieve very little in the presence of such processes. It may be possible to interrupt these kinds of processing biases by using interventions that, as it were, work on the same level as them. If they are indeed implicit processes then it follows that intervention may be achieved by the use of some kind of implicit stimulus. Masked, non-planar objects within a VE could form the basis of such implicit intervention by carrying greater plausibility than explicit, planar stimuli presented via non-interactive media lacking any illusion of place. This would be predicated on demonstrating that such stimuli could actually modify affective cognition.

9.3 Method

This experiment used a 2 condition (expression: angry-looking vs. smiling static) between-participants design, with a reaction time measure (CBM letter discrimination) to test for between-condition effects. After a practice period, participants completed 40 trials in total with 40 repetitions per expression level and 64 trials for the reaction test.

9.3.1 Participants

Eighteen participants were recruited for the experiment, drawn from posters and word of mouth. All complied with the experimental protocol. The datasets represented six right-handed females, one left-handed female, one left-handed male with the remainder being right-handed males.

9.3.2 Materials

The hardware and software platforms used were those described in the Methodology set out in Chapter 3.

There were two VEs used in this experiment, each explicitly the same in form. As with the previous experiments, both of the VEs consisted of a series of spaces connected by doorways. At the far end of each room, opposite to the point of entry, were the two doors. Passing through either door repeatedly led the participants on into further similar spaces. For this experiment, the VEs borrowed elements from the VEs found in Experiment 5. The VEs for Experiment 6 were closed rooms, similar to the previous Dystopian condition and other VEs used earlier. However, the lighting and surface textures resembled those of the Utopian condition. Both of the VEs were lit as if on a bright sunny day. The walls of the rooms resembled hedges, lending the VEs the feel of a tall garden maze, see figure 9.1, below.

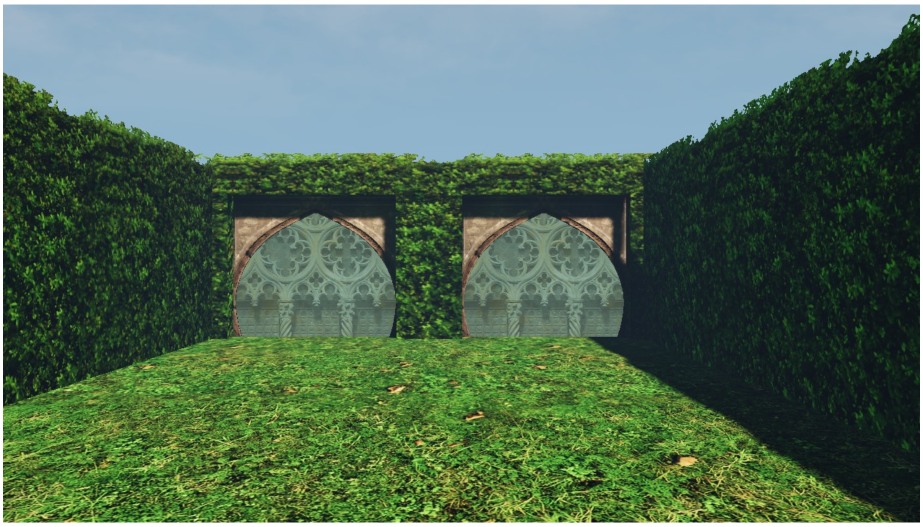


Figure 9.1. A view of the VE.

The two VEs, although having the same explicit appearance, were different in the masked targets that they contained. One included angry-looking masked expressions, the other smiling masked expressions.

The participant was instructed to make their way through the VE as quickly as possible. Regarding each pair of doors, one doorway would lead on to the next room, the other would lead to a dead end. Both doors blinked, synchronously, as if masking a target but only one did. The doorway that led to the next room would be the one that contained the masked face, the other door contained no target. The target expressions were therefore intended to act as implicit primes, pointing the way out to the next room. As the participant's goal was to navigate to the end of the VE as quickly as possible, processing which took account of the position of the targets should aid faster completion. In this way, participants may become implicitly primed to respond for a particular facial expression in a scene. There were twenty rooms per VE.

Following exposure to a priming VE, participant reaction times were collected during a CBM test. The test was based on a form used in CBM research to bias attention away from one expression type and towards another (Beard *et al.* 2011). A succession of pairs of faces, angry and smiling, were displayed on a screen for 500ms. Each expression was randomly displayed on either the left or the right of the screen. As both faces disappeared, they were replaced by a single letter, either 'E' or 'F' in place of one of the faces. Participants were instructed to press a letter key corresponding to the letter on screen. However, the letters always actually appeared under the smiling face. In this way, participant attention was primed towards smiling faces and away from the angry expressions.

This CBM task is normally used as an initial priming condition of an experiment, as the basis of an independent variable. In this study, though, it was used as a test to generate dependant variable figures. It is theorised that participants who experienced smiling expressions in their VE task will be pre-primed to search for the congruent expressions, smiling faces, of the CBM task. It follows that the reaction times of that group of participants should be lower than those who experience angry targets in the VE. If such a between-groups difference were found then this would be evidence of implicit modification of social cognition.

The CBM test was constructed in Affect4.0 (Spruyt *et al.* 2010). See Appendix 6 for examples of the CBM test stimuli.

9.3.3 Measures

There was one independent, fixed effect: Expression. Expression had two levels: smiling and angry, equating to the two VE conditions and the two final analysis groups.

The dependant variable comprised of reaction times recorded during the CBM test task. There were 64 test trials per participant.

This experiment was a repeated-measures, between-participants design as each participant experienced only one condition but multiple reaction times were recorded in the test task.

9.3.4 Procedures

The order of experimental events for each participant was as follows:

1. Participants were briefed about their task in the experiment.

2. They were then introduced the VE and given a short practice to become accustomed to the controls.

3. Participants were randomly assigned to a condition then carried out the first task, which was simply to navigate through the 20 rooms of their assigned VE.

4. A short break of a few minutes allowed participants to rest their eyes from the stereoscopic display.

5. Step 3 was repeated, with the same VE.

6. The CBM test was explained and the task then followed.

7. Participants were debriefed.

9.4 Hypothesis

Hypothesis 6

The reaction times of the smiling prime group will be significantly lower than those of the angry prime group.

9.5 Results

If a participant chose the incorrect letter key, this was recorded as an error. Errors were removed from the data and the number of times per participant was truncated to that of the most error prone participants. This ensured that the overall dataset contained a balanced number of trials per participant: sixty-one. The eighteen participants were evenly distributed between the two groups.

9.5.1 Descriptive Statistics

The overall error rate was 2.5%, approximately balanced across both groups. The Smiling group mean was observed to be 595.5 with a standard deviation of 158.1. The Angry group mean was 693.7 with a standard deviation of 240.6.

9.5.2 Inferential Statistics

The data were analysed in two ways:

1. The mean reaction times per participant were aggregated to form two groups: smiling and angry, corresponding to the VE that each participant had experienced. When analysed via a t test the results were $p = 0.04$. This was found to be the case whether equal variance was assumed within the groups or not.

2. All of the participant times were aggregated to form the two groups smiling and angry. These two groups were then subjected to a t test, the outcome of which was $p = 0.00$. As with the first analysis, the result was found to be the same whether equal variance was assumed within the groups or not.

Cohen’s d was found to be 0.47, which would constitute a small to medium effect by traditional standards.

9.6 Discussion

The hypothesis was supported, whether data were analysed using scores collapsed into two groups or using two groups of participant means. The effect size found is encouraging.

Repeated measures analyses may hide error variance if all participant scores are treated as if coming from one participant. In effect, inter-participant error becomes divided up into the analysis groups, increasing the possibility of a type 1 error. Analysis of participant means can provide an extra check. With this experiment, both forms of analysis proved significant.

When carrying out the CBM task, responses appeared to show a priming effect, rendering participants more vigilant for the particular masked expression type to which they were exposed within the VE.

This is an interesting effect not least because it indicates a transfer of learning from implicit to explicit processes, the implicit priming from the VE generating an explicit effect in the CBM task. However, the picture may be more complex than this. It is not clear that the CBM task itself does not make any use of implicit processes. That it would be expected to would be a reasonable assumption given the mean reaction speeds of the participants, in either of the two groups. For the CBM task, participants would logically have to:

a. move their eyes from the fixation cross when the faces appear;

b. attend to one or other face to read its expression;

c. if angry, move to the other face, if smiling keep gazed fixed on that spot;

d. when the letter appears, recognise which it is;

e. respond with the appropriate letter on the keyboard.

Additionally, somewhere around the first step an implicit, visual scan of the whole screen would possibly be taking place in parallel. As all of this took place in a little over 500ms, it is unlikely that there would not be enough time to allow for conscious deliberation to occur during *all* of the steps outlined. It follows that for at least part of the time, implicit processes would be involved i.e. the participant would arrive at a conclusion without explicitly knowing how they had arrived at it. It is likely that the recognition of a particular expression type would be an example of this, utilising implicit social processes for instance. Attention is not merely a passive process as active biases can increase vigilance towards a particular stimulus type (see section 2.3.4 for instance). It is within these processes that implicit priming can have its effect.

9.7 Conclusions

Given the above discussion, it would be reasonable to assume that implicit as well as explicit processes were at work in the CBM task. This blend of process types may provide one explanation as to how implicit learning can appear to be transferred to an explicit task. Another explanation could be grounded in the assumption that the difference between explicit and implicit processes arises due a category error, imposed by current theoretical leanings: the brain may simply not recognise this dichotomy, at an operational level at least, below the creation of the experience of being conscious. When engaged in the CBM task, participants may enter a state similar to Csikszentmihalyi's 'flow' (1991).

However, regardless of the model of underlying process activity, the above VEs have been shown to be capable of modifying social processes by implicit means. It may be particularly hard to treat some sufferers of social affect disorder and the like, due to the interference of cognitions that prevent the explicit acceptance of treatment. Undermining these cognitions by implicit means may be possible, as the results of this experiment suggest.

More generally, the modification of cognition by implicit means offers the hope that VEs can be constructed which are capable of disrupting processes that keep any type of affect disorder stable. Exposure to masked of facial expressions could be but one of many techniques that could be applied although it may difficult to find stimuli with the same degree of effect.

10 Conclusions, Impact and Future Work

10.1 General Conclusions

This thesis attempted to achieve several goals. It set out to test whether visual masking was possible within a stereoscopic VE, using non-planar targets. Variables were tested ranging from varying participant instructions to include a threat, varying virtual surface qualities, varying target dynamics and varying target type. Building on this technical work, the thesis also considered the affective effects of non-planar, virtual targets. Either expressive faces were paired against domestic objects or positive expressions were paired against negative expressions. Affective constructs, such as mood, were homed in on via several types of measure that included behavioural, cognitive and self-reflective responses.

Several effects were found. Visual making has been shown to be possible within a stereoscopic VE using non-planar targets and that moderate depth disparity between target and mask does not act to disrupt this. It was also shown that such perceptual effects could be modulated by what could be considered construction variables such as target exposure or target and mask surface qualities. Colour saturation, surface texture fluency, target form, exposure of target form, and environment context have all been shown to have some effect.

Some affective effects were found: choice behaviour appeared to favour masked targets showing positive and neutral expressions over those with negative facial expressions. In addition, a cognitive bias appears to have been implicitly induced and transferred to explicit behaviour. The bias in question has been associated, by other research, with improvements in affect.

The effects found tended to be small to medium, in traditional terms. This could have been due to one or more factors:

- **a.** It may be that the effect sizes were a product of the numbers of participants used in the experiments and that having a greater number of participants could increase effects. This is problematic though as it implies that the participants were not typical and that including more may result in increased variance. Indeed, it could also be that testing more participants would lower variance, possibly leading to no observed effects. This remains as an issue for all research that uses statistical methods.

- **b.** Alternatively, it is possible that the VEs were not constructed in such a way as to engender large effects. This issue was recognised by several experiments in this thesis that attempted to manipulate construction variables in order to modulate effects (see Chapters 5, 6 & 7). Further work may prove fruitful in this direction as better VE construction knowledge may lead to larger affective effects.

c. Yet another possibility is that the effects themselves may just be diminutive in nature and so no amount of modulation or variable tweaking would amplify them. This is doubtful because variance was observed by some construction variable change.

Although some experiments found effects, others gave unexpected results. Experiment 4, for instance, did not find that dynamic targets increased effect, quite the opposite in fact: dynamic targets appeared to perform poorly in comparison to static, suggesting that the number of exposures to a full expression might be important. It did not appear that dynamic targets were implicitly read by participants as animated (which would be expected to increase the effect). Further work would be needed, perhaps using neuro-imaging, to support or weaken this position.

Participant numbers should have been sufficient in all experiments to expose any effects, with the possible exception of Experiment 5. Even though most experiments in this thesis used a within-participants design, the last two did not. This may have been a particular issue for Experiment 5 that only found environmental and expression effects and no *significant* word valence effect. It may be that greater numbers in this mixed model experiment would expose an effect. This suggests that a future repeat of this experiment, using higher numbers, may yield a result but this is by no means certain.

Generally, those experiments that focused on technical aspects have tended to yield stronger results than those that focused on affect. However, in terms of the overall goal of this thesis, the most promising results appear to be those produced by the 6th experiment. These results support the idea that it is possible to construct VEs that can implicitly alter or introduce a cognitive bias, in particular in social processes affected by stimuli with an affective valence. The 'facial scanning' behaviour of people suffering from social anxiety disorder could be vulnerable to these techniques. These techniques could therefore prove effective for the development of an intervention, given the propensity of some people to resist treatment.

Beyond social cognitions, one suggestion would be to use masked phobic stimuli prior to using explicit versions. Whilst masked phobic stimuli certainly do not appear to have the same strength of effect as their explicit counterpart, there may be enough of an effect to use masked stimuli to 'prepare' a patient prior to explicit treatment. An implicit lowering of a patient's panic-inducing cognitions might make it more likely that they will continue treatment by participating in an explicit programme, lowering the risk of general patient dropout.

10.2 Impact

10.2.1 Research Impact

This research manifests several areas of impact. Firstly, it supports the approach that possibilities can arise from reviewing the history of psychological science, and its cognates, for effects that can be applied within VE science (see, for example, Yuan & Steed 2010). This thesis is part of that on-going effort. Such syntheses help to further both sciences by opening up new research perspectives. Performing visual masking within VEs, for example, not only points to further research possibilities within visual masking but also should support thinking about VEs as being more than simulations of the real world. A whole host of established psychological effects, perceptual and otherwise, are available which might benefit from the unique qualities that a VE exhibits beyond simulation.

Specific to visual masking within a VE, modulators did produce effects. The impact of this is to open up the possibility of either refining these modulators, or finding others, with goal of amplifying effects to the point of usefulness in creating a robust and consistent affective tool.

As noted in the section describing the scope of this thesis there will be some impact on visual masking research in general (see section 1.5.2). Predominantly, the issue of depth disparity has been opened up. The internal depth disparity of non-planar targets has been shown not to disrupt the masking process within a VE. It remains to be seen if there are limits to this, in particular with the kind of depth produced by stereoscopic displays: as noted in the Methodology (section 3.1), this is not the same category of depth as that produced by viewing the real world. In addition, it has been demonstrated that there are limits to the disruption of masking from external disparity between target and mask. Parts of the non-planar targets sat on approximately the same depth plane as the planar masks whilst other parts clearly did not. Regarding this, it remains to be seen whether the brain is using the continuity of the target surface to prevent masking disruption to the more distant parts from the mask.

This thesis may also have an impact in affective cognition research. Aside from adding to the already considerable evidence for actions being driven by implicit processes, support has been given to the idea implicit processes can affect explicit learning. Experiment 6 demonstrated an effect whereby an implicit, learned bias towards positive facial manifested itself in the speed of learning of a similar, explicit bias. Of interest, is the fact that both the implicit VE task and the explicit measure carry essentially the same affective meaning, i.e. a series of valent facial expressions, but are different in form: the former used non-planar, target objects embedded within a VE while the latter used two-dimensional images presented in a simple context (a black background). It is assumed, therefore, that any effect that the former task had on the latter was due to the similarities between tasks. In other words, a bias had been established implicitly and was subsequently demonstrated explicitly.

Although Experiment 5 found no significant expression or environment effects with the inventories, it nevertheless produced promising results with the Emotional Stroop test. Here, there was a significant interaction between these two factors. This appears to indicate the possibility of a congruency effect where affective cognitions are affected more when facial expression matches context. It is suggested that this warrants further investigation. Further investigation is also warranted following Experiment 6. A cognitive bias effect was observed yet within this there may have been a congruency effect of the type observed in Experiment 5.

10.2.2 Wider Application to Psychotherapy

One of the main aims of this thesis has been to demonstrate that visual masking in a VE could provide the basis for a psychotherapeutic tool. As changes are wrought implicitly, explicit resistance to changes by a patient should be ameliorated. For instance, it is possible to embed the masked objects within a gaming VE of flexible form that can be tailored to fit a specific patient profile, thereby, garnering increased patient engagement, lowering overall treatment dropout rates.

Being able to alter background affect, either directly or indirectly via cognitive modification, could serve as a useful adjunct to, or replacement for, some forms of psychopharmacological intervention. This may be useful for those patients who are unwilling to take drugs or derive no benefit from the pharmacology available. Additionally, there appears to be no reason why such a VE tool could not be used alongside drugs in an initial intervention strategy. Several affect conditions may benefit from such complex intervention strategies. Social anxiety disorder is one such candidate but there are many others.

VEs containing visually masked stimuli may also be helpful during the body of an intervention when a patient may be undergoing some form of psychotherapy such as CBT. Regular use of the VE to stabilise mood could prove supportive for the therapeutic work. It would allow the patient and therapist to concentrate on tackling the cognitive precursors to negative affect without affective disruption and thus improve the longer-term prognosis.

Besides being a useful adjunct to initial therapy strategies, such VEs may also be useful in stabilising mood as a patient weans himself or herself off from a particular drug, whether the dependency is psychological in nature, chemical or both.

Beyond psychotherapeutic use, it should be possible to adapt the tool for a range of uses including training simulations, entertainment and so on. Affect induction could prove useful if, for example, a trainee for a particular role needed to demonstrate a capability to carry out duties under stressful conditions. With current simulators, stressors are generally explicit even if some are less so than others such as background audio designed to enhance certain moods. Trainees are therefore potentially aware of the ‘mechanics’ of the training VE to an extent and this may influence the production of behaviour not encountered in reality. It has already been noted that VE simulations of real world event can engender a dampened form of behaviour normally encountered (see section 2.2.3.1). Visually masked targets embedded within a training VE may be able to counter this confound by introducing implicit stressors that could go some way to offsetting the influence of a trainee’s cognition that the VE was not real.

Yet another example lies with the induction of mood to enhance a gaming experience. Applying implicit induction to match that intended by explicit constructs at various stages of a gaming narrative could serve to amplify the satisfaction of the overall experience. Having such a device could certainly confer an advantage to a particular game in the modern, competitive market, especially if used as a marketing tool.

The ethics of the two previous examples would need careful consideration. For instance, with training simulators, providing trainees with prior knowledge of the experience they were about to undertake could be considered an ethical approach to take but may be thought to undermine any advantage gained by implicit affect induction. However, as some of the work in this thesis has shown, even possessing some knowledge of the stimuli does not appear to detract from their efficacy completely. For example, in the first experiment (see Chapter 4) participants were aware that they were to look for masked objects in the doorways. In this way, some foreknowledge could be imparted prior to a simulation session, enough to satisfy ethical considerations but not enough to eliminate any effects. Future work could further examine the effects of foreknowledge of the masked objects on results as some interesting implications arise. For example, if explicit foreknowledge has little effect on implicit processes, how does this sit with the results of Experiment 6 where *implicit* processes appeared to have an effect on *explicit* processes?

Nevertheless, even taking account of the ethical complexities, implicit mood induction via visually masked objects within VE opens up exciting possibilities for the future.

10.3 Future Work

Various possibilities for future work have been noted in the previous sections. These will be taken up here, alongside additional suggestions for further research. The ideas can be grouped into three categories: perception, higher cognition and affect.

10.3.1 Perception Questions

Extending the exploration of target mask geometry, with regard to internal depth disparity, may prove fruitful. The current work used planar masks with non-planar targets. There remains, therefore possibilities of further combinations that involve non-planar masks, alternated with planar and non-planar targets. A cogent aim would be to see any of these further combinations had a disruptive effect on visual masking within a VE.

Related to the above, further exploration of external depth disparity, between target and mask, could also be investigated. Given any one of the above combinations, at what point does visual masking break down, within a VE, as external depth disparity increases? Are there relative factors involved?

The surface continuity of virtual objects could also prove a useful area to investigate. Supposing that either or both mask and target are non-planar, does increasing either or both internal depth disparity values disrupt the masking process? If it is assumed that it does, then what are the conditions under which this occurs?

Could eye dominance effects be lessening any masking effects? As noted, the stereoscopic system used for the current thesis generates a 60Hz display by alternating blank frames with image frames at 120Hz for each eye. When one eye is seeing a blank frame, the other is seeing an image frame and vice versa. As each eye sees a slightly different image for each of the 60Hz frames, eye dominance could be affecting the masking. For example, if the right eye of a right eye dominant participant is seeing a blank frame then that may take processing precedence over the left eye image frame. Effectively this could dampen any effect, behavioural or otherwise, produced by particular targets because, as Experiment 4 seems to indicate, the number of exposure times is an important factor in the influence of targets on participants. Investigating ways around the problem of eye dominance may serve to increase future effect sizes. One solution could be to source systems that project full displays to each eye separately, without blanking, and that are capable of the frame rates required for visual masking.

Experiment 4 supports the idea that exposure time to target is important. Therefore, researching other factors that might affect exposure, such as participant eye saccades or scan paths, could also bring benefits. This may necessitate some form of eye tracking in future work if we were to allow participants an equivalent freedom of movement to that which they experienced in the experiments.

Another direction leading from this current work could be to look more closely at the geometry of the VE. Participants had the freedom to approach the masking doors at a range of angles but this was not used as a factor for the analysis. Approach angle may alter the visibility or other effects of a target. Also, within a VE objects appear to scale-up in size as they are approached (measured by the angle subtended on the retina), simulating real-world perspective. There may even be an interaction between approach angle and mask/target scale on a dependant variable.

10.3.2 Higher Cognition Questions

Experiment 4 showed dynamic targets appearing to be implicitly read as series of separate images and not as a series of animated frames. To support this hypothesis, a neuro-imaging study could place markers in places of expected increased activity in brain areas associated with biological motion. An absence of activity when presented with dynamic, masked stimuli would support the hypothesis whereas observing activity could suggest a more complex picture, e.g. dynamic targets may be initially processed as biological motion but this signal is then dampened by later processes.

Varying levels of immersion may serve to influence the strength, or indeed existence, of the effects of target form on behaviour, cognition and affect. Sourcing a HMD capable of the high frame rates demanded by visual masking would be a practical consideration here.

Related to the above, varying levels of Psi and PI may vary the efficacy of the visual masking effects observed from the current work. One hypothesis to test would be an assumption made within this thesis, namely that Psi is more important for target form to influence behaviour, cognition and affect than PI.

10.3.3 Affective Processing Questions

Experiment 6 suggests several possibilities for further work. Two different instruments could be used, as alternatives to the explicit bias modification test. Testing for affect change with an emphasis on social processes could yield some results. For example, social anxiety inventories may help to expose any changes in social cognition that have occurred because of exposure to the stimuli. Similarly, Implicit Association tests might reveal changes to a participant's implicit self-esteem.

One issue not covered by the current thesis is that of longevity of effect. Recidivism is an important issue in affect intervention work and has been the basis of on-going criticism for various therapies including psychopharmacology. It can be argued that a therapeutic adjunct does not need to be effective in the long term, merely that it effectively facilitates or supports any therapy it runs alongside (which itself would need to be efficacious in the long term). However, counter to this it could be said that any combination of therapies needs to produce a long term effect as a whole and that if one or more constituent parts are not contributing to this in some way, or even working to the detriment of this goal, then the use of those parts needs to be questioned. If visual masking in a VE is to form the basis of a psychotherapeutically useful tool then a good starting point would surely be to see if the effects found within this thesis are capable of acting over a sustained period of time. It should be straightforward to test initially how long a single session exposure lasts. Following this, work could be undertaken to look at extending this time. It is suggested that this could be tested by using some form of exposure regime, run over several months, where participants regularly encounter the stimuli. In addition, looking at how these stimuli affect more traditional therapy programmes, such as CBT, would be yet another direction to take. This latter approach is perhaps the most important as it is not envisaged that the stimuli could constitute a standalone intervention of and in themselves in all but the mildest cases of affect disorder.

Further work on the congruency effects seen in Experiment 5 may prove fruitful. Neuro-imaging would perhaps be a useful option to see if activity were increased in the absence or presence of visually masked stimuli either matching the valence of explicit context or displaying opposing valence. Variation could be introduced by ranging masked expression from negative, via neutral, to positive and similarly with the explicit context. A goal here would be to look for combinations of values that enhanced effects.

Finally, how does foreknowledge of the VE masking stimuli affect the results? Various levels of foreknowledge could be tested, forming factor levels in an independent variable. For the experiments in this thesis, participants had little foreknowledge of masking although pilot trials, perhaps surprisingly, showed that such knowledge did not necessarily extinguish effects. If a tool along the lines proposed were to be taken up beyond the laboratory then it could not be assumed that future potential patients would have no knowledge of how it worked. It is therefore important that the effects of foreknowledge be known early on.

11 Thesis Validation

11.1 Construct validity

The constructs assumed by this research have been matched to the measures. For example, the affect measurement inventories used in Experiment 5 were standard forms whose validity has been tested by prior research.

In addition, it is understood that some constructs are not necessarily exposed through a single measure. Measurement of more than one modality has been used throughout this thesis. For instance, behavioural measures, cognitive/affective tests and inventories have been used where appropriate.

11.2 Internal validity

Care has been taken to construct appropriate controls for all of the experiments carried out. For example, thought was been given as to how mood induction experiments can be bookended by inventories without introducing confounding variance. Other mood induction studies do not always consider this as an issue.

Using faces to induce emotion or affect cognition, it can be argued, offers greater validity than using other methods such as music video etc. Faces and expression recognition appears early on, developmentally speaking, for humans whereas other modes used to induce moods tend to rely on later, learnt contexts. We can therefore claim there is something profound and fundamental about using faces as emotional stimuli, eclipsing validity claims made for other stimulus types. Using non-planar virtual expressions within an interactive, stereoscopic VE further improves validity over planar images shown in a non-interactive context.

11.3 External validation

Experiments 1 & 2:

Drummond, J., Berthouze, N. & Steed, A.J. (2011). Affective Reactions to Visually Masked Stimuli within a Virtual Environment. *Annual Review of Cybertherapy and Telemedicine*, p116.

The above material was also the subject of a talk given at the CyberTherapy 16 conference, Canada 2011.

Experiment 4:

Drummond, J. & Steed, A.J. (2011) Dynamic Visual Masking in a Virtual Environments, *Perceptual Illusions in Virtual Reality proceedings*, IEEE VR2011.

Experiments 3 & 4:

Drummond, J. & Steed, A.J. (2013) Visual Masking Parameters for Virtual Environments, *conference proceedings*, IEEE VR2013.

11.4 Conclusion validity

The conclusions are supported by the statistical analysis of experimental data, using standard methods, both descriptive and inferential. Section 3.6 briefly lists the inferential forms of analysis used within this thesis. It can be seen that they are all of a type practiced by researchers for a long time

For further detail of the statistical analysis used in each of the experiments, please see the individual experiment chapters.

Interpretations of the experimental results also rest upon the conclusions and assumptions drawn from the Literature Review, Chapter 2, which deals with prior research in the subject areas underpinning this thesis.

12 References

Adolphs, R., Tranel, D., Damasio, H. & Damasio, A. (1994). Impaired Recognition of Emotion in Facial Expressions Following Bilateral Damage to the Human Amygdala. *Nature*, 372, pp. 669-672.

Alcañiz, M., Baños, R., Botella, C. & Rey, B. (2003). The EMMA Project: Emotions as a Determinant of Presence. *PsychNology Journal*, 1, 2, pp.141 - 150.

Ambadar, Z., Schooler, J.W. & Cohn, J.F. (2005). Deciphering the Enigmatic Face: The Importance of Facial Dynamics in Interpreting Subtle Facial Expressions. *Psychological Science*, 16, 5, pp. 403-410.

Anderson, P., Rothbaum, B. & Hodges, L. (2003). Virtual reality in the treatment of social anxiety: Two case reports, *Cognitive and Behavioral Practice*, 10**,** 3, pp. 240-247.

Andréasson, P. & Dimberg, U. (2008). Emotional Empathy and Facial Feedback. *Journal of Nonverbal Behavior*, 32, 4, pp. 215-224.

Ansorge, U., Francis, G., Herzog, M.H. & Ogmen, H. (2007). Visual masking and the dynamics of human perception, cognition, and consciousness: A century of progress, a contemporary synthesis, and future directions, *Advances in Cognitive psychology*, volume 3, no 1-2, 1-8.

Ansorge, U., Klotz, W. & Neumann, O. (1998). Manual and verbal responses to completely masked (unreportable) stimuli: Exploring some conditions for the metacontrast dissociation, *Perception*, 27, pp. 1177-1189.

Baccus, JR., Baldwin, MW. & Packer, DJ. (2004). Increasing Implicit Self-Esteem through Classical Conditioning. *Psychological Science,* 15, 7, pp. 498-502.

Bachorowski, J. & Owren, M.J. (2008). Vocal Expressions of Emotion, in Lewis, M., Haviland-Jones, J.M. & Feldman Barrett, L. (Eds.) *Handbook of Emotions*, The Guilford Press, New York.

Bahrami, B., Lavie, N. & Rees, G. (2007). Attentional Load Modulates Responses of Human Primary Visual Cortex to Invisible Stimuli, Current Biology, 17, 6, pp. 509-513.

Bailenson, J., Yee, N., Merget, D. & Schroeder, R. (2006). The Effect of Behavioral Realism and Form Realism of Real-Time Avatar Faces on Verbal Disclosure, Nonverbal Disclosure, Emotion Recognition, and Copresence in Dyadic Interaction. *Presence: Teleoperators and Virtual Environments*, 15, 4.

Baldwin, MW., & Baccus, JR. (2003). Relational Knowledge and Expectancy Value Approach to Self-esteem. In Spencer, S. & Fein, S. (Eds.) *Motivated social perception: The Ninth Ontario Symposium*. Available from http://selfesteemgames.mcgill.ca/resources/Baldwin_Baccus2004.pdf (accessed Sept 2007).

Baños, R., García-Soriano, G., Botella, C., Oliver, E., Etchemendy, E., Bretón, J., & Alcañiz, M. (2009). Positive Mood Induction and Well Being, *Proceedings Human System Interaction*, Catania.

Baños, R., Liaño, V., Botella, C., Alcañiz, M., Guerrero, B. & Rey, B. (2006). Changing Induced Moods Via Virtual Reality, in IJsselsteijn, W.A., De Kort, Y.A.W., Midden, C., Eggen, B. & Van den Hoven, E. (Eds.), *Persuasive 2006*, LNCS 3962, pp. 7–15.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition, *Journal of Cognitive Neuroscience* , **15**(4), 600-609.

Barbot, A. & Kouider, S. (2011) Longer is not better: nonconscious overstimulation reverses priming influences under interocular suppression, *Attention & Perception Psychophysics*, DOI 10.3758/s13414-011-0226-3.

Baron-Cohen, S., Wheelwright, S. & Joliffe, T. (1997). Is there a 'language of the eyes'? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition*, **4**, pp. 311-331.

Bartlett, F.C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press UK.

Baumeister,RF., Campbell, JD., Krueger, JI., & Vohs, KD. (2005). Exploding the Self-Esteem Myth. *Scientific American*, January edition.

Beard, C., Weisberg, R.B. & Amir, N. (2011). Combined Cognitive Bias Modification Treatment for Social Anxiety Disorder: A Pilot Trial. *Depression and Anxiety*, 28, 11, pp. 981–988.

Beck, A.T. (1967). *The Diagnosis and Management of Depression*. University of Pennsylvania Press, USA.

Begeer, S., Rieffe, C., Terwogt, M.M. & Stockmann, L. (2006). Attention to facial emotion expressions in children with autism. *Autism*, **10,** 1, pp. 37-51.

Bentall, R.P. (2003). *Madness Explained*. Penguin, London.

Bevan, W., Jonides, J. & Collyer, S.C. (1970). Chromatic relationships in metacontrast suppression, *Psychonomic Science*, 19, pp. 367-368.

Bohner, G., Crow, K., Erb H.P., & Schwarz, N. (1992). Affect and Persuasion: Mood Effects on the Processing of Message Content and Context Cues and on Subsequent Behavior. *European Journal of Social Psychology*, 22, 6, pp. 511-530.

Bonneh,Y. S.,Cooperman , A. & Sagi,D. (2001). Motion-induced blindness in normal observers, *Nature*, 411, pp. 798–801.

Bower, G.H., Gilligan, S.G. & Monteiro, K.P. (1981). Selectivity of Learning Caused by Affective States. *Journal of Experimental Psychology: General*, 110, 4, pp. 451-473.

Bradley, M.M., & Lang, P.J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.

Breitmeyer, B. G. (1984). *Visual masking: An integrative approach*, Oxford University Press, UK.

Breitmeyer B.G., Ehrenstein, A., Pritchard, K., Hiscock, M. & Cristan, J. (1999). The Roles of Location Specificity and Masking Mechanisms in the Attentional Blink, *Perception Psychophysics*, 61, pp. 798-809.

Breitmeyer B.G. & Öğmen, H. (2006). *Visual Masking: Time Slices Through Conscious and Unconscious Vision,* Oxford University Press, UK

Breitmeyer B.G., Öğmen, H. & Chen, J. (2004a). Unconscious priming by color and form: Different processes and levels, *Consciousness and Cognition*, 13, pp. 138–157.

Breitmeyer, B. G., Ro, T. & Singhal, N. (2004b). Unconscious priming with chromatic stimuli occurs at stimulus- not percept-dependent levels of visual processing, *Psychological Science*, 15, pp. 198-202.

Brewer, D., Doughtie, E.B. & Lubin, B. (1980). Induction of Mood and Mood Shift. *The Journal of Clinical Psychology*, 36, pp. 215-26.

Brewin, C. (2001) A Cognitive Neuroscience Account of PTSD and its Treatment. *Behaviour Research and Therapy*, 39, pp. 373-393.

Britton, JC., Shin, LM., Feldman Barrett, L., Rauch, SL., & Wright, CI. (2008). Amygdala and Fusiform Gyrus Temporal Dynamics: Responses to Negative Facial Expressions. *BMC Neuroscience*, 9, 44.

Browning, M., Holmes, E.A & Harmer, C.J. (2010). The Modification of Attentional Bias to Emotional Information: A Review of the Techniques, Mechanisms, and Relevance to Emotional Disorders. *Cognitive, Affective & Behavioural Neuroscience*, 10, 1, pp. 8–20.

Bruce, V. & Young, A. W. (1986). Understanding Face Recognition. *British Journal of Psychology*, **77**, pp. 305–327.

Brunicardi, F.C., Shavelle, D.M. & Andersen, D.K (1995). Neural regulation of the endocrine pancreas, *International Journal of Gastrointestinal Cancer*, Volume 18, Number 3, pp. 177-195.

Bunzeck, N. & Düzel, E. (2006). Absolute Coding of Stimulus Novelty in the Human Substantia Nigra/VTA, *Neuron*, 51, 369–379.

Burns, E., Razzaque, S., Panter, A.T., Whitton, M.C., McCallus, M.R. & Brooks, F.P. Jr. (2005). The Hand Is More Easily Fooled than the Eye: Users Are More Sensitive to Visual Interpenetration than to Visual-Proprioceptive Discrepancy, *Presence*, 15, 1, pp. 1-15.

Burton, A.M., Bruce, V. & Johnston, R.A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*, 81, 3, pp. 361-380.

Buss, DM. (2004). *Evolutionary Psychology: the New Science of the Mind*, Pearson Education Inc.

Cacioppo, J.T., Petty, R.E., Losch, M.E. & Kim, H.S. (1986). Electromyographic Activity over Facial Muscle Regions can Differentiate the Valence and Intensity of Affective Reactions. *Journal of Personality and Social Psychology*, 50, 2, pp. 260-268.

Calder, A.J. & Young, A.W. (2005). Understanding The Recognition Of Facial Identity And Facial Expression. *Nature Reviews: Neuroscience*, 6, pp. 641-651.

Calder, A.J., Burton, A.M., Miller, P., Young, A.W. & Akamatsu, S. (2001). A Principal Component Analysis of Facial Expressions. *Vision research*, 41, pp. 1179-1208.

Capps, L., Yirmiya' N. & Sigman, M. (2006). Understanding of Simple and Complex Emotions in Non-retarded Children with Autism. *Journal of Child Psychology and Psychiatry*, **33**, 7, pp. 1169-1182.

Carlin, A.S., Hoffman, H.G., Weghorst, S. (1997). Virtual reality and tactile augmentation in the treatment of spider phobia: a case report, *Behaviour Research and Therapy*, 35, 2.

Causse, M., Pavard, B., Sénard, JM., Démonet, JF., Pastor, J. (2007). Emotion Induction through Virtual Avatars and its Impact on Reasoning: Evidence from Autonomous Nervous System Measurements and Cognitive Assessment. *VRIC'07*, Laval, France.

Chen, M. & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, 25, 2, pp. 215-224.

Clark, D.M. (2001). A Cognitive Perspective on Social Phobia, in Crozier, W.R. & Alden L.E. (Eds.), *International Handbook of Social Anxiety: Concepts, Research and Interventions Relating to the Self and Shyness*, John Wiley & Sons Ltd, UK.

Clark, T.F., Winkielman, P. & McIntosh D.N. (2008). Autism and the Extraction of Emotion From Briefly Presented Facial Expressions: Stumbling at the First Step of Empathy. *Emotion*, 8, 6, pp. 803-809.

Conway, C.A., Jones, B.C., Debruine, L.M., Little, A.C., Hay, J., Welling, L.L.M., Perrett, D.I. & Feinberg, D.R. (2007). Integrating physical and social cues when forming face preferences: Differences among low and high-anxiety individuals. *Social Neuroscience*, 1, pp. 1-7.

CORE (2009). http://coreims-online.co.uk/about_core_core_system.htm (accessed November 2009)

Coyle (2009). https://www.cs.tcd.ie/David.Coyle/playWrite.htm (accessed September 2009)

Coyle, D., Doherty, G. & Sharry, J. (2009). An Evaluation of a Solution Focused Computer Game in Adolescent Interventions, *Clinical Child Psychology and Psychiatry*, 14, 3.

Csikszentmihalyi, M. (1991). *Flow: The psychology of optimal experience.* Harper & Row, New York.

Damasio, A.R. (1994). *Descartes' Error: emotion, reason, and the human brain*, Grosset/Putnam USA.

Damasio, A.R., Damasio, H., Van Hoesen, G.W. (1982). Prosopagnosia: anatomic basis and behavioral mechanisms. *Neurology*, 32, pp. 331-341.

Dandeneau, SD., Baldwin, MW. (2004). The Inhibition of Socially Rejecting Information Among People with High Versus Low Self-esteem: the Role of Attentional Bias and the Effects of Bias Reduction Training. *Journal of Social and Clinical Psychology*, 23, 4, pp. 584-602.

Dandeneau, SD., Baldwin, MW., Baccus, JR., Sakellaropoulo, M., & Pruessner, JC. (2007). Cutting Stress Off at the Pass: Reducing Vigilance and Responsiveness to Social Threat by Manipulating Attention. *Journal of Personality and Social Psychology*, 93, pp. 651-666.

Darwin, C. (1899 edition) *The Expression of Emotion in Man and Animals* (Authorised Edition), D.Appleton & Co., New York.

De Haan, E.H.F., Young, A.W. & Newcombe, F. (1991). Covert and Overt Recognition in Prosopagnosia. *Brain*, 114, pp. 2575-2591.

De Houwer, J. (2007). A Conceptual and Theoretical Analysis of Evaluative Conditioning. *The Spanish Journal of Psychology*, 10, 2, pp. 230-241.

De Kort, YAW. & Ijsselsteijn, WA. (2006). Reality Check: The Role of Realism in Stress Reduction Using Media Technology. *CyberPsychology & Behavior*, 9, 2, pp. 230-233.

De Kort, YWA., Meijnders, AL., Sponselee, AAG., IJsselsteijn, WA. (2006). What’s wrong with virtual trees? Restoring from stress in a mediated environment. *Journal of Environmental Psychology*, 26, pp. 309-320.

De la Rosa, S., Schneider, B. & Moraglia, G. (2008). Binocular unmasking is important for the detection, categorization and identification of noise-masked real-life objects. *International Journal of Psychology,* 29, 3-4, pp. 766.

Detenber, B.H., Simons R.F., & Bennett G.G., Jr. (1998). Roll 'Em!: The Effects of Picture Motion on Emotional Responses. *Journal of Broadcasting and Electronic Media*, 42, 1, pp. 113-127.

Difede, J., Hoffman, H., Jaysinghe, N. (2002). Innovative use of virtual reality technology in the treatment of PTSD in the aftermath of September 11. *Journal of Psychiatric Services*, Multimedia Reviews*,* 53(9): 1083-1085.

Dijksterhuis, A. (2004). I Like Myself but I Don't Know Why: Enhancing Implicit Self-Esteem by Subliminal Evaluative Conditioning. *Journal of Personality and Social Psychology*, 86, 2, pp. 345–355.

Dijksterhuis, A., Aarts, H., & Smith, P. K. (2005). The power of the subliminal: Subliminal perception and possible applications. In R. Hassin, J. Uleman, & J. A. Bargh (Eds.), *The New Unconscious*, Oxford University Press, New York.

Dimberg, U., Thunberg, M. & Elmehed, K. (2000). Unconscious Facial Reactions to Emotional Facial Expressions. *Psychological Science*, 11, 1, pp. 86-89.

Dolan, R.J. (2002). Emotion, Cognition, and Behavior. *Science,* **298**, 1191.

DSM-IV-TR (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th Edition, Text Revision) the American Psychiatric Association.

Düzel, E., Bunzeck, N., Guitart-Masip, M. & Duzel, S. (2010). NOvelty-related Motivation of Anticipation and exploration by Dopamine (NOMAD): Implications for healthy aging. *Neuroscience and Biobehavioral Reviews*, 34, 660–669.

Dumoulin, S.O. & Hess, R.F. (2006). Modulation of V1 Activity by Shape: Image-Statistics or Shape-Based Perception? *Journal of Neurophysiology*, 95, pp. 3654–3664.

Dyck, M., Loughead, J., Kellermann, T., Boers, F., Gur, R.C. & Mathiak, K. (2011). Cognitive versus automatic mechanisms of mood induction differentially activate left and right amygdala. *Neuroimage*, 54, pp. 2503-2513.

Ekman, P. (1972). Universals And Cultural Differences In Facial Expressions Of Emotions. In J. Cole (ed.), *Nebraska Symposium on Motivation*, 1971. Lincoln, Neb.: University of Nebraska Press, pp. 207- 283.

Ekman, P (2009). http://www.face-and-emotion.com/dataface/facs/description.jsp (accessed November 2009).

Ekman, P. & Freisen, W.V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, 32, 1, pp. 88-106.

Ekman, P. & Rosenberg, E.L. (2005). *What the Face Reveals* (2nd edition), Oxford University Press, US.

Enns, J.T. (2000). Visual Binding in the Standing Wave Illusion, *Psychonomic Bulletin and Review*, 9, pp. 489-496.

Enns, J.T. & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, 4, 9, pp. 345-352.

Farnham, SD., Greenwald, AG., & Banaji, MR. (1999). Implicit Self Esteem. In D. Abrams & M. Hogg (Eds.), *Social Identity and Social Cognition* (pp. 230-248). Oxford, UK:Blackwell.

Fiedler, K. (1990). Mood-dependent Selectivity in Social Cognition. *European Review of Social Psychology*, 1, pp. 1-32.

Ferrer-Garcia, M. & Gutiérrez-Maldonado, J. (2009). Anxiety and depression responses to virtual environments in patients with eating disorders and controls, *RAVE 2009*, Barcelona.

Ferrer-García, M., Gutiérrez-Maldonado, J. & Letosa-Porta, A. (2005). Body Image Assessment Software: A new program for assessing body-image disturbance using adjustable partial image distortion*, European Council on Eating Disorders (ECED),* Innsbruck 2005.

Ferster, C.B., & Skinner, B.F. (1957). *Schedules of reinforcement*, Appleton-Century-Crofts, New York.

Ferwerer, J.A., Pattanaik, S.N., Shirley, P. & Greenberg, D.P. (1997). A Model of Visual Masking for Computer Graphics, *International Conference on Computer Graphics and Interactive Techniques*, Proceedings of the 24th annual conference on Computer graphics and interactive techniques, 1997, Pages: 143 - 152.

Finkbeiner, M. & Palermo, R. (2009). The Role of Spatial Attention in Nonconscious Processing: A Comparison of Face and Nonface Stimuli. *Psychological Science*, 20, 1, pp. 42-51.

Flake, G.W. (1998). The Computational Beauty of Nature, A Bradford Book/MIT Press, UK/USA.

Forgas, J.P. (1999). On Feeling Good and Being Rude: Affective Influences on Language Use and Request Formulations. *Journal of Personality and Social Psychology*, **76**, 6, pp. 928-939.

Foster, D.H (1976). Rod-cone Interaction in the After-flash effect, *Vision research*, 16, pp. 393-396.

Fox, E. (2008). *Emotion Science*, Palgrave Macmillan UK.

Fox, J., Bailenson, J. & Binney, J. (2009). The Virtual Experiences, Physical Behaviours: The Effect of Presence on Imitation of an Eating Avatar, *Presence*, Vol. 18, No. 4., 294–303.

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 & Virtual Environments*, 8, 1, pp. 1-13.

Gallagher, S. (2007). Simulation Trouble. *Social Neuroscience,* 2, 3, pp. 353-365.

Garau, M., Slater, M., Pertaub, D.P. & Razzaque, S. (2005). The responses of people to virtual humans in an immersive virtual environment Presence-Teleoperators and Virtual Environments 14: 1. 104-116.

Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A, Steed, A. & Sasse, M.A. (2003). The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 529-536.

Gardner, P. & Oei, TPS. (1981). Depression and Self-esteem: An Investigation that used Behavioural and Cognitive Approaches to the Treatment of Clinically Depressed Clients. *Journal of Clinical Psychology*, 37, 1, pp.128-135.

Gasper, K. & Clore, G. (2002). Attending to the Big Picture: Mood and Global Versus Local Processing of Visual Information. *Psychological Science*, **13**, 1, pp. 34-40.

Gilboa-Schechtman, E., Revelle, E.W. & Gotlib, I.H. (2000). Stroop Interference following Mood Induction: Emotionality, Mood Congruence, and Concern Relevance, *Cognitive Therapy and Research*, 24, 5, pp. 491-502.

Gilet, A.L. (2008). Procédures d'Induction d'Humeurs en Laboratoire: Une Revue Critique, *L'Encéphale*, 34, 233-239.

Glantz, K., Durlach, N.I., Barnett, R.C. & Aviles, W.A. (1996). Virtual reality (VR) for psychotherapy: From the physical to the social environment. *Psychotherapy: Theory, Research, Practice, Training,* 33, 3, pp. 464-473.

Gasper, K. & Clore, G. (2002). Attending to the Big Picture: Mood and Global Versus Local Processing of Visual Information, *Psychological Science*, **13**(1), 34-40.

Gauthier, I. & Tarr, M.J. (1997). Becoming a Greeble Expert: Exploring Mechanisms for Face Recognition. *Vision Research*, 37, pp. 1673-1682.

Giesbrecht, B. & Di Lollo, V. (1998). Beyond the Attentional Blink: Visual Masking by Object Substitution, Journal of Experimental Psychology: Human Perception and Performance, vol. 24, No. 5,1454-1466.

Gorini, A., Griez, E., Petrova, A. & Riva, G. (2010). Assessement of the emotional responses produced by exposure to real food, virtual food and photographs of food in patients affected by eating disorders, *Annals of General Psychiatry,* 9:30.

Grazebrook, K., Garland, A. & the Board of British Association for Behavioural and Cognitive Psychotherapies (2005). What Is CBT? http://www.babcp.com/silo/files/what-is-cbt.pdf (accessed October 2010).

Greenwald, A.G. & Farnham, S.D. (2000). Using the Implicit Association Test to Measure Self-Esteem and Self-Concept, *Journal of Personality and Social Psychology*, vol.79, no.6, 1022-1038

Grill-Spector, K. (2003). The neural basis of object perception, *Current Opinion in Neurobiology*, 13, pp. 1–8.

Gross, T.F (2004). The Perception of Four Basic Emotions in Human and Nonhuman Faces by ChildrenWith Autism and Other Developmental Disabilities, *Journal of Abnormal Child Psychology*, **32**, 5, pp. 469–480.

Gross, T.F. (2008). Recognition of Immaturity and Emotional Expressions in Blended Faces by Children with Autism and Other Developmental Disabilities. *Journal of Autism and Developmental Disorders*, **38**, pp. 297-311.

Grossman , R.B. & Tager-Flusberg, H. (2008). Reading faces for information about words and emotions in adolescents with autism. *Research in Autism Spectrum Disorders*, **2**, pp. 681-695.

Gur, R. C, & Sackeim, H. A., (1979). Self-deception: A concept in search of a phenomenon. *Journal of Personality and Social Psychology*, 37, 147-169.

Habel, U., Klein, M., Kellermann, T., Shah, N.J., & Schneider, F. (2005). Same or Different? Neural Correlates of Happy and Sad Mood in Healthy Males. *Neuroimage*, 26, pp. 206-214.

Hammen, C. & Watkins, E. (2008). *Depression* (2nd edition), Psychology Press, UK.

Hatfield, E., Cacioppo, J.T. & Rapson, RL. (1994). *Emotional contagion*, New York: Cambridge University Press.

Haxby, J. V., Hoffman, E. A. & Gobbini, M. I. (2000). The Distributed Human Neural System for Face Perception. *Trends in Cognitive Science*, **4**, pp. 223–233.

Hermens, F. & Ernst, U. (2007). Visual Backward Masking: Modeling Spatial and Temporal Aspects. *Advances in Cognitive Psychology***,** 3, 1-2, pp. 93-105.

Hermens, F., Luksys, G., Gerstner, W., Herzog, M.H. & Ernst, U. (2008). Modeling Spatial and Temporal Aspects of Visual Backward Masking, Psychological Review, 115, 1, pp. 83–100.

Hertel, P.T. & Mathews, A. (2011). Cognitive Bias Modification: Past Perspectives, Current Findings, and Future Applications. *Psychology Faculty Research*, Paper 14, http://digitalcommons.trinity.edu/psych_faculty/14 (accessed 23/09/2011).

Herzog, M.H. (2007). Spatial Processing and Visual Backward Masking. *Advances in Cognitive Psychology*, 3, 1-2, pp. 85-92.

Hill, SE. & Buss, DM. (2006). The Evolution of Self-Esteem. In Kernis, M., (Ed.), *Self Esteem: Issues and Answers: A Sourcebook of Current Perspectives,* Psychology Press, New York, pp. 328-333.

Hoffman, H.G., Patterson, D.R., Carrougher, G.J. & Sharar, S.R. (2001). Effectiveness of Virtual Reality-Based Pain Control With Multiple Treatments, *The Clinical Journal of Pain*, 17, 3.

Holender, D. (1986). Conceptual, Experimental, and Theoretical Indeterminacies in Research on Semantic Activation without Conscious Identification, *Behavioral & Brain Sciences*, 9, pp. 50-66.

Holender, D. & Duscherer, K. (2004). Unconscious Perception: The Need for a Paradigm Shift, *Perception & Psychophysics*, 66, 5, pp. 872-881.

Holmes E.A., James E.L., Coode-Bate T., Deeprose C. (2009). Can Playing the Computer Game ‘‘Tetris’’ Reduce the Build-Up of Flashbacks for Trauma? A Proposal from Cognitive Science. *PLoS ONE*, 4, 1, e4153.

Humphreys, G.W., Donnelly, N. & Riddoch, M.J. (1993). Expression is Computed Separately from Facial Identity, and it is Computed Separately for Moving and Static Faces: Neuropsychological Evidence. *Neuropsychologia*, 31, 2, pp. 173-181.

Ijsselsteijn, W.A., De Ridder, H., Freeman, J., Avons, S.E. & Bouwhuis, D. (2001). Effects of Stereoscopic Presentation, Image Motion and Screen Size on Subjective and Objective Corroborative Measures of Presence. *Presence: Teleoperators & Virtual Environments*, 10, 3, pp. 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, 4, pp. 207-214.

Iwamoto, T., Tatezono, M, and Shinoda, H. (2008). Non-Contact Method for Producing Tactile Sensation Using Airborne Ultrasound, *Proc. EuroHaptics 2008*, LNCS 5024, pp. 504-513.

Jacoby, L.L & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, 118, pp. 126-135.

James, L.K., Lin, C. Y., Steed, A., Swapp, D. & Slater, M. (2003). Social Anxiety in Virtual Environments: Results of a Pilot Study. *CyberPsychology & Behavior, 6, 3*, pp. 237-243.

Johnson, M.H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 90, pp. 1-19.

Juruena , M.F., Giampietro, V.P., Smith , S.D., Surguladze , S.A., Dalton , J.A., Benson , P.J., Cleare , A.J. & Fu, C.H.Y. (2010). Amygdala activation to masked happy facial expressions. *Journal of the International Neuropsychological Society*, 16, pp. 1-5.

Kahneman, D. (1964). Temporal Summation in an Acuity Task at Different Energy Levels – a Study of the Determinants of Summation, *Vision Research*, 4, pp. 557-566.

Kahneman, D. (1967). An onset-onset law for one case of apparent motion and metacontrast, *Perception and Psychophysics*, 2, pp. 577–584.

Kammer, T. (2007). Visual Masking by Transcranial Magnetic Stimulation in the First 80 Milliseconds, *Advances in Cognitive Psychology, 3*, 177-179.

Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception, *The Journal of Neuroscience*, 17, pp. 4302-4311.

Kaplan, S. (1992). Environmental Preference in a Knowledge-seeking, Knowledge-using Organism. In Barkow, J.H., Cosmides, L. & Tooby, J. (Eds.) *The Adapted Mind*, Oxford University Press, New York.

Kaplan S. (1995). The Restorative Benefits of Nature: Toward an Integrative Framework. The Journal of Environmental Psychology,15, pp.169-182.

Kenealy, P.M. (1986). The Velten Mood Induction Procedure: A Methodological Review. *Motivation and Emotion*, 10, 4.

Klinger, E., Légeron, P., Roy, S., Chemin, I., Lauer, F. & Nuges, P. (2006). Virtual Reality Exposure in the Treatment of Social Phobia, in *Cybertherapy - Internet and Virtual Reality as Assessment and Rehabilitation Tools for Clinical Psychology and Neuroscience,* G. Riva, C. Botella, P. Légeron and G. Optale (Eds.), Amsterdam, IOS Press.

Klotz, W. & Wolfe, P. (1995). The effect of a masked stimulus on the response to the masking stimulus, *Psychological Research*, 58, pp. 92-101.

Killgore, W.D.S & Yurgelun-Todd, D.A. (2004). Activation of the Amygdala and Anterior Cingulate During Nonconscious Processing of Sad Versus Happy Faces. *NeuroImage*, 21, pp. 1215–1223.

Klein, R.A. (2000). Virtual Reality Exposure Therapy in the Treatment of Fear of Flying, *Journal of Contemporary Psychotherapy*, Vol. 30, No. 2.

Korff, M. & Simon, G. (1996). The Relationship Between Pain and Depression, The British Journal of Psychiatry, Volume 168, Supplement 30, pp 101-108.

Koster, E.H.W., Fox, E. & MacLeod, C. (2009). Introduction to the Special Section on Cognitive Bias Modification in Emotional Disorders, *Journal of Abnormal Psychology*, 118, 1, pp. 1-4.

Krumhuber, E., Cosker, D., Manstead, A., Marshall, D., & Rosin, P.L. (2005). Temporal Dynamics of Smiling: Human versus Synthetic Faces. *XIVth Conference of the International Society for Research on Emotions*, poster.

Lachter, J., Forster, K.I., & Ruthruff, E. (2004). Forty-five years after Broadbent: Still no identification without attention, *Psychological Review*, 111, 880–913.

Lamme, V.A.F (2004). Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness, *Neural Networks*, 17, pp. 861–872.

Lamme, V.A.F., Super, H., Landman, R., Roelfsema, P.R. & Spekreijse, H. (2000). The role of primary visual cortex (V1) in visual awareness, *Vision Research*, 40, pp. 1507-1521.

Lamson, R.J. (2002). Virtual Reality Immersion Therapy for Treating Psychological, Psychiatric, Medical, Educational and Self-Help Problems, *US Patent No. 6425764 B1, July 2002*

Larsen, R.J., Mercer, K.A. & Balota, D.A. (2006). Lexical Characteristics of Words Used in Emotional Stroop Experiments. *Emotion*, 6, 1, pp. 62-72.

Larsen, R. J., Mercer, K. A., Balota, D. A. & Strube, M. J. (2008). Not All Negative Words Slow Down Lexical Decision and Naming Speed: Importance of Word Arousal. *Emotion*, 8, pp. 445-452.

Lavie, N. (1995). Perceptual Load as a Necessary Condition for Selective Attention, *Journal of experimental Psychology: Human Perception and Performance*, **21**, 451-468.

Layard, P.R.G (Chair), Bell, S., Clark, D., Knapp, M., Meacher, M.C., Priebe, S., Thornicroft, G., Turnberg, L. & Wright, B. (2006). *The Depression report: A New Deal for Depression and Anxiety Disorders*, London School of Economics.

Leary, MR. & Baumeister, RF. (2000). The Nature and Function of Self-esteem: Sociometer Theory, in Zanna, M.P. (Ed.) *Advances in Experimental Social Psychology*, 32, pp. 1-62.

Leberman, S., Mcdonald, L. & Doyle, S. (2006). *The Transfer of Learning*, Gower, UK.

Lee, I., Tyrer, P. & Horn, S. (1983). A Comparison of Subliminal, Supraliminal and Faded Phobic Cine-films in the Treatment of Agoraphobia, *British Journal of Psychiatry, 143*, pp. 356–361.

Lehmkuhle, S. & Fox, R. (1980). Effect of Depth Separation on Metacontrast Masking. *Journal of Experimental Psychology: Human Perception and Performance*, 6,4, pp. 605-621.

Levine, L.J. & Pizarro, D.A. (2004). Emotion and Memory Research: A Grumpy Overview. *Social Cognition*, 22, pp. 530-554.

Lewis, M., Haviland-Jones, J.M. & Feldman Barrett, L. (Eds.) **(2008).** *The Handbook of Emotions*, The Gilford Press, NY.

Lofler, G., Gordon, G.E., Wilkinson, F., Goren D. & Wilson H.R. (2005). Configural masking of faces: Evidence for high-level interactions in face perception, *Vision Research*, 45 (2005). 2287–2297.

Loomis, J.M. (2002). Using Immersive Virtual Reality to Study Visual Space Perception, Visual Control of Locomotion, and Visually-based Navigation. *Journal of Vision*, 2, 10, pp. 3.

Lotto, R.B. and Purves, D. (2000). An empirical explanation of color contrast. *Proceedings of the National Academy of Science USA,* 97, pp. 12834-12839.

McIntosh, D.N. (1996). Facial Feedback Hypotheses: Evidence, Implications, and Directions. *Motivation and Emotion*, 20, 2.

McNair, D.M., Lorr, M. & Droppelman, L.F. (1971). *Manual for the profile of mood states*. San Diego, CA: Educational and Industrial Testing Service.

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., & Holker, L. (2002). Selective attention and emotional vulnerability: Assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of Abnormal Psychology*, 111, pp. 107–123.

MacKnik, S.L. & Livingstone, M.S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system, *Nature: Neuroscience*, 1, 2, pp. 144-149.

Macknik, S.L. & Martinez-Conde, S. (2007). The role of feedback in visual masking and visual processing, *Advances in Cognitive Psychology*, volume 3, 1-2, pp. 125-152.

Marks, I.M., Mataix-Cols, D., Kenwright, M., Cameron, R., Hirsch, S. & Gega, L. (2003). Pragmatic evaluation of computer-aided self-help for anxiety and depression, *British Journal of Psychiatry*, 183, 57-65.

Martin, N.G. (2006). *Human Neuropsychology*, Prentice Hall, Europe.

Maslow, A.H. (1943). A Theory of Human Motivation. *Psychological Review*, **50**, 4, pp. 370-96.

Mathews, A. & MacLeod, C. (2002). Induced processing biases have causal effects on anxiety. *Cognition and Emotion*, 16, 3, pp. 331–354.

Mathews, A., & MacLeod, C. (2005). Cognitive Vulnerability to Emotional Disorders. *Annual Review of Clinical Psychology*, 1, pp. 167–195.

Mayer, B. & Merckelbach H. (1999). Unconscious processes, subliminal stimulation, and anxiety, *Clinical Psychology Review*, 19, 5, pp. 571–590.

Mayer, B., Merckelbach, H., de Jong, P. J., & Leeuw, I. (1998). Skin conductance responses of spider phobics to backwardly masked phobic cues, *Journal of Psychophysiology*, 13, 3, pp. 152-159.

Mayer, J.D., Allen, J.P. & Beauregard, K. (1995). Mood Inductions for Four Specific Moods: a Procedure Employing Guided Imagery Vignettes with Music. *Journal of Mental Imagery*, 19, 2, pp. 133-150.

Mealey, L. & Theis, P. (1995). The Relationship between Mood and Preferences among Natural Landscapes: An Evolutionary Perspective. *Ethology and Sociobiology*, 16, 3, pp. 247-256.

Meehan, M., Insko,B., Whitton, M. & Brooks, FP. Jr. (2002). Physiological Measures of Presence in Stressful Virtual Environments, *ACM Transactions on Graphics*, 21, 3.

Meijer, F., Geudeke, B.L. & Van den Broek, E.L. (2009). Navigating through Virtual Environments: Visual Realism Improves Spatial Cognition,Cyberpsychology & Behavior, 12, 5, pp. 517-521.

Messinger, D.S., Fogel, A. & Dickson, K.L. (1999). What's in a Smile? *Developmental Psychology*, 35, 3, pp. 701-708.

Milgram S. (1963). Behavioral study of obedience, *Journal of Abnormal Social Psychology*, 67, pp. 371-8.

Miller, J. (1989). The control of attention by abrupt visual onsets and offsets, *Perception & Psychophysics*, 45, pp. 567-571.

Minsky, M. (1980). Telepresence. *Omni*, June 1980, pp. 45-51.

Mitroff, S.R. & Scholl, B.J. (2005). Forming and Updating Object Representations Without Awareness: Evidence from Motion-induced Blindness. *Vision Research*, 45, pp. 961–967.

Moraglia, G. & Schneider, B (1990). Effects of Direction and Magnitude of Horizontal Disparities on Binocular Unmasking. *Perception*, 19, pp. 581-593.

Moore, T.E. (1992). Subliminal Perception: Facts and Fallacies, *Skeptical Enquirer*, 16, 3.

Moser, E., Derntl, B., Robinson, S., Fink, B., Gur, R.C., Grammer, K. (2006). Amygdala Activation at 3T in Response to Human and Avatar Facial Expressions of Emotions. *Journal of Neuroscience Methods*, 161, 1, pp. 126-133.

Mühlberger, A., Pauli, P. & Wiedemann, G. (2003). Efficacy of a One-session Virtual Reality Exposure Treatment for Fear of Flying, *Psychotherapy Research*, 1, 3, pp. 323–336.

Mühlberger, A., Weik, A., Pauli, P. & Wiedemann, G. (2006). One-session Virtual Reality Exposure Treatment for Fear of Flying: 1-year Follow-up and Graduation Flight Accompaniment Effects, *Psychotherapy Research, January*, 16, 1, pp. 26-40.

Murry, S.O., Kersten, D., Olshausen, B.A., Schrater, P. & Woods, D.L. (2002). Shape Perception Reduces Activity in Human Primary Visual Cortex. *Procedures of the National Academy of Science U.S.A*, 99, pp. 15164-15169.

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention, *Psychological Science*, 13, pp. 416–424.

Neenan, M. & Dryden, W. (2004). *Cognitive Therapy: 100 Key Points & Techniques*, Brunner-Routledge UK

Neumann, O. (1990). Direct parameter specification and the concept of perception. *Psychological Research, 52*, pp. 207-215.

Neumann, O. & Klotz, W. (1994). Motor Responses to Nonreportable, Masked Stimuli. Where is the Limit of Direct Parameter Specification? In Umilta, C. & Moscovitch, M. (Eds.), *Attention and Performance XV*, MIT Press, USA.

Neumann, R., & Strack, F. (2000). "Mood contagion": The Automatic Transfer of Mood Between Persons. *Journal of Personality and Social Psychology,* 79, 2, pp. 211-223.

NHS Choices: CBT (2010). http://www.nhs.uk/Conditions/Cognitive-behavioural-therapy/Pages/What-is-it-used-for.aspx (4/11/2010).

NHS Choices: Depression Treatments (2010).
http://www.nhs.uk/Conditions/Depression/Pages/Treatment.aspx (accessed November 2010).

NHS Choices: Generalised Anxiety Disorder Treatment Map (2010).
http://healthguides.mapofmedicine.com/choices/map/anxiety4.html (accessed 04/11/2010).

NICE (2008). Computerised cognitive behaviour therapy for depression and anxiety: Review of Technology Appraisal 51, www.nice.org.uk (accessed 13/10/2011).

NICE (2010). National Institute for Health & Clinical Excellence Guidance for Treating Depression in Adults
http://www.nice.org.uk/nicemedia/live/12329/45891/45891.pdf (accessed 04/11/2010).

Nokes, T. J. (2009). Mechanisms of knowledge transfer, Thinking & Reasoning, 15, 1-36.

O'Brien, M. & Houston, G. (2000). *Integrative Therapy, a Practitioner's Guide*, Sage.

Öhman, A. (2002). Automaticity and the Amygdala: Nonconscious Responses to Emotional Faces. *Current Directions in Psychological Science*, 11, 2, pp. 62-66.

Öhman, A. & Soares, J.J.F. (1994). Unconscious Anxiety: Phobic Responses to Masked Stimuli. *Journal of Abnormal Psychology*, 103, 2, pp. 231-240.

Orians, GH., & Heerwagen, JH. (1992). Evolved Responses to Landscapes, in Barkow, JH., Cosmides, L. & Tooby, J. (Eds.) *The Adapted Mind,* Oxford University Press, New York.

Ozonoff, S., Pennington, B. & Rogers, S. (1990). Are there emotion perception deficits in young autistic children? *Journal of Child Psychology & Psychiatry*, 31, pp. 343-363.

Pan, X., Gillies M., and Slater, M. (2008). Male Bodily Responses During an Interaction with a Virtual Woman. In *IVA 2008*, Tokyo, Japan.

Pan, X. and Slater, M. (2011). Confronting a Moral Dilemma in Virtual Reality: A Pilot Study in Human-Computer Interaction, *25th BCS Conference on Human-Computer Interaction (HCI).*

Panksepp, J. (1998). *Affective Neuroscience – the Foundations of Human and Animal Emotions*, Oxford University Press, USA.

Pasetto, M. & Barbati, S.D. (2011). How the Interpretation of Drivers' Behavior in Virtual Environment Can Become a Road Design Tool: A Case Study, *Advances in Human-Computer Interaction*, Volume 2011, Article ID 673585.

Pertaub, D.P., Slater, M. & Baker, C. (2001). An Experiment on Public Speaking Anxiety in Response to Three Different Types of Virtual Audiences. *Presence: Tele-operators and Virtual Environments*, 11, pp. 68-78.

Pessoa, L., McKenna, M., Gutierrez, E. & UngerleiderL.G. (2002). Neural Processing of Emotional Faces Requires Attention. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 17, pp. 11458-11463.

Pfungst, O. (1911). *Clever Hans (The Horse of Mr. Von Osten): A Contribution to Experimental Animal and Human Psychology*, Henry Holt & Co./Project Gutenberg EBook, http://www.gutenberg.org/files/33936/33936-h/33936-h.htm (accessed 25/09/2012).

Przeworski, A. & Newman, M.G. (2006). Efficacy and utility of computer-assisted cognitive behavioural therapy for anxiety disorders. *Clinical Psychologist*, 10, 2, pp. 43-53.

Posner, M.I. & Boies, S.J. (1971). Components of Attention. *Psychological Review*, 78, 5, pp. 391-408.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting of attention, *Journal of Neuroscience*, 4, 1863-1874.

Purcell, D.G., Stewart, A.L. & Brunner, R.L. (1974). Metacontrast Target Detection under Light and Dark Adaptation, *Bulletin of the Psychonomic Society*, 3, pp. 199-201.

Rachman, S. (2004). *Anxiety* (2nd edition), Psychology Press UK.

Rains, G.D. (2002). *Principles of Human Neuropsychology*. McGraw-Hill, USA.

Ramachandran, V.S. & Cobb, S. (1995). Visual Attention Modulates Metacontrast Masking, *Nature*, 373, 5, pp. 66-68.

Rastad, C., Ulfberg, J. & Lindberg, P. (2011). Improvement in Fatigue, Sleepiness, and Health-Related Quality of Life with Bright Light Treatment in Persons with Seasonal Affective Disorder and Subsyndromal SAD. Depression Research and Treatment, 2011, pp. 1-10.

Reber, R. & Schwarz, N. (1999). Effects of Perceptual Fluency on Judgments of Truth, *Consciousness and Cognition*, 8, 3, pp. 338-342.

Reber, R., Winkielman P., & Schwarz N. (1998). Effects of perceptual fluency of affective judgments. *Psychological Science*, 9, 1, pp. 45-48.

Rizzo, AA., Neumann, U., Enciso, R., Fidaleo, D. & Noh, JY. (2001). Performance-Driven Facial Animation: Basic Research on Human Judgments of Emotional State in Facial Avatars. *CyberPsychology & Behavior*, 4, 4, pp. 471-487.

Rizzo, AA., Pair, J., Graap, K., Manson, B., McNerne, P.J., Wiederhold, B., Wiederhold, M. & Spira, J. (2005). A Virtual Reality Exposure Therapy Application for Iraq War Military Personnel with Post Traumatic Stress Disorder: From Training to Toy to Treatment, Roy. M. (Ed.). *NATO Advanced Research Workshop on Novel Approaches to the Diagnosis and Treatment of Posttraumatic Stress Disorder*, IOS Press, Washington D.C., pp. 235-250.

Ro, T., Singhal, N.S., Breitmeyer, B.G. & Garcia, J.O. (2009). Unconscious processing of color and form in metacontrast masking. *Attention, Perception, & Psychophysics*, 71, 1, pp. 95-103.

Rolls, E.T., Tovée, M.J. & Panzeri, S. (1999). The Neurophysiology of Backward Visual Masking: Information Analysis, *Journal of Cognitive Neuroscience*, 11, 3, pp. 300–311.

Rottenberg, J., Ray, R.R. & Gross, J.J. (2007). Emotion Elicitation Using Films, in Coan, J.A & Allen, J.J.B. (Eds.) *The Handbook of Emotion Elicitation and Assessment*, Oxford University Press, New York.

Roy, S. (2003). State of the art of virtual reality therapy (VRT) in phobic disorders, *Psychology Journal*, 1, 2, pp. 176 – 183.

Saddik, E. (2007). The Potential of Haptics Technologies, IEEE Instrumentation & Measurement Magazine, 10, 1, pp. 10-17.

Saxe, R. & Baron-Cohen, S. (2006). Editorial: The neuroscience of theory of mind, *Social Neuroscience*, 1 (3 - 4), i - ix.

Sayer, E., Lerner, A., Cohen-Or, D., Chrysanthou, Y. & Deussen, O. (2004). Aggressive Visibility for Rendering Extremely Complex Foliage Scenes, *Proceedings of IK* 2004.

Schartau, E.S. , Dalgleish, T. & Dunn B.D. (2009). Seeing the Bigger Picture: Training in Perspective Broadening Reduces Self-Reported Affect and Psychophysiological Response to Distressing Films and Autobiographical Memories, Journal of Abnormal Psychology, 118, 1, pp. 15–27.

Schere, K (2005). What are emotions? And how can they be measured? *Social Science Information*, 44, 4, pp. 695-729.

Schlaghecken, F., Rowley, L., Sembi, S., Simmons, R. & Whitcomb, D. (2007). The negative compatibility effect: A case for self-inhibition, *Advances in Cognitive Psychology*, 3, 1-2, pp. 227-240.

Schmidt, T. (2000). Visual Perception Without Awareness: Priming Responses by Color, in Metzinger, T. (Ed.) *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, MIT Press, pp. 157-170.

Schmidt, N.B., Richey, A.J., Buckner, J.D & Timpano, K.R. (2009). Attention Training for Generalized Social Anxiety Disorder, *Journal of Abnormal Psychology*, 118, 1, pp. 5–14.

Schmitz, N, Kugler, J, & Rollnik, J. (2003). On the Relation between Neuroticism, Self-esteem, and Depression: Results from the National Co-morbidity Survey. *Comprehensive Psychiatry*, 44, 3, pp. 169-176.

Schneider, F, Gur, R.C., Gur, R.E. & Muenz, L.R. (1992). Standardized Mood Induction With Happy and Sad Facial Expressions. *Psychiatry Research*, 51, 19-31.

Schwabe, L., Obermayer, K., Angelucci, A. & Bressloff P.C. (2006). The Role of Feedback in Shaping the Extra-Classical Receptive Field of Cortical Neurons: A Recurrent Network Model, *The Journal of Neuroscience*, 26, 36, pp. 9117–9129.

Schwarz, N. (2010). Feelings-as-Information Theory, in Van Lange, P., Kruglanski, A. & Higgins, E. T. (Eds.), *Handbook of theories of social psychology*, Sage USA.

Schwarz, N., & Clore, G. L. (1983). Mood, Misattribution, and Judgments of Well-being: Informative and Directive Functions of Affective States. *Journal of Personality and Social Psychology*, **45**, pp. 513–523.

Schwarz, N., & Clore, G. L. (2003). Mood as Information: 20 Years Later. *Psychological Inquiry*, 14, 3&4, pp. 296-303.

Shapiro, K. L., Raymond, J. E., & Amell, K. M. (1994). Attention to visual pattern information produces the attentional blink in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, 20, pp. 357-371.

Sigman, M., Spence, S.J. & TingWang, A. (2006). Autism for Developmental and Neuropsychological Perspectives. *Annual Review of Clinical Psychology*, 2, pp. 327-355.

Skalska, B., Jaśkowski, P. & Van der Lubbe, R.H. (2006). The role of direct parameter specification and attentional capture in near threshold priming of motor reactions, *Advances in Cognitive Psychology*, 2, 1, pp. 47-59.

Slater, M. (2004). How Colorful Was Your Day? Why Questionnaires Cannot Assess Presence in Virtual Environments. Presence, 13, 4, pp. 484-493.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments, *Philosophical Transactions of the Royal Society*, 364, 1535, pp. 3549-3557.

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., Pistrang, N. & Sanchez-Vives, M.V. (2006). A Virtual Reprise of the Stanley Milgram Obedience Experiments*, PLoS ONE*, Issue 1, e39.

Slater, M. & Garau, M. (2007). The Use Of Questionnaire Data In Presence Studies: Do Not Seriously Likert. Presence: Teleoperators and Virtual Environments, 16, 4, pp. 447-456.

Slater, M., Khanna, P., Mortensen, J. & Yu, I. (2009). Visual Realism Enhances Realistic Response in an Immersive Virtual Environment, IEEE Computer Graphics Applications, 29, 3, pp. 76-84.

Slater, M., Spanlang, B., Corominas, D. (2010). Simulating Virtual Environments within Virtual Environments as the Basis for a Psychophysics of Presence. *ACM Trans. Graph. 29*, 4, Article 92.

Slater, M., Spanlang, B., Sanchez-Vives, M.V. & Blanke, O. (2010). First Person Experience of Body Transfer in Virtual Reality. *PLoS ONE*, 5, 5, e10564.

Skalska, B., Jaśkowski, P. & Van der Lubbe, R.H.J. (2006). The role of direct parameter specification and attentional capture in near-threshold priming of motor reactions, *Advances in Cognitive Psychology,* 2, 1, pp. 47-59.

Smith, C.A. & Lazarus, R.S. (1990). Emotion and Adaption, Pervin, L.A. (Ed.) *Handbook of Personality: Theory and Research*, Guilford, New York.

Souman, J.L., Frissen, I., Sreenivasa, M.N. & Ernst M.O. (2009). Walking Straight into Circles, *Current Biology,* 19, pp. 1538–1542.

Spencer, T.J. & Shuntich, R (1970). Evidence for an Interuption Theory of Backward masking, *Journal of Experimental Psychology*, 85, pp. 198-203.

Spruyt, A., Clarysse, J., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2010). Affect 4.0: A Free Software Package for Implementing Psychological and Psychophysiological Experiments. Experimental Psychology, 57, 1, pp. 36-45.

Sutherland, G., Newman, B., Rachman, S. (1982). Experimental Investigations of the Relation Between Mood and Intensive Unwanted Cognitions. *British Journal of Medical Psychology*, 55, pp. 127-38.

Suls & Krizan (2005). On the Relations Between Explicit and Implicit Global Self Esteem and Personality, *The New Frontiers of Self Research*, Information Age Publishing.

Tamietto, M. & De Gelder, B. (2009) Emotional Contagion for Unseen Bodily Expressions: Evidence from Facial EMG. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 42, pp. 17661-17666.

Taylor, A.F., Kuo, F.E. & Sullivan, W.C. (2001) Views of Nature: Evidence from Inner City Children. *Journal of Environmental Psychology*, 21, pp. 1-15.

Taylor, DN. & Del Pilar, J. (1992). Self-esteem, Anxiety, and Drug Use. *Psychological reports*, 71, 3, pp. 896-898.

Tedeschi, J. T., Schlenker, B. R., & Bonoma, T. V. (1971). Cognitive dissonance: Private ratiocination or public spectacle? *American Psychologist*, 26, pp. 685-695.

Tipper, S.P. (1985). The negative priming effect: Inhibitory priming with to be ignored objects, *The Quarterly Journal of Experimental Psychology*, 37A, pp. 571-590.

Treisman, A.M. & Gelade, G. (1980). A feature-integration theory of attention, *Cognitive Psychology*, 12, 1, pp. 97-136.

Triesman, A.M. & Schmidt, H. (1982). Illusory Conjunctions in the Perception of Objects, Cognitive Psychology, 14, 1, pp. 107-141.

Tse, P.U., Martinez-Conde, S., Schlegel, A.A. & Macknik, S.L. (2005). Visibility, visual awareness, and visual masking of simple unattended targets are confined to areas in the occipital cortex beyond human V1□V2, *Proceedings of the National Academy of Sciences of the United States of America*, 102, 47, pp. 17178-17183.

Turvey, M.T. (1973). On peripheral and Central Processes in Vision: Inferences From an Informational Processing Analysis of Masking With Patterned Stimuli, Psychological Review, 80, 1, pp. 1-52.

Ulrich, R.S. (1984). View Through a Window May Influence Recovery from Surgery. *Science*, 224, pp. 420-421.

Ulrich, RS. (1986). Human Responses to Vegetation and Landscapes. *Landscape and Urban Planning*, 13, pp. 29-44.

Ulrich, RS., Simons, RF., Losito, BD., Fiorito, E., Miles, MA., & Zelson, M., (1991). Stress Recovery During Exposure to Natural and Urban Environments. *Journal of Environmental Psychology*, 11, pp. 201-230.

Ulrich-Lai Y.M., Herman J.P. (2009). Neural regulation of endocrine and autonomic stress responses. *Nature Reviews Neuroscience*, 10, 6, pp. 397-409.

Valstar, M.F. & Pantic, M. (2010). Induced Disgust, Happiness and Surprise: an Addition to the MMI Facial Expression Database. *Proceedings of the International Conference of Language Resources and Evaluation, Workshop on EMOTION*, pp. 65-70.

Valstar, M.F. & Pantic, M. (2012). Fully Automatic Recognition of the Temporal Phases of Facial Actions. *IEEE Transactions On Systems, Man, And Cybernetics—Part B: Cybernetics*, 42, 1, pp. 28-43.

Van Santen, J.P.H. & Jonides, J. (1978). A Replication of the Face-superiority Effect, *Bulletin of the Psychonomic Society*, 12, 5, pp. 378-380.

Velten, E. (1968). A Laboratory Task for Induction of Mood States. *Behaviour Research and Therapy*, 6, pp. 473-82.

VEPSY (2009). http://www.cybertherapy.info/ (accessed December 2009).

Viggiano, M.P. & Marzi, T. (2010). Context and Social Effects on Face Recognition. In Balcetis, E. & Lassiter, G.D. (Eds.), *Social Psychology of Visual Perception*, Psychology Press, UK.

Vinayagamoorthy, V., Brogni, A., Gillies, M., Slater, M., Steed, A.J. (2004). An Investigation of Presence Response across Variations in Visual Realism In: *the 7th International Conference on Presence, Presence 2004,* pp. 148-155.

Vicory, A.C (1968). *A Brief History of Aircraft Identification Training*. George Washington University Human Resource Research Office, Professional Paper 27-68.

Watson, D. & Clark, L.A. (1999) *The PANAS-X Manual for the Positive and Negative Affect Schedule - Expanded Form* (1994 edition updated 8/1999), The University of Iowa, USA.

Weizenbaum, J. (1966). ELIZA — A Computer Program For the Study of Natural Language Communication Between Man And Machine, *Communications of the ACM,* 9, 1, pp. 36–45.

Wells, A. (1997). *Cognitive Therapy of Anxiety Disorders*, John Wiley & Sons UK.

Wells, N.M. & Evans, G.W. (2003). Nearby Nature: A Buffer of Life Stress Among Rural Children. *Environment & Behaviour*, 35, 3, pp. 311-330.

Weir, W. (1984). Another look at subliminal "facts," *Advertising Age*, October, pp. 46.

Wesstein, N. (1972). Metacontrast, in Jameson, D. & Hurvich, L.M. (Eds.) *Visual Psychophysics*, Springer, New York.

Weisstein, N., Williams, M.C. & Harris, C.S. (1982). Depth, connectedness, and structural relevance in the object-superiority effect: line segments are harder to see in flatter patterns, *Perception,* 11, 1, pp. 5-17.

Weyers, P., Muhlberger, A., Hefele, C. & Pauli, P. (2006). Electromyographic Responses to Static and Dynamic Avatar Emotional Facial Expressions. *Psychophysiology*, 43, pp. 450–453.

Whalen, P.J., Rauch, S.L., Etcoff, N.L., McInerney, S.C., Lee, M.B., & Jenike, M.A. (1998). Masked Presentations of Emotional Facial Expressions Modulate Amygdala Activity without Explicit Knowledge, *The Journal of Neuroscience*, 18, 1, pp. 411–418.

Whitney, D. (2009). Neuroscience: Toward Unbinding the Binding Problem, Current Biology, 19, 6, pp. R251-R253.

Wickens, C.D. (1990). Three-dimensional stereoscopic display implementation: guidelines derived from human visual capabilities, *Stereoscopic Displays and Applications*, Proceedings of SPIE, 1256, pp. 2-11.

Wild, B., Erb, M. & Bartels, M. (2001). Are Emotions Contagious? Evoked Emotions While Viewing Emotionally Expressive Faces: Quality, Quantity, Time Course and Gender Differences. *Psychiatry Research*, 102, 2, pp. 109-124.

Wilhelm, FH., Pfaltz, MC., Gross, JJ., Mauss, IB., Kim, SI. & Wiederhold, BK. (2005). Mechanisms of Virtual Reality Exposure Therapy: The Role of the Behavioral Activation and Behavioral Inhibition Systems. *Applied Psychophysiology and Biofeedback,* 30, 3.

Williams, J.M.G., Mathews, A. & MacLeod, C. (1996). The Emotional Stroop Task and Psychopathology. *Psychological Bulletin*, 120, 1, pp. 3-24.

Winkielman, P. & Berridge, KC. (2004). Unconscious Emotion, *American Psychological Society*, 13, 3.

Wolfe, J.M., Cave, K.R. & Franzel, S.L. (1989). Guided search: an alternative to the feature integration model for visual search, *Journal of Experimental Psycholgy: Human Perception and Performance*, 15, 3, pp. 419-433.

Wong, P.S. & Root, J.C. (2003). Dynamic variations in affective priming, *Consciousness and Cognition*, 12, pp. 147–168.

Woods, A.J. & Helliwell, J. (2012). Investigating the cross-compatibility of IR-controlled active shutter glasses. *Proceedings SPIE: Stereoscopic Displays and Applications XXIII*, pp. 8288-1C.

Wright, J. & Mischel, W. (1982). Influence of Affect on Cognitive Social learning Variables. *Journal of Personality and Social Psychology*, 43, pp. 901-14.

Wyer, R. S., & Carlston, D. E. (1979) *Social cognition, inference, and attribution*, Lawrence Erlbaum Associates, Inc.

Wyszecki, G. & Stiles, W. S. (2000). *Color Science: Concepts and Methods, Quantitative Data and Formulae*, Wiley-Blackwell, UK.

Yuan, Y. & Steed, A. (2010). Is the Rubber Hand Illusion Induced by Immersive Virtual Reality? *IEEE VR*, proceedings.

Yun, C., Deng, Z. & Hiscock, M. (2009). Can Local Avatars Satisfy a Global Audience? A Case Study of High-fidelity 3D Facial Avatar Animation in Subject Identification and Emotion Perception by US and International Groups. *Computers in Entertainment*, 7, 2.

Zajonc, R. (1980). Feeling and Thinking: Preferences Need No Inferences. *American Psychologist*, 33, 2, pp. 151-175.

Zimmons, P. & Panter, A. (2003). The Influence of Rendering Quality on Presence and Task Performance in a Virtual Environment. *IEEE Virtual Reality, Proceeding,* pp. 293-294.

13 Appendices

13.1 Appendix 1 - Participant Profile Questionnaire

Choice behaviour within a virtual environment

1. Participant number: ....

2. Gender: ....

3. Handedness: ....

4. Do you regard yourself as a regular computer game or video game player? ....

    4a. How much time have you spent playing computer or video games in the last month? ....

    4b. How much time have you spent playing 3D computer games or video games in the last month (as opposed to 2D or puzzle games)? ....

5. Do you have regular experience of immersion in virtual reality? ....

    5a. How many hours have you experienced stereoscopic graphics systems in the last month? ....

    5b. How many hours have you experience high end virtual environment systems such as a Cave, a head mounted display or similar in the last month? ....

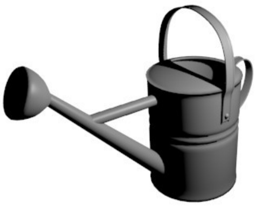
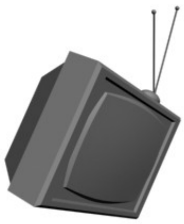
13.2 Appendix 2 - Experiment 1 Record Sheet

Room	Any object perceived - yes or no?	Object type
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		

13.3 Appendix 3 - Experiment 1 Object Reference Sheet

Visual masking within a virtual environment

Object list

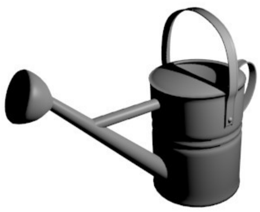
 chair	 teapot
 wateringcan	 tap
 lamp	 phone
 cup	 pan
 bicycle	 TV

Please use the labels on this page when entering the object type the objects on the questionnaire.

13.4 Appendix 4 – Experiment 3 Object Reference Sheet

Visual masking parameters within a virtual environment

Object reference list

 chair	 teapot
 wateringcan	 cup
 lamp	 pan

Please use the labels on this page when describing the object type to the experimenter.

13.5 Appendix 5 – Experiment 3 Expression Reference Sheet

Visual masking parameters within a virtual environment

Expression reference list

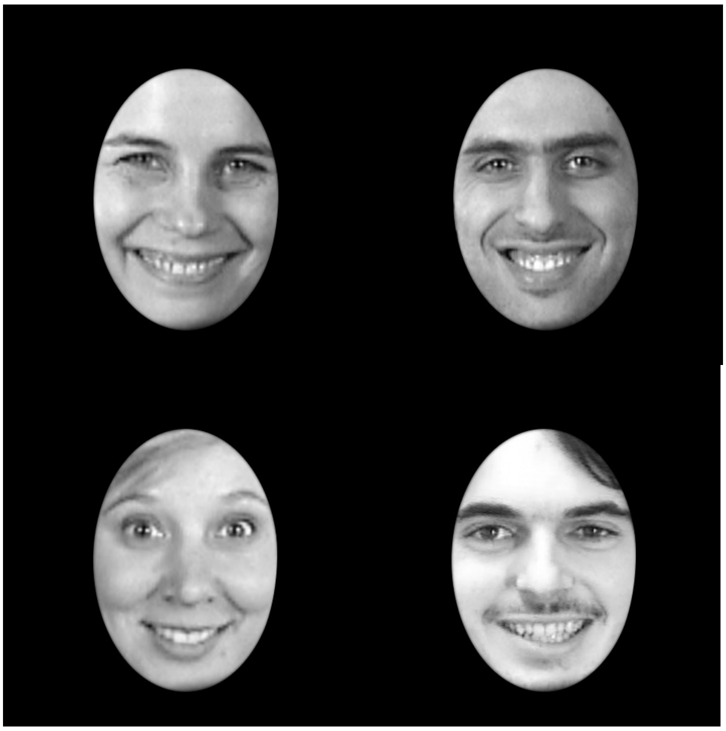
 happy	 angry
 sad	 surprised
 disgusted	 afraid

Please use the labels on this page when describing the expression type to the experimenter.

13.6 Appendix 6 – CBM Test Stimuli

These facial expression images were derived from the MMI Facial Expression Database (Valstar & Pantic 2010).

Smiling faces:



Angry looking faces: