INVESTIGATING THE LIMITS OF HOW EXPECTATION CAN SHAPE AFFECTIVE JUDGEMENT

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Medical and Human Sciences

2016

ADAM LAWRENCE

SCHOOL OF PSYCHOLOGICAL SCIENCES
CONTENTS

LIST OF FIGURES .............................................................................................................................. 8
LIST OF TABLES ................................................................................................................................. 10
ABBREVIATIONS ............................................................................................................................... 11
ABSTRACT ........................................................................................................................................ 12
DECLARATION .................................................................................................................................... 13
COPYRIGHT STATEMENT .................................................................................................................. 13
ACKNOWLEDGEMENTS ................................................................................................................... 14

1 GENERAL INTRODUCTION ........................................................................................................ 15
  1.1 Preface ...................................................................................................................................... 15
  1.2 Assimilation-Contrast Theory ................................................................................................. 19
      1.2.1 Defining assimilation and contrast effects ........................................................................ 19
      1.2.2 Investigating the time course of assimilation effects ...................................................... 20
      1.2.3 Establishing the need to determine the boundaries of assimilation effects ....................... 22
  1.3 Predictive coding and the prediction error hypothesis ........................................................... 24
      1.3.1 Predictive coding of liking ............................................................................................... 24
      1.3.2 Predictive coding as Bayesian inference ......................................................................... 26
      1.3.3 Prediction errors – A neurological marker of expectation? ............................................ 27
  1.4 Promoting affective expectations ............................................................................................ 29
      1.4.1 Affective priming ............................................................................................................. 29
      1.4.2 Priming affective values and measuring changes in stimulus liking ............................... 30
  1.5 The aims of the current investigation ..................................................................................... 31

2 GENERAL METHODS .................................................................................................................. 32
  2.1 Participants .............................................................................................................................. 32
  2.2 Questionnaire measures ......................................................................................................... 32
      2.2.1 The LOT-R ...................................................................................................................... 33
  2.3 Experimental behavioural paradigms ...................................................................................... 33
  2.4 Implicit attitude measures ....................................................................................................... 34
      2.4.1 The IAT .......................................................................................................................... 34
      2.4.2 The GNAT ...................................................................................................................... 36
  2.5 The Psychometric Function .................................................................................................... 38
  2.6 Statistical analysis .................................................................................................................... 40
3 INVESTIGATING THE INFLUENCE OF AFFECTIVE EXPECTATIONS USING ATTRACTIVE AND UNATTRACTIVE FACES .................................................................................................................. 41

3.1 Introduction .................................................................................................................. 41

Experiment 1 ..................................................................................................................... 45

3.2 Methods ....................................................................................................................... 45
  3.2.1 Participants .............................................................................................................. 45
  3.2.2 Stimuli and apparatus ............................................................................................ 45
  3.2.3 Design and procedure ........................................................................................... 47
  3.2.4 Data analysis .......................................................................................................... 49

3.3 Results ......................................................................................................................... 50
  3.3.1 Attractiveness ratings for parent faces ................................................................. 50
  3.3.2 Attractiveness ratings for 30% and 70% anchors ................................................. 50
  3.3.3 Attractiveness ratings for presentations between 40% and 60% morph ......... 51

3.4 Interim discussion ...................................................................................................... 53

Experiment 2 ..................................................................................................................... 55

3.5 Amended design ........................................................................................................ 55

3.6 Methods ....................................................................................................................... 56
  3.6.1 Participants .............................................................................................................. 56
  3.6.2 Stimuli and apparatus ............................................................................................ 56
  3.6.3 Design and procedure ........................................................................................... 57
  3.6.4 Data analysis .......................................................................................................... 59

3.7 Results ......................................................................................................................... 61
  3.7.1 Liking ratings for each attractiveness level ......................................................... 61
  3.7.2 The effect of gender ............................................................................................... 61
  3.7.3 The effect of cue .................................................................................................... 62

3.8 Discussion .................................................................................................................... 63

3.9 Conclusion .................................................................................................................... 65

4 AFFECTIVE ASSIMILATION AND CONTRAST EFFECTS THROUGH THE OBSERVATION OF HAPPY AND SAD EMOTIONAL FACES ........................................................................ 66

4.1 Introduction .................................................................................................................. 66

Experiment 3 ..................................................................................................................... 71

4.2 Methods ....................................................................................................................... 71
  4.2.1 Participants .............................................................................................................. 71
  4.2.2 Stimuli and apparatus ............................................................................................ 71
  4.2.3 Design and procedure ........................................................................................... 72
  4.2.4 Presentation conditions ......................................................................................... 74
  4.2.5 Trial breakdown .................................................................................................... 76
  4.2.6 Data analysis .......................................................................................................... 76

4.3 Results .......................................................................................................................... 77
  4.3.1 All participants at 100ms ...................................................................................... 77
  4.3.2 Data from participants who completed 50ms and 100ms trials ....................... 82

4
8.6.5 Mood Manipulations ................................................................. 172
8.6.6 Imaging paradigms ................................................................. 172
8.7 Conclusion .............................................................................. 173

REFERENCES ................................................................................. 174

APPENDIX A ................................................................................. 187
A.1 Preliminary Study 3a: Identifying attractive and unattractive faces ........................................ 187
A.2 Preliminary Study 3b: Establishing identity boundaries for composite morphs .................. 190
A.3 Preliminary Study 3c: Testing the assumption that attractive faces are more liked ................ 195

APPENDIX B ................................................................................. 200
B.1 Example page from participant response sheet ................................................................... 200
B.2 Questionnaire item: Life orientation test revised (LOT-R) ................................................ 201

APPENDIX C ................................................................................. 203
C.1 List of images used in the GNAT and rating tasks ............................................................. 203
C.2 List of words used in the GNAT task ............................................................................. 208
C.3 Example Functions from intermediate food responses ......................................................... 209

APPENDIX D ................................................................................. 210
D.1 Preliminary Study 6a: Confirming the preference for curvature .......................................... 210
D.2 Preliminary Study 6b: Pattern differentiation task ............................................................ 212
D.3 Sharp Patterns used in pattern judgement tasks ............................................................... 215
D.4 List of words used in the IAT task portion of pattern judgement task ................................ 216
D.5 Excluded participants from Experiment 6 ........................................................................ 217
D.6 Graphs of average liking ratings for each group in experiment 6 ...................................... 220

APPENDIX E ................................................................................. 222
E.1 Trial breakdown for session 2 and 3 (congruent trials) ....................................................... 222
E.2 Trial breakdown for session 3 (minor discrepancy trials) ................................................... 223
E.3 Trial breakdown for session 3 (extreme discrepancy trials) ................................................ 224

Word count: 69,282
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Theories of disconfirmation of expectations. Adapted from Anderson (1973).</td>
<td>20</td>
</tr>
<tr>
<td>1.2</td>
<td>The principle of predictive coding based upon the extraction of low spatial</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>frequency information.</td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td>General architecture of a hierarchical predictive coding model. Adapted</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>from Rao and Ballard (1999).</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>Trial sequence from experiment 1</td>
<td>49</td>
</tr>
<tr>
<td>3.2</td>
<td>Results of paired samples t-tests conducted using the attractiveness ratings</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>of the male and female 30%/70% face pairs.</td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>Face Classification for stimuli used in experiment 2</td>
<td>57</td>
</tr>
<tr>
<td>3.4</td>
<td>Trial sequence from experiment 2</td>
<td>58</td>
</tr>
<tr>
<td>3.5</td>
<td>Example psychometric functions illustrating responses from a single</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>participant in experiment 2</td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>Functions for moderately attractive male and female faces presented in</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>attractive and unattractive cue conditions</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>Examples of faces shown during the training procedure</td>
<td>72</td>
</tr>
<tr>
<td>4.2</td>
<td>Sequence of events from experiment 3.</td>
<td>74</td>
</tr>
<tr>
<td>4.3</td>
<td>Emotional intensities used in each presentation condition</td>
<td>75</td>
</tr>
<tr>
<td>4.4</td>
<td>The influence of discrepancy direction and emotion condition on the</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>accuracy of participant responses for 100ms presentations</td>
<td></td>
</tr>
<tr>
<td>4.5</td>
<td>Optimism x discrepancy size x discrepancy direction interaction for 100ms</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>presentations</td>
<td></td>
</tr>
<tr>
<td>4.6</td>
<td>The interaction between emotion, discrepancy type, and discrepancy size for</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>100ms presentations</td>
<td></td>
</tr>
<tr>
<td>4.7</td>
<td>The influence of discrepancy direction and emotion condition on the</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>accuracy of participant responses at 50ms and 100ms.</td>
<td></td>
</tr>
<tr>
<td>4.8</td>
<td>The interaction between discrepancy size, discrepancy direction, and emotion</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>condition for 100ms and 50ms presentations</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>Valence ratings for food groups used in experiment 4</td>
<td>96</td>
</tr>
<tr>
<td>5.2</td>
<td>Example trial sequence from either rating phase of experiment 4</td>
<td>99</td>
</tr>
<tr>
<td>5.3</td>
<td>Examples of correct and incorrect responses to ‘Food + Good Word’ sequence</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>GNAT trials</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.4  Non-significant interaction between session and cue for ratings of intermediate food objects .......................................................... 104
Figure 6.1  Examples of the sharp, smooth and intermediate versions of one pattern used in experiment 5 ........................................................................................................ 115
Figure 6.2  Example trial sequence from the conditioning task ................................. 117
Figure 6.3  Example trial sequence from the rating task .............................................. 118
Figure 6.4  Liking ratings for intermediate stimulus levels from experiment 5 collapsed across session and cue type .................................................................................................. 122
Figure 6.5  Functions for cued and non-cued responses from participants presented across session 1 and session 2 in experiment 5 ........................................................................ 123
Figure 6.6  Sample set of stimuli from experiment 6 ........................................................ 125
Figure 6.7  Example of stimulus increments and cues used in the cued condition. .......... 128
Figure 6.8  Example trial sequence from the rating phase of experiment 6 in the non-cued condition .................................................................................................................. 129
Figure 6.9  Liking ratings for intermediate stimulus levels from experiment 6 collapsed across session (1 and 2) and group (cued and non-cued) .................................................................. 132
Figure 6.10 Group-level functions for cued and non-cued participants presented across session 1 and session 2 from experiment 6 .............................................................................. 133
Figure 7.1  Mean reproductions of movement amplitudes from Hollingworth (1909). Diagram taken from Laming (1997) .................................................................................................. 141
Figure 7.2  Example trial sequence from experiment 7. .............................................. 146
Figure 7.3  Time intervals used for short, medium and long presentations .................. 147
Figure 7.4  Response data from congruent trials in session 3 ........................................ 153
Figure 7.5  Response data from minor discrepancy trials in session 3 ............................ 154
Figure 7.6  Response data from extreme discrepancy trials in session 3 ....................... 155
LIST OF TABLES

Table 2.1. Trial Sequence for the standard IAT design adapted from the Bush vs. Gore design described in Greenwald, Nosek, and Banaji (2003) .............................................................................................. 35
Table 2.2 Trial sequence for Dog vs. Cat GNAT ........................................................................................................ 37
Table 3.1 Attractive / unattractive face pairs from experiment 1 ........................................................................... 46
Table 3.2 Mean percentage of liked responses for the moderately attractive sub-groups .......... 61
Table 5.1 Design of GNAT ........................................................................................................................................ 100
Table 5.2 Pairwise comparisons for the main effect of food group ......................................................... 103
Table 6.1 Design of IAT task .................................................................................................................................. 120
Table 6.2 Design of rating task ............................................................................................................................. 127
Table 7.1 Results of One sample t-tests of regressions coefficients for congruent and discrepant trial types .......................................................................................................................... 152
# ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2AFC</td>
<td>2 alternative forced choice</td>
</tr>
<tr>
<td>ACT</td>
<td>Assimilation-contrast theory</td>
</tr>
<tr>
<td>AEM</td>
<td>Affective Expectation Model</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>ANCOVA</td>
<td>Analysis of covariance</td>
</tr>
<tr>
<td>EEG</td>
<td>Electroencephalography</td>
</tr>
<tr>
<td>fMRI</td>
<td>Functional magnetic resonance imaging</td>
</tr>
<tr>
<td>FRIDa</td>
<td>Foodcast Research Image Database</td>
</tr>
<tr>
<td>JND</td>
<td>Just noticeable difference</td>
</tr>
<tr>
<td>KDEF</td>
<td>Karolinska Directed Emotional Face Database</td>
</tr>
<tr>
<td>LOT-R</td>
<td>Life Orientation Test Revised</td>
</tr>
<tr>
<td>LSF</td>
<td>Low spatial frequencies</td>
</tr>
<tr>
<td>ms</td>
<td>Millisecond</td>
</tr>
<tr>
<td>PSE</td>
<td>Point of subjective equality</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SDT</td>
<td>Signal detection theory</td>
</tr>
<tr>
<td>SE</td>
<td>Standard error</td>
</tr>
<tr>
<td>VAS</td>
<td>Visual analogue scale</td>
</tr>
</tbody>
</table>
The generation of predictions shapes our experience of the world around us. By making inferences about what is likely to happen within a given scenario, we can conserve cognitive resources and enhance our prospects of survival. Predictive coding accounts of perception indicate that this is achieved by minimally processing information that is consistent with our expectations, and prioritising the processing of unexpected or meaningful information. Predictions are also beneficial in situations where accurate perception is difficult, and clues like contextual information allow expectations to ‘fill in the blanks’ when sensory information is noisy or ambiguous. This comes at a cost, however, and a reliance upon expectations can lead to perceptual biases, and in certain cases misperceptions.

According to Assimilation Contrast Theory (ACT) and the Affective Expectation Model (AEM), when we attempt to judge affectively ambiguous stimuli, our judgements are biased by expectations in a similar manner. If stimuli are within an acceptable range of an existing expectation, minor discrepancies will be ignored and judgements of those stimuli will fall in line with expectations (assimilation). Alternatively, if the affective discrepancy between expectation and stimulus is so large that it is acknowledged, the extent of that discrepancy will be exaggerated instead (contrast). This thesis aimed to investigate the boundaries and time-course of these effects.

A series of behavioural experiments were conducted to investigate: (i) whether predictive cues promoted a state of affective readiness, where judgements across a range of stimuli were biased based upon the assumption that they were broadly part of a positive or negative category (chapters 3 and 4); (ii) whether affective biases (assimilation effects) persisted over time (chapters 5 and 6); and (iii) whether the boundaries of affective and perceptual assimilation effects remained consistent over time (chapter 6 and 7).

Psychophysical measures of affective bias indicated that predictive cues influenced participants to judge the same stimuli differently, according to whether they expected those stimuli to be positive or negative. Furthermore, after expectations were learned, judgements of the same stimuli continued to be biased toward expectations after a period of one week. When stimuli from affectively or perceptually distinct categories were manipulated slowly over time, to the point where they became identical, judgements of those stimuli continued to be influenced by the expectation that they should remain distinct. These findings indicate that the boundaries of perceptual and affective assimilation effects may not be static, and if deviations from expectation are small enough to go generally unnoticed, people may update their internal representations of items over time, and the boundaries of acceptance which surround those representations.
DECLARATION

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT STATEMENT

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DoculInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.
I must first offer my thanks to my supervisors Professor Wael El-Deredy, and Professor Rebecca Elliott. Their support during my PhD has been invaluable, and they have been extremely generous in sharing both their time, and their expertise throughout the project. I would also like to thank Dr Marco Hoeksma from Unilever R&D for organising my industrial placements, and all of the staff at Unilever for being so accommodating during the months that I spent there.

Secondly, I’d like to thank Dr Gemma Barnacle and Dr Stacey Humphries who had the dubious pleasure of sharing an office with me for the majority of our studies. I couldn’t have wished for a pair of nicer office buddies. If I ever need to convince anyone that I’m Irish, or if I need to find someone that’s secretly almost as angry as I am about nearly every aspect of social media, you guys will always be my first port of call. I’d also like to thank my original office buddy Dr Jenna ‘J-Bags’ Roberts, for welcoming me into the fold when I first started.

Thanks the Thursday drinks crew (past and present): Emma Davis, George ‘Dr Nesh’ Farmer, Kiki Giannou, Dr Michelle Hall, Robert ‘Dr Bob’ Hoskin, Dr Lizzi Lewis, Dr Oana Lindner, Dr Dan Poole, Dr Zoe Tsivou, and Dr Cliff Workman. Special thanks to Dr Sam Couth for starting the snow globe trend that has slowly overtaken my desk, and to Rachel Ashworth and Dr Jeff Wood who have both been firm friends ever since we began our studies at the University together in 2008. All of you have made the last 4 years immensely enjoyable, here’s to more cycling holidays, half marathons (maybe) and other random nights out in the future.

Additional thanks go to Karen Lopez who helped with data collection for the timing interval study, and to my friends from back home in Hawarden and elsewhere around the world, who I don’t get to see as often as I would like.

Finally, I would like to thank my parents David and Jacqui Lawrence who have remained endlessly supportive throughout this project, and all of my other endeavours. You guys really are the best 😊
1 GENERAL INTRODUCTION

1.1 Preface

A growing body of literature suggests that when a strong expectation exists, that expectation can be sufficient to bias perception, and in many cases the effect of this expectation can be beneficial in our interactions with the outside world. For example, when entering a familiar environment, we instinctively know that the regularities of that environment dictate that the placement of objects, textures and colours is likely to be similar to the last time that we were there. Since we expect these features to remain consistent, we do not then have to commit additional resources to assessing each of them in detail (Summerfield & Egner, 2009). However, sometimes we can rely too heavily upon our expectations which may lead to perceptual biases, and in some circumstances, misperceptions (incorrect interpretations of sensory information).

The role of expectation in informing perception can be seen as part of a top-down process where prior information is used to modulate the perceptual processing of incoming bottom-up sensory information. Bias, and eventually misperception, can occur when the bottom-up signal is weak or ambiguous, and thus cannot be relied upon. In such cases, the perceptual system must instead rely more heavily upon a strong, and sometimes wrong, top-down signal (Summerfield, Egner, Mangels, & Hirsch, 2006).

There are numerous reasons why sensory information could be described as ambiguous and these include: a lack of attention on the part of the observer, difficult observation conditions (such as attempting to see in the dark or attempting to hear in a noisy environment), or sensory information only being present for a brief period of time (Król & El-Deredy, 2011a). Furthermore, the strength of an expectation, and relatedly, the extent to which it is relied upon to influence perception, also depends upon a number of factors. These factors may include the scope of prior knowledge, and the degree to which inferences can be made about the available sensory information. For example, a number of studies have illustrated that recognition of ambiguous objects is facilitated when they are presented in the correct context, compared to when they are presented in isolation or out of context. When an item is presented in the correct context, the strength of expectation is increased because an observer can draw upon their prior experience of the visual world to guide predictions about the identity of the ambiguous object (for review see Bar, 2004).

Predictions are beneficial because they provide advanced information about an incoming stimulus, due to the processing of that stimulus being partially completed prior to its arrival. By
making predictions, surprise and uncertainty can be reduced as perceptual resources are diverted toward only the unexpected or important features of the environment. This conserves effort and allows additional time to be allocated to the production of behavioural reactions, which, in turn, facilitate environmental interactions and increases overall survival prospects (Kveraga, Ghuman, & Bar, 2007; Schultz, 1998). However, as the example of misperceptions highlights, predictions cannot always be relied upon. This raises questions about the extent to which we rely upon expectation to determine perception. How ambiguous does sensory information have to be before we become more heavily reliant upon prior knowledge, and what is the level of confidence needed in an expectation before it is relied upon to determine perception?

A number of theoretical models exist which have attempted to understand these very questions, namely the extent to which our perceptual experience is determined by top-down predictions (Bar, 2003; Friston, 2005; Grossberg, 1982; Mumford, 1992; Ullman, 1995). Whilst these models do display certain differences, they generally support the notion that top-down predictions are used to aid in the processing of bottom-up information. Being primarily based upon visual processing where there is a hierarchy of processing regions, these models indicate that higher-level regions use prior experience to generate predictions which are projected to lower-level processing regions. These predictions guide the interpretation of an observed input by reducing the number of possible interpretations associated with it. This means that processing can be completed more quickly because fewer interpretations of the sensory data need to be considered. The utility of these top-down predictions is thus affected by the ambiguity of the sensory information. When this information is very clear, very few interpretations of the information need to be considered.

These models also suggest mechanisms for adaptation when discrepancies occur between the top-down predictions offered by these higher-level regions and the incoming bottom-up sensory information. For example, when we see an unfamiliar object placed in a familiar context, our top-down predictions offer a limited number of possibilities about the identity of that object. However, if that same object is placed in an unfamiliar context, it will not be located within the associated list of possibilities and a discrepancy will be observed between what is expected, and what is actually observed. A common feature of the models discussed here, is that they all resolve these discrepancies by attempting to match predictions against actual experience in an iterative fashion. Again, these models display differences in the precise mechanisms by which they resolve discrepancies, but broadly speaking, each model indicates
that when a prediction is offered by a higher-level region it is compared to the available sensory information; if a discrepancy between the two exists, then this discrepancy is projected back to the higher-level area where a new prediction is generated. This process repeats until the discrepancy is eliminated and the observed object can be identified correctly (Kveraga et al., 2007).

Although the models discussed here deal with the influence of expectation and top-down predictions upon sensory perception, Bar (2009b) suggests that the alleged boundaries between perception (the use of the senses to analyse information pertaining to the physical world via the sensory cortex) and cognition (the internal processes that go beyond what is needed to simply perceive the world around us) are artificially constructed. To illustrate this point he uses the example of recognition which he considers unique in that it requires elements of perceptual and cognitive processes in order to be accomplished. Rather than purely examining object recognition from a perceptual perspective and asking “what is this?” Bar instead suggests we should instead view object recognition as asking the question “what is this like?” In this way, object recognition attempts link the observed sensory information, to prior knowledge of similar examples held within memory by means of the cognitive process of association forming. After observing an ambiguous presentation of an object, the brain will attempt to locate an example of a similar object from memory. Even if the two objects are not exactly the same, a prediction can be made about the identity of the currently observed object.

Elsewhere, Bar (2009a) has claimed that, just as physicists can explain complex systems by virtue of a small set of equations, predictions can be seen as a similarly universal principle in explaining how the brain operates. Even when the brain is in its resting state, or what has been termed the brain’s default mode (Raichle et al., 2001), the brain does not wait idly for sensory input, instead the brain acts in a proactive manner continually forming associations and making predictions about the relevant near-future (Bar, Aminoff, Mason, & Fenske, 2007). In this way, whenever we encounter a particular stimulus, our interactions with that stimulus are coloured by our prior experience, since the brain is already comparing it to similar stimuli that we have encountered in the past and attempting to make accurate predictions. Examples of the alleged universality of predictions can be seen in a number of other studies where Bar has illustrated how this framework may be used to explain cognitive processes such as: how we form first impressions (Bar, Neta, & Linz, 2006) and how we recognise threatening stimuli within the surrounding environment (Bar & Neta, 2006).
This thesis aims to assess how expectation and prediction generation influence cognition in terms of our affective experience, specifically, how they affect the development and maintenance of liking. When we discuss liking, we often do so only in terms of the sensory information relating to how an item is perceived at a particular moment in time. Following this line of thought, one would assume that the evaluation of an item is determined solely by its intrinsic properties, and the present state of the person interacting with it. For example, the liking of a specific meal would be determined only by the molecular composition of that meal and the level of hunger present in the person consuming it (Plassmann, O’Doherty, Shiv, & Rangel, 2008). Increasingly, however, as with perception related studies, researchers are acknowledging the influence of both bottom-up, and top-down determinants of affective experience.

A number of studies have illustrated that the presence of a strong expectation has an impact upon the hedonic ratings of certain products, particularly when these products are paired with an external cue where an expectation already exists (for review see Deliza & MacFie, 1996). What is not clear, however, is the extent of expectation’s influence in determining changes in liking. As has already been discussed, Bar (2009b) states that the brain makes predictions by attempting to compare incoming sensory information with prior experience. Even if an exact match for an observed input is not located, a prediction can be generated based upon prior experience with inputs of a similar type. If, as has already been posited, expectation serves to guide the processing of sensory information, then how far can that sensory information deviate from prior experience before the expectation of liking is rejected? Additionally, if the level of liking for a superior or inferior stimulus can be influenced by strong expectations of liking or disliking; will this influence be maintained indefinitely? Even if this is not the case, there is currently little research relating to how long this influence will be maintained for. This thesis hopes to address these questions in an effort to better understand the extent to which affective experiences are influenced by expectation, and whether this influence remains consistent over time.

In order to address these questions, this introductory chapter discusses how we can assess the influence of expectation upon liking, and how we can possibly instigate changes in liking over time, in terms of three main areas:

- Assimilation-Contrast Theory
- Predictive coding and the prediction error hypothesis
- Promoting affective expectations
1.2 Assimilation-Contrast Theory

1.2.1 Defining assimilation and contrast effects

One way in which we can attempt to predict the influence of expectation upon liking is through the use of two complimentary theories: assimilation theory, and contrast theory (R. E. Anderson, 1973; Bohner & Dickel, 2011; Cardello & Sawyer, 1992; Kuenzel, Zandstra, El-Deredy, Blanchette, & Thomas, 2011; Schifferstein, Kole, & Mojet, 1999). Assimilation theory, developed from Festinger's (1957) theory of cognitive dissonance, proposes that, in order to maintain cognitive stability or "consonance", humans are motivated to seek out information which confirms their existing attitudes and expectations. When confronted with new information that contradicts these attitudes and expectations, humans distort their perceptions to avoid the mental discomfort caused by disconfirmation. Therefore, when assessing a particular product, regardless of whether the performance of that product falls above or below expectation, consumers will assimilate performance to the level of expectation.

Alternatively, contrast theory proposes that humans expect a certain level of stimulation from a product, if the performance of that product deviates from their expectation, the perceived difference between the product expected and the product received will be exaggerated. This can have both positive and negative effects upon product perception. If the quality of a product is undersold, consumer satisfaction may be increased by virtue of lowered expectations, and if a product fails to meet expectations the reverse effect is likely to be observed.

The complimentary nature of these theories has given rise to Assimilation-Contrast Theory (ACT) (Hovland, Harvey, & Sherif, 1957) which incorporates both assimilation and contrast theories. The important distinction between ACT and assimilation and contrast theories is that in ACT, the size of the discrepancy between the expected and the actual stimulus is what determines whether it will be assimilated toward expectation or repelled away from it. If the discrepancy is sufficiently small, perception will shift and the discrepancy will be ignored with the stimulus falling into the "latitude of acceptance." If there is a larger discrepancy between the expected and actual stimulus, then the perceived discrepancy will be magnified, and the stimulus will fall into the "latitude of rejection" (see Figure 1.1).
Assimilation and contrast effects have been observed in a number of studies investigating the liking of particular products based upon expectation relating to various extrinsic factors, including: methods of production (Caporale & Monteleone, 2004; Schifferstein et al., 1999; Van Wezemaal et al., 2012) perceived ingredients (L. Lee, Frederick, & Ariely, 2006; Torres-moreno, Tarrega, Torrescasana, & Blanch, 2012) or region of origin (Stefani, Romano, & Cavicchi, 2006). However, as Kuenzel et al. (2011) note, these have predominantly been studies based upon purely categorical differences, for example, whether a particular ingredient is present or not. Therefore, a key part of the current investigation is to attempt to determine the level to which stimuli can deviate from expectation along an interval scale, whilst still remaining within the latitude of acceptance.

1.2.2 Investigating the time course of assimilation effects

As has been discussed already, a common finding from within the related literature is that expectation can influence perception in the form of bias, misperceptions etc. There is also emerging evidence to suggest that expectation can induce similar changes in liking during individual controlled trials when subjects are faced with ambiguous stimuli. For example, L. Lee, Frederick, and Ariely (2006) conducted a study where participants were asked to compare two different beers: a regular beer, and a beer and vinegar solution dubbed "MIT Brew". Pilot
testing had indicated that the presence of the vinegar would be conceptually offensive to most beer drinkers. This suggested that the expectation that MIT brew would taste poorer than regular beer would be enough to alter the perceptual and affective experiences of the participants. In line with the results from the pilot data, the experimenters observed that liking of the MIT brew was significantly reduced when participants were informed of the presence of the vinegar prior to tasting. Interestingly however, the same reduction in preference was not observed when participants were alerted to the presence of the vinegar after tasting. Therefore, it seems that if participants were not initially aware of the difference between the two beers that they sampled, if they developed a preference for the MIT brew, then learning about the potentially offensive ingredient that it contained, did not cause them to retrospectively reinterpret their experience. Whilst this example highlights how expectations can influence and sometimes promote liking in the short term, fewer studies have examined the extent to which expectation induced liking (or disliking) persists over a longer period of time. If these findings are to have any kind of practical application outside of a controlled setting, this is an area that requires further investigation.

A recent study conducted by E. Robinson, Blissett, and Higgs (2013) indicates that, when a strong expectation of liking or disliking exists, hedonic judgements will remain relatively stable over time. In a first experiment, participants rated expected liking for several snack foods, then tasted samples of those foods, with the sample that produced the most disappointing experience being classified as the “disappointing food”. Expected liking for the disappointing food was reduced if a follow-up session was conducted one day later, but it returned to baseline levels if the session occurred one week later. In a second experiment it was revealed that this effect was modulated by product familiarity. Reductions in expected liking after a disappointing experience persisted for a period of one week, for foods that were infrequently eaten. The same effect was not observed for frequently eaten foods, and expected liking reductions after a disappointing experience, returned to baseline levels one week after that experience.

Elsewhere, the results of a study conducted by Tuorila, Cardello, and Lesher (1994) support the claim that liking remains a relatively stable construct over time. They found that the expectation that fat-free crackers would taste poorer than regular-fat crackers, persisted over a period of one month, even when participants were faced with information that disconfirmed this expectation. Like the studies discussed previously, it was found that affective judgements remained stable when there was a strong expectation in place, and that a participant’s expected liking of a product was best predicted by their existing familiarity with the product.
Learning from past experiences with familiar products is difficult, since liking is often associated with particular emotions, and because emotional experiences are episodic in nature they are not well represented in memory (M. D. Robinson & Clore, 2002). As time passes from an individual experience, people tend to forget some of the specific details about that experience and instead rely upon more general knowledge about the event in order to reconstruct how they must have felt about it (Norbert Schwarz & Xu, 2011). Evidence of this phenomenon can be seen in Kuenzel et al. (2010) who found that the emotional associations which affect liking could be conditioned to novel flavours, but not familiar flavours where strong associations already existed. This finding raises questions about how far the expectation of liking can be manipulated. If a consumer already holds a strong expectation of liking about a particular product, how far does their actual experience with that product have to deviate before they reject the expectation of liking? It is also possible that liking could also be influenced by the mere-exposure effect where a preference for a particular stimulus is developed purely by repeated exposure to that stimulus. Regardless of whether the sensory properties of a stimulus are altered for better or worse, would repeated exposure alone be enough for someone to update their existing expectations over time?

An additional question to consider is, if we are able to induce liking via expectation in the face of impoverished or contradictory sensory information, how long will this effect last for? Placebo related literature has suggested that similarly induced biases in other domains (the placebo analgesia effect) can last for a period of several days. Colloca and Benedetti (2006) conducted a study where a group of patients were given an inert treatment, and were told that the treatment would have an analgesic effect. After the treatment was applied, and painful stimulation was administered, the intensity of the stimulation was covertly reduced to promote the belief that the treatment was effective. When painful stimulation was applied again at a later time, this manipulation produced a strong placebo effect which persisted up to seven days later. It is currently unclear, however, whether the maintenance of liking follows a similar pattern to placebo analgesia.

1.2.3 Establishing the need to determine the boundaries of assimilation effects

Whilst manufacturers might prefer not to alter products where a strong expectation of liking already exists, an increasing public awareness of the on-going global problems associated with obesity, means that many manufacturers are under increasing pressure from governments and consumers to make their products healthier (Zandstra, Lion, & Newson, 2016). According to the World Health Organisation (2016), since 1980 worldwide obesity has more than doubled. In
2014 they estimated that over 600 million adults were classified as obese, which represents approximately 11% of men and 15% of women. A further 1.9 billion adults aged 18 years and over could be classed as overweight. In the UK, rates of obesity have increased dramatically in recent times, from 15% in 1993 to 26% in 2014. These figures represent a major public health concern since obesity has been associated with poorer quality of life and premature mortality, and problems associated with obesity have been linked to a number of diseases including: type 2 diabetes, hypertension, cardiovascular disease, and cancer (NHS, 2016).

As consumers display an increasing awareness of the products that they purchase, and Western governments strive to promote healthier lifestyle choices, product manufacturers have had to adapt to these relatively new challenges. For example, in the snack foods industry, the late twentieth century saw a surge in the development and marketing of reduced-fat and fat-free products with research in this area only being outweighed by research into new flavours (Levis & Chambers, 1997). Outside of the external pressures to create healthier products, however, manufacturers may also see additional commercial benefits from the development of healthier products because product intake has been observed to increase when a product is regarded as healthy rather than unhealthy (Gravel et al., 2012; Pechmann & Catlin, 2016; Provencher, Polivy, & Herman, 2009). An additional reason for pursuing research in this area, relates to how the development of healthier versions of existing products can be beneficial for those who are attempting to initiate dietary change. Some dieticians have claimed that people find it easier to adapt to, and maintain, dietary changes when they switch to lower-fat versions of traditionally eaten foods, rather than avoiding these foods altogether (Phelan, Lang, Jordan, & Wing, 2009).

As the results of numerous studies of dietary habits and attitudes have indicated, health related issues (e.g. the calorie content of foods and beverages) are now a primary driver of food choice for many consumers (see Kiszko, Martinez, Abrams, & Elbel, 2014 for review). Yet manufacturers have learned that it is often not sufficient to simply make products healthier. Products must also conform with expectations, and maintain a level of pleasant taste, otherwise sales will decline (Rappoport, Peters, Huff-Corzine, & Downey, 1992; Roininen & Tuorila, 1999). As the study conducted by Tuorila et al. (1994) illustrates, many consumers harbour the notion that healthier alternatives automatically equate to being poorer tasting products. Furthermore, if the sensory attributes of a new product or a new variant of an existing product, are not adequate to confound these expectations, then remembered enjoyment will likely cause the consumer to avoid the product or similar products in the future (E. Robinson, Blissett, & Higgs, 2012). Whilst this is problematic for manufacturers, it is also problematic from a public health
perspective as it has been reported that consumers often see this association as a barrier to eating more healthily (Lloyd, Paisley, & Mela, 1995).

A more recent example of this association is provided by Norton, Fryer, and Parkinson (2013) who observed that expected liking ratings for reduced-fat labelled chocolate were significantly lower than those of regular chocolate, even when the two products did not differ physically. However, whilst the authors observed that reduced-fat labelling of chocolate had a negative impact upon expected liking, they also observed no significant differences in actual liking when chocolate was labelled as reduced-fat compared to when it was not. These results suggest that if the sensory attributes of a healthier version of a product can be sufficiently matched to the familiar non-healthy alternative, then the effects of reduced expectations may be offset by contrast effects.

1.3 Predictive coding and the prediction error hypothesis

1.3.1 Predictive coding of liking

An explanation for why liking can be affected by assimilation and contrast effects can be found in so-called 'predictive coding' accounts of perception, which emphasise the role of expectation in determining what we perceive. The visual object recognition model suggested by Bar (2003) which has been discussed previously, is an example of such an account. According to this model, visual object recognition proceeds in three stages. Firstly, once a visual stimulus has been detected, low spatial frequency (LSF) information is rapidly extracted from the image. This information is then projected to the prefrontal cortex (which is in contrast to bottom-up models of recognition. Such models indicate that all of the features of an input need to be analysed before recognition can occur). Secondly, the rapid extraction of low spatial frequencies activates multiple interpretations in the prefrontal cortex about the identity of the object. Thirdly, these "initial guesses" are back-projected to the inferior temporal cortex where they activate known representations of other objects to be integrated with the rest of the bottom-up information (see Figure 1.2). According to Bar (2003) object recognition typically occurs within 150-200 milliseconds. Since recognition attempts to rapidly acknowledge the identity of a particular object on the basis of limited sensory information and prior expectation, one can easily see why perceptual errors can occur when an object is considered 'close enough' to an existing representation.
Figure 1.2 The principle of predictive coding based upon the extraction of low spatial frequency (LSF) information.
Bar (2007, 2009b) argues that these ideas are not unique to visual processing, and that the same ideas can be adapted to understand other cognitive processes. When a novel input is encountered, not all of the sensory information associated with that input is extracted and analysed, instead broad "gist" information is extracted. Like low spatial frequencies, this is very basic information that is minimally processed, and it is used to create analogies which allow the features of the novel input to be compared to similar existing representations held in memory. Thus, the creation of analogies allows people to make predictions about the relevant near-future because the brain continually anticipates future events and makes predictive judgements about upcoming experiences. For example, if we consider the liking of a particular edible stimulus, that liking may depend on a number of visual, olfactory and taste cues. However, since a rapid judgement is required to determine whether that stimulus is something that is good or bad, all of the information relating to the stimulus cannot be analysed in detail. Instead, any initial assessment of the stimulus will likely be based upon gist information and prior knowledge.

Predictive coding models also provide a potential explanation for why it may be difficult to alter liking for familiar products, since strong analogies already exist between observed inputs and existing representations held within memory. However, models based upon predictive coding frameworks like those suggested by Bar, (2003) or Friston (2005), also indicate that humans adopt a Bayesian inferential approach to shape and refine predictions and expectations by gradually incorporating newly learned information and statistics. If this is the case then it seems reasonable to assume that, if a sufficient amount of information is received that is expectation inconsistent, representations relating to the affective value of familiar products could eventually be updated according to that information (Brown & Brüne, 2012).

1.3.2 Predictive coding as Bayesian inference

Predictive coding accounts indicate that perception can be understood as a continuous process of hierarchical Bayesian inference. Bayes rule is defined thusly:

\[ P(A|B) = \frac{P(B|A) P(A)}{P(B)} \]

In the study of perception, Bayes rule indicates that an observer’s best estimate about the probability of an outcome is based upon both their existing knowledge, and the currently observed sensory information. P(A) represents the prior, (the observer’s expectations) based upon their knowledge about the probability of different outcomes, P(B|A) represents the
likelihood, or the probability of obtaining the currently observed evidence, given a number of potential outcomes, and \( P(A|B) \) represents the posterior, or the current best estimate based upon both the prior and the likelihood. To use an example from the perception literature, several authors have presented participants with images that have been distorted to the point where 2 unique interpretations of the same stimulus are equally probable. Both Hohwy, Roepstorff, and Friston (2008), and Summerfield, Egner, Mangels, and Hirsch (2006) have done this using images of houses and faces. If participants are presented with images of houses and faces, followed by an ambiguous image (potentially house or face), then their interpretation of that image will likely be biased according to whether more houses, or more faces, have been presented beforehand. If the experiment continues and either unambiguous houses or faces continue to be presented with greater frequency, then interpretations of the ambiguous images will continue to be biased in the same direction. However, if in reality the ambiguous image actually depicts an example consistent with the opposite interpretation, and if the image begins to be presented in a manner that is less ambiguous, then participants’ responses will update in light of this new evidence. At first, responses may continue to be biased in the same direction as participants’ expectations, but eventually, once the ambiguous image has been disambiguated sufficiently, it will become increasingly clear that there are many aspects that are inconsistent with the expected interpretation.

1.3.3 Prediction errors – A neurological marker of expectation?

In the study of expectations, some researchers have sought to quantify the effects of expectations through paradigms which have measured prediction errors. According to the prediction error hypothesis, when a discrepancy is detected between a top-down prediction and a bottom-up observation, surprise is generated, and this is highly unwelcome (Bar-Anan, Wilson, & Gilbert, 2009). The discrepancy creates an error signal which is proportional to the magnitude of the difference between what was expected, and what was actually experienced. When the size of the prediction error is substantial enough, the mismatch is sent back to the higher level regions, so that a new prediction can be generated. As the prediction becomes closer to the observed input, the prediction error is reduced. Consequently, the mismatch between the expectation and the input continues to be processed in an iterative fashion, until a final interpretation of the sensory input is reached (Morton, El-Deredy, Morton, Elliott, & Jones, 2011; Summerfield & Egner, 2009).
Prediction errors are encoded by the mesencephalic dopamine system which has long been associated with reward-based learning, and evidence for such an association comes from a number of different sources. Firstly, studies which have investigated the addictive properties of drugs such as amphetamines and cocaine have observed that their effectiveness (the reward that they offer to the user) lies in their ability to enhance dopaminergic function and block dopamine uptake transporters (Koob & Bloom, 1988). A second source of evidence for a link between dopamine and reward-based learning comes from studies involving electrical self-stimulation in rats. These studies have observed that electrical stimulation via electrodes implanted into the brains of rats, is particularly reinforcing when delivered to the neural pathways associated with dopamine neurons. So much so that rats will often forego other rewards in favour of stimulating these areas via a bar pressing task (Olds, 1958; Phillips, Carter, & Fibiger, 1976). A third and final source of evidence for the association between dopaminergic function and reward-based learning, comes from the observation that when animals are treated with neuroleptic drugs, which are known to block dopamine receptors, they exhibit inhibited reward learning and produce greater reaction times in response to food rewards (Wise, 1982). Similar dopamine blocking drugs have also been shown to inhibit the previously discussed reinforcing effects of electrical self-stimulation in animals (Fouriezos & Wise, 1976).

Prior to the mid-1990s, it had been thought that dopamine mediated the feelings of pleasure experienced by an animal upon receipt of a reward (Wise, 1982; Wise, Spindler, DeWit, & Gerberg, 1978). However, the “hedonia” hypothesis has since been replaced by the prediction error hypothesis as the dominant theory concerning dopaminergic function, thanks in part to a series of experiments involving monkeys conducted by Wolfram Schultz and colleagues (for

---

**Figure 1.3** General architecture of a hierarchical predictive coding model. Adapted from Rao and Ballard (1999)
review see Schultz, 1998). First observed by Schultz, Apicella, & Ljungberg, (1993), and Mirennowicz & Schultz (1994), an effect was found where a rewarding stimulus (apple juice) dispensed to monkeys, caused an increase in dopaminergic activity. During these experiments, stimulus dispersal was reliably preceded by an audible tone and initially, neuronal firing in the brains of the monkeys followed stimulus delivery, but not the auditory predictor of stimulus delivery. However, as the monkeys learned that the tone acted as a predictive cue for stimulus delivery, dopamine firing in response to the rewarding stimulus itself returned to baseline levels, and the previously observed spike in dopaminergic activity shifted to coincide with the timing of the predictive auditory cue. Furthermore, if the predictive cue was presented in isolation and rewards were withheld, a reduction in the background level of dopaminergic activity was observed at the time of the expected reward.

Therefore, unlike the hedonia hypothesis that suggested dopamine essentially transformed experience into reward, the prediction error hypothesis indicates that it is the unpredictability of a reward that causes the increase in dopamine firing upon receipt of that reward. When reward delivery remains predictable, and an animal becomes conditioned to expect a reward to follow a predictive cue, then the increase in dopamine firing will move in line with that cue. However, if a reward fails to materialise at the expected time after the cue, or at a level above or below that which is expected, dopamine neurons will encode an error in the prediction of that reward. By encoding the difference between expected and observed outcomes, the animal is able to learn to optimise behaviour to ensure that the delivery of rewards is maximised (Caplin & Dean, 2008; Schultz, 1998).

1.4 Promoting affective expectations

1.4.1 Affective priming

An effect has been observed where the mere observation of an affect-loaded stimulus can activate attitudes from memory, and this can impact upon an individual’s ability to judge subsequently presented stimuli (Hermans, De Houwer, & Eelen, 1994). Fazio, Sanbonmatsu, Powell, and Kardes (1986) investigated this ‘affective priming’ effect and observed that attitudes were activated automatically when participants were presented with strongly affect-loaded priming words. Using pairs of words, they found that when participants were first presented with a prime noun that was either positively (e.g. gift or cake) or negatively (e.g. death or hell) weighted, their judgements of subsequently presented target adjectives were affected by the congruency of the priming word. After presentation of the prime word, participants were asked to recite the prime word aloud, and make a binary choice regarding whether the target
adjective could be considered positive or negative in affect by pressing a button. The investigators found that, in comparison to when neutral words were presented (nonsense strings of three letters e.g. BBB), responses times were facilitated by congruent primes and inhibited by incongruent primes. These results support the assertion that the attitude related to the prime was automatically activated upon presentation of that prime, and this effect has since been replicated outside of simple valenced word pairs in studies involving: pairs of coloured slides (Hermans et al., 1994) valenced words paired with face stimuli (Morton et al., 2011) symbols paired with particular foods (Kuenzel, Blanchette, et al., 2011) and odours paired with particular foods (Hermans, Baeyens, Lamote, Spruyt, & Eelen, 2005).

Affective priming is also often used as a means indirectly assessing implicit attitudes toward items. Rather than explicitly asking a participant about their feelings toward items, attitudes are inferred through their responses toward related stimuli. In a typical affective priming study, participants are presented with a positive or negative prime, and then a target. Attitudes can be inferred by the time taken to evaluate the target object. If the prime and target share the same valence, then quicker responses are observed (Hermans et al., 2005; Hermans, De Houwer, & Eelen, 2001).

1.4.2 Priming affective values and measuring changes in stimulus liking

Whilst affective priming studies initially involved simple visual stimuli, there are now a growing number of studies where affective priming has been used as a means of influencing attitudes toward particular consumable stimuli. For example, Hermans et al. (2005) investigated the liking of originally neutral foods where affect was acquired via odour conditioning. In a first session, participants learned to associate different brands of yoghurt with either positive or negative odours. In a subsequent session where brands of yoghurt were paired with valenced words, the yoghurts were used as primes to activate the associated attitudes. Whilst the yoghurts were originally considered a neutral stimulus, after conditioning, the presentation of the yoghurts activated attitudes relating to the associated odour. During the second affective priming phase, the authors observed that participants were quicker to respond to congruent pairs of stimuli (e.g. yoghurts associated with negative odours paired with negatively valenced words) than incongruent pairs (e.g. yoghurts associated with positive odours paired with negatively valenced words).

In order to prime affective attitudes toward groups of stimuli, the studies presented in this thesis capitalised upon the concept of judgement ‘anchoring’. According to Tversky and Kahneman (1974) stimulus judgements are biased by the value of an initially encountered
stimulus. When the value of subsequently encountered stimuli is acknowledged as being different from the initially encountered value, adjustments are made to correct for this difference, but these adjustments are typically insufficient. Consequently, judgements are biased toward the initially encountered value. In terms of biasing effects, anchors are most effective when (i) they are sufficiently attended to, and (ii) when the value of an anchor is considered sufficiently compatible with a subsequently presented item (Paek, Yoon, & Hove, 2011; T. D. Wilson, Houston, & Brekke, 1996). The current research exploited this fact by first presenting participants with anchor items whose affective value was relatively easy to discern, and then slowly introducing additional items whose affective value was less certain; with the aim of investigating how closely judgements of the uncertain items assimilated toward the values of the anchors.

1.5 The aims of the current investigation

In light of the research discussed in the sections above, the goals of the studies conducted as part of the current investigation can be summarised as follows:

1. To test whether hedonic cues bias liking in the same way that sensory cues have been shown to bias perception.
2. To test how long this placebo-like effect (increased liking induced via prediction) persists.
3. To measure the bounds of this bias, i.e. how far sensory information can deviate from a prediction before the expectation of liking is rejected (contrast).
2 GENERAL METHODS

2.1 Participants
Participants in all studies were healthy volunteers aged between 18 and 49, and studies were advertised through a number of mediums: Posters were placed at various locations around the University of Manchester, and advertisements were placed online using both the Psychological Sciences BSc Psychology and Cognitive Neuroscience Experiment Sign-up System, and the University of Manchester Research Volunteering website. Anyone expressing an interest in participating in one of the studies was first provided with an electronic copy of the associated information sheet to consult beforehand, and testing sessions were arranged either through email, or through the experiment sign-up system.

All studies were approved by the University of Manchester Research Ethics committee. Written informed consent was obtained from all volunteers prior to participation, and the study information was confirmed by the experimenter at the start of each testing session.

Participants were also reminded at the start of all testing sessions that they could withdraw at any time without detriment or explanation.

Participants could be included in the studies if they met the following inclusion criteria:

- Aged 18-50 years
- In good general health
- Normal, or corrected to normal vision
- Fluent in English
- No history of neurological or psychiatric conditions (including epilepsy; depression; schizophrenia, bipolar disorder)
- Had not seen a professional in relation to drug use, alcohol consumption or gambling

2.2 Questionnaire measures
Most of the studies did not employ questionnaire measures, but one questionnaire measure, the Life Orientation Test Revised (LOT-R) (Scheier, Carver, & Bridges, 1994), was administered during the experiment described in chapter 4.
2.2.1 The LOT-R
The LOT-R (see Appendix B.2) is a 10 item measure of generalised optimism versus pessimism, and each item is a statement such as ‘I’m always optimistic about my future’. For each item participants have to state the extent to which they agree with a statement, using a five point scale ranging from ‘I agree a lot’ to ‘I disagree a lot’. Three of the items are positively scored, 3 are reverse scored, and 4 are filler items. Responses are coded with higher LOT-R scores corresponding to higher levels of generalised optimism.

2.3 Experimental behavioural paradigms
A comprehensive description of the behavioural paradigms used across the programme of study is provided within each of the experimental chapters (chapters 3 – 7). For each study, the information sheets provided to participants before attending the experimental sessions offered a broad outline of the tasks that would be used, and comprehensive instructions were programmed into the start of all computerised experimental tasks. An experimenter was always on hand at the start of experimental sessions to provide further verbal instructions, and to ensure that participants had fully understood those instructions. Additionally, practice trials were included at the start of the majority of experimental tasks so that participants could become accustomed to the demands of the tasks.

In chapters 3-6, computerised tasks were used to collect explicit measures of attitudes toward groups of stimuli: attractive and unattractive faces (chapter 3), happy and unhappy emotional faces (chapter 4), food items of varying pleasantness (chapter 5), and smooth and sharp abstract patterns (chapter 6). In chapter 7 computerised tasks were instead used to collect data about the perceived length of time intervals of varying lengths.

The experimental tasks described in chapter 3 were written in E-Prime (Psychology Software Tools, Pittsburgh, PA), and all of the tasks featured in the remaining experimental chapters were written in PsychoPy (Peirce, 2007). The majority of experimental tasks were performed in a dimly lit testing cubicle within the Zochonis building at the University of Manchester. Participants were seated at a desk in front of a PC attached to either a 23-inch or 17-inch monitor, and responses were collected via keyboard inputs or mouse clicks. The only exception was the study described in chapter 4, which was conducted in a lecture theatre, where data was collected from multiple participants at once. In this study, stimuli were presented on a projector.
screen at the front of the theatre, and participants responded using a pen and a printed response sheet.

2.4 Implicit attitude measures

In addition to the behavioural tasks described in section 2.3, which were used to measure explicit attitudes toward groups of stimuli, standardised implicit attitude measures were used during chapters 5 and 6. In chapter 5, implicit attitudes toward images of foods were collected using the Go/No-Go Association Task (GNAT) (Nosek & Banaji, 2001); and in chapter 6, implicit attitudes toward groups of abstract patterns were collected using the Implicit Association Task (IAT) (Greenwald, McGhee, & Schwartz, 1998).

2.4.1 The IAT

The IAT is a computerised task which measures implicit attitudes toward 2 competing categories of items. Throughout the task participants are presented with either: positively or negatively valenced words (attribute items), exemplars of the 2 competing categories (target items), or a combination of both words and exemplars. In each trial, labels relating to the words or the category exemplars are presented in the top-left, and top-right corners of the screen; and participants are expected to classify items which appear in the centre of the screen (words or category exemplars) as quickly as possible according to those labels. This is normally achieved by instructing participants to respond using keys at opposite ends of the keyboard.

The IAT is based around a 7 block design with certain blocks functioning as practice blocks and certain blocks functioning as test blocks (see Table 2.1). In the first 2 practice blocks participants are only expected to classify items of a single type, positively/negatively valenced words, or exemplars from category A/category B. In the test blocks, items of both types (words and category exemplars) are presented, and the labels in either corner of the screen are combined to classify one set of words, or one set of category exemplars, e.g. left label: positive words or exemplars from category A; right label: negative words or exemplars from category B. In the latter half of the task this procedure is repeated, but the word type/category exemplar pairings are reversed.
Table 2.1. Trial Sequence for the standard IAT design adapted from the Bush vs. Gore design described in Greenwald, Nosek, and Banaji (2003).

<table>
<thead>
<tr>
<th>Block</th>
<th>Trials</th>
<th>Function</th>
<th>left response key</th>
<th>right response key</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>Practice</td>
<td>Category A items</td>
<td>Category B items</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>Practice</td>
<td>Pleasant Words</td>
<td>Unpleasant Words</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>Practice</td>
<td>Pleasant + A</td>
<td>Unpleasant + B</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>Test</td>
<td>Pleasant + A</td>
<td>Unpleasant + B</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>Practice</td>
<td>Category B items</td>
<td>Category A items</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>Practice</td>
<td>Pleasant + B</td>
<td>Unpleasant + A</td>
</tr>
<tr>
<td>7</td>
<td>40</td>
<td>Practice</td>
<td>Pleasant + B</td>
<td>Unpleasant + A</td>
</tr>
</tbody>
</table>

Implicit attitudes toward the target items are measured using response latencies. If attitudes toward the target items from each category (exemplars) are differentially associated with the positive and negative attribute items (words), then participants should find that one version of the combined label task is considerably easier, resulting in reduced response latencies for this condition.

In chapter 6, IAT scores for each participant in both experiments were computed based upon the improved IAT scoring algorithm described by Greenwald et al. (2003). (i) Response latencies above 10,000ms were removed, and participants with more than 10% of latencies below 300ms were also removed (this did not apply to any of our participants). (ii) The mean latency for blocks 3, 4, 6 and 7 was then computed, and (iii) a pooled SD was computed for blocks 3 and 6, and blocks 4 and 7. (iv) Each error latency within blocks 3, 4, 5 and 6 was replaced by the mean of the corresponding block, and then (v) an amended overall mean latency was computed for each block. (vi) The amended mean latency for block 3 was subtracted from the amended mean latency from block 6, and the same procedure was applied to blocks 4 and 7. (vii) The two values calculated in step 6 were then divided by their associated pooled SDs calculated in step 3. (viii) Finally, the two values calculated in step 7 were averaged to create an IAT score. This resulted in a value that ranged between -1 (indicating more positive implicit attitudes toward category A) and 1 (indicating more positive implicit attitudes toward category B).
2.4.2 The GNAT

The GNAT is a response inhibition task derived from the IAT, however unlike the IAT which requires 2 competing categories of target items, the GNAT allows an experimenter to assess the implicit association between attribute items, and target items from a single category. In the standard GNAT design participants are asked to quickly respond to items belonging to the target category (Go trials), and to not respond to distractor items (No-Go trials). This is usually achieved by the participant pressing, or failing to press, a specified key on the keyboard.

A single GNAT consists of a series of practice trials followed by two blocks which are visually similar to the IAT (see table 2.2). In the first block, after the practice trials, participants are presented with target items, distractor items, and positive/negative attitude items. If, for example, one was interested in attitudes toward dogs, then dogs would be the target items, cats could act as distractor items, and the attitude items would be positive or negative words. In block 1, participants would be asked to respond to only images of dogs and positive words, and the labels ‘dog’ and ‘positive’ would be presented in the upper corners of the screen throughout the block. In each trial, an item would appear in the centre of the screen, and that item could be either: one of the target items, one of the distractor items, or one of the positive/negative attitude items. If the participant responded when an image of a dog or a positive word was presented, then their response would be classed as a hit (correct), if the participant failed to respond to either of these items their response would be classed as a miss (incorrect). Conversely, if the participant responded when an image of a cat or a negative word was presented, then their response would be classed as a false alarm (incorrect), and if they withheld their response when either of these items was presented, then their response would be classed as a correct rejection (correct).
Table 2.2  

<table>
<thead>
<tr>
<th>Block</th>
<th>Trials</th>
<th>Function</th>
<th>Go Trials</th>
<th>No-Go Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>Practice</td>
<td>Dogs + Pos words</td>
<td>Cats + Neg Words</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>Test</td>
<td>Dogs + Pos words</td>
<td>Cats + Neg Words</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>Practice</td>
<td>Dogs + Neg words</td>
<td>Cats + Pos Words</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>Test</td>
<td>Dogs + Neg words</td>
<td>Cats + Pos Words</td>
</tr>
</tbody>
</table>

Performance in the GNAT can be measured by two dependant variables: sensitivity and response latency. Like signal detection theory (SDT), sensitivity in the GNAT is measured by analysing correct, and incorrect responses, and indexed by d-prime (d'), a measure of how well participants are able to discriminate targets (signal) from distractors (noise). Alternatively, response latency can be used to assess how quickly participants are able to respond to target items drawn from a single category (e.g. dogs), when they are paired with positive and negative attitude items. Performance in either measure should be improved (increased sensitivity or reduced response latency), when participants are asked to respond to target items paired with attitude items that are strongly associated, rather than weakly associated.

In chapter 5, performance in the GNAT was calculated using response latencies, following the guidelines outlined by Nosek and Banaji (2001) and Greenwald, Nosek, and Banaji (2003). For each participant, data from incorrect response trials, and distractor trials were removed first. Data were then checked to see whether any participant displayed unusually high error rates (above 40%), or unusually fast responses (less than 300ms). After the data were checked, the average response latency for each participant was calculated for each of the critical blocks, and a $D$ value (comparable to Cohen’s $d$ see Greenwald et al., 2003) was calculated by dividing the average difference in reaction time across blocks by the standard deviation across blocks. Higher $D$ values indicated a stronger association between the target items and the positive attitude items.
2.5 The Psychometric Function

Psychometric functions are a tool commonly used within perceptual research to relate responses from a psychophysical task (e.g. reporting whether a stimulus is more or less intense than remembered) to the physical properties of a stimulus (e.g. actual stimulus intensity in each trial). Data from the task are plotted at various points corresponding to the proportion of correct responses (along the y axis) at each level of presented stimulus intensity (along the x axis). A psychometric function can then be fitted, based upon these data points (see Figure 2.1).

![Psychometric Function Diagram]

Figure 2.1 Example of a psychometric function

The functions fitted to participant data in chapters 3, 5 and 6 of this thesis were all fitted using Matlab (The MathWorks, Natick, MA), and the Palamedes toolbox for Matlab (Prins & Kingdom, 2009). Prior to fitting each function, estimates of a number of parameters were specified including the threshold ($\alpha$), and the slope ($\beta$) of each function. The threshold represents the point at which the proportion of correct responses reaches some predetermined level, and the threshold of interest may vary according to the objectives of the study. For example, an experimenter may wish determine the level of stimulus intensity required so that it is perceived correctly 75% of the time, in order to calculate a measure of just noticeable difference (JND). For the experiments discussed within this thesis, the threshold of interest was the point of subjective equality (PSE), this is the 0.5 point along the y axis, or (in the case of the example above) the point at which correct/incorrect responses are given 50% of the time (see Figure 2.2). The slope or gradient of the function is a measure of a participants’ sensitivity to changes...
in the stimuli. Steeper slopes indicate that a participant discriminated consistently between the presented stimulus intensities, and the opposite is indicated by shallower slopes (Herbert, Bertenshaw, Zandstra, & Brunstrom, 2014). Broad estimates of general PSE and slope were used to guide the fitting process, but both parameters were free to vary during the fitting process.

**Figure 2.2**  Example of a psychometric function including estimates of both PSE ($\alpha$), and slope ($\beta$)

Two additional parameters were specified prior to the fitting process: the guess rate ($\gamma$), and the lapse rate ($\lambda$). The guess rate represents the probability that a response would be correct if an observer were to guess in a particular trial. The lapse rate represents the proportion of trials where responses are provided independently of the presented stimulus intensity (e.g. if a participant was not paying attention.).

Kingdom and Prins (2010) state that there are five different types of psychometric function that can be fitted to data: Cumulative Normal; Logistic; Weibull; Gumbel; and Hyperbolic Secant. All of these functions result in a similar sigmoidal shape, and estimates of threshold and slope will also be relatively similar for all types.
For the studies described in chapters 3 and 6, logistic functions were fitted to the data:

\[ F_L(x;\alpha,\beta) = \frac{1}{1 + \exp(-\beta(x - \alpha))} \]

Since lapse rate and guess rate were not of interest in any of the studies, both parameters were fixed to conservative estimates close to 0. In line with the guidelines provided by Kingdom and Prins (2010). These same guidelines also influenced the design of the studies, and the number of iterations needed within each experimental condition to acquire a reasonable estimate of PSE and slope.

For each of the studies discussed in chapters 3, 5 and 6, the value along the x axis of a psychometric function represents the affective value of the stimulus that was being presented, and the value along the y axis represents the proportion of liked responses. Consequently, the PSE represents the affective value of a given stimulus, required for a participant to shift from providing a higher proportion of disliked responses, to providing a higher proportion of liked responses.

### 2.6 Statistical analysis

All statistical analysis was performed using SPSS for Windows version 22 (SPSS Inc., Chicago, IL, USA). In the majority of cases, data were analysed using a univariate repeated measures analysis of variance (ANOVA), and planned and post-hoc analyses were performed using t-tests. In order to avoid potential type I errors, Bonferroni corrections were applied when multiple comparisons were used. In chapter 4, data were analysed using a repeated measures analysis of covariance (ANCOVA), to assess the influence of variables which could not be controlled as part of the experimental manipulation (generalised optimism, and participant bias). In chapter 6, since participants were recruited in 2 different groups which were differentially cued, but who performed the same basic group of tasks, data were analysed using a mixed design ANOVA. Further details of the analysis performed for each study are provided in the data analysis sections of the experimental chapters.
3 INVESTIGATING THE INFLUENCE OF AFFECTIVE EXPECTATIONS USING ATTRACTIVE AND UNATTRACTIVE FACES

3.1 Introduction

The idea of the predictive brain is becoming increasingly influential for researchers interested in cognition, perception, and decision making. The implication of this idea is that the brain is not an organ that sits by idly waiting for input, instead it acts in a proactive manner continually forming associations between incoming sensory information and prior experience (Kveraga et al., 2007). Indeed the generation of predictions is now being suggested as the core function of the brain, with some asserting that at the most basic level, the brain can be considered as essentially being a 'prediction machine' (Clark, 2013). When predictions are being discussed in this sense they do not refer to longer-term prospections about the likelihood of events occurring in the far distant or even relatively near future. Instead the term predictions is used to refer to what Panichello, Cheung, and Bar (2012) describe as, “Expectations about the immediate sensory environment based upon previous experience and learning.” (p.1).

Predictive coding theory posits that perception is a process of active inference, constructed through the formation of top-down predictions, and the computation of prediction errors (Huang & Rao, 2011; Rao & Ballard, 1999). Our brains continually capture information about the rules and regularities which govern our world, and this information is used to build and store internal models, or representations, of various world states. When we are placed in a situation that is the same or similar to one that has been encountered before, these internal representations are used to make predictions. During perception these predictions are refined in order to minimise prediction errors which represent the difference between top-down predictions, and bottom-up sensory experience (Barrett & Bar, 2009; Friston, 2003; Friston, Mattout, & Kilner, 2011; Schultz & Dickinson, 2000).

There are immediate benefits from this type of future-oriented thinking; surprise and uncertainty are unwelcome, and can lead to inefficient responding which is costly in many situations. By pre-empting what is likely to be experienced in the near future, we can attempt to reduce uncertainty, and can thus respond more quickly to meaningful stimuli in the surrounding environment like threats or rewards (Brown & Brüne, 2012). Furthermore, predictive processing offers the added benefit of promoting efficiency by reducing the cognitive resources required to process expected inputs. Expectations constrain the number of possible interpretations of a given input to what is possible or probable within a given scenario. As such, only the
unexpected features of an input need to be processed in detail, while the features which
conform to prior expectations do not require the same level of processing because internal
representations can be used as seemingly reliable estimates of those features (Kveraga et al.,
2007; Mumford, 1992; Summerfield & Egner, 2009).

The construction of perception through prediction is not without drawbacks, however, and
perceptual biases can sometimes emerge as a result of this need for fast and efficient
processing. Such biases may be considered understandable if they are based upon knowledge
that has been accepted as accurate, and has thus been internalised over a lifetime of
experience; yet several studies have shown that similar biases can also arise from knowledge
that has only been acquired relatively recently. Studies involving bistable images (images which
have two unique interpretations) for example, have observed that the presentation of
semantically related primes before an image can bias perception toward a particular
interpretation of that image (Balcetis & Dale, 2007; Bugelski & Alampay, 1961; Goolkasian &
Woodberry, 2010), while binocular rivalry studies have shown that perception can be
dominated by one stimulus, if that stimulus is presented with greater frequency beforehand
(Chopin & Mamassian, 2012).

In some cases it has been shown that recently acquired knowledge can even override strongly
held longer-term biases. One example of such a bias is the ‘light from above’ prior, which refers
to the fact that inferences about the depth of ambiguously shaded two-dimensional images are
ordinarily based upon the assumption that light comes from above (Sun & Perona, 1998).
Adams, Graf, and Ernst (2004) showed that this bias could be reversed though experience with a
task which contradicts this assumption. It is perhaps, logical that this should occur, given that
predictions represent the current best attempt to model the world, and to remain accurate,
those models need to be continually updated in light of new information.

The reliance upon existing knowledge can sometimes go even beyond bias, and can instead lead
to perceptual errors or misperceptions. It is thought that misperceptions occur as the result of a
strong top-down signal – existing knowledge, and an impoverished bottom-up signal –
ambiguous sensory information (Król & El-Deredy, 2011a; Summerfield et al., 2006). An
illustration of such a phenomenon is provided by Gregory (1997) who notes that when
individuals observe a face painted onto the inside of a hollow mask, they will report seeing the
mask as a normal convex face. The reason for this illusion is that viewers are attempting to infer
details about the properties of the three dimensional world from the limited flat images
provided by the eyes. Various visual cues like shading and shadow might signal that the mask is
hollow if they are considered, but in order to facilitate survival, object recognition must be performed quickly. Therefore, in the case of the mask, viewers instead infer its properties based upon the assumption that most faces are convex.

As these examples highlight, perceptual biases often arise from the need to process information quickly, and an increased reliance upon prior knowledge. However, the extent of this bias may also be determined, at least in part, by how closely a stimulus adheres to the prior expectation.

ACT (Hovland et al., 1957) proposes that if a given stimulus adheres fairly closely to an expectation, assimilation will occur. Perception will shift, and small discrepancies will fall into the ‘latitude of acceptance’, where any conflict with the judge’s standards is relatively minor (Davidenko et al., 2015). Alternatively, when larger discrepancies occur between expectation and experience, they instead fall into the ‘latitude of rejection’ and produce contrast effects. Such discrepancies cause surprise and uncertainty, and in cases where the discrepancy between expectation and experience is so large that it must be acknowledged, contrast effects magnify the extent of the discrepancy.

ACT has been explored primarily within the product design and marketing literature, and studies within this area have attempted to investigate these effects by adjusting the properties of various products like the saltiness of foods (Dijksterhuis, Boucon, & Le Berre, 2014; Herbert et al., 2014) or the sweetness of drinks (Kuenzel, Zandstra, et al., 2011; Woods, Poliakoff, Lloyd, Dijksterhuis, & Thomas, 2010) without a user’s knowledge. Small deviations from the expected product appear to be ignored due to the assumption that the product will remain homogenous, however, at a certain point the discrepancy between the received product and the expected product becomes too great to be masked by this assumption. It has been argued that, if this happens repeatedly, the effect of expectations will diminish as the consumer reconsolidates their memories of a product to incorporate these inconsistencies (J. L. C. Lee, 2009). This leads more specifically to the questions that the current study aims to address. If strongly held expectations can lead to perceptual biases, do expectations also lead to affective biases and misjudgements, and can the boundaries of these effects be measured?

The influence of top-down processes over our affective reactions has previously been acknowledged by the Affective Expectation Model (AEM) which claims that affective reactions are driven by how people expect to feel about an experience or stimulus. This expectation promotes a state of ‘affective readiness’ where an individual is primed to accept a stimulus of a certain affective value (T. D. Wilson, Lisle, Kraft, & Wetzel, 1989). This state of affective readiness means that, just as similar-to-expected sensory information may result in sensory
biases, if the affective value of a stimulus or event adheres closely to an expected value, then minor discrepancies between the two may be ignored (assimilation effects). Alternatively, if the difference between the expected affective value and the received affective value is so great that it is consciously acknowledged, then affective reactions will instead veer away from the affective expectation (contrast effects) (Geers & Lassiter, 1999, 2005).

The aim of the experiments presented in this chapter was to test whether predictive cues could be used to promote a state of affective readiness, where participants’ affective responses to stimuli fell in line with their expectations. To achieve this predictive cues were paired with a series of attractive, unattractive, and moderately attractive faces. The studies discussed within this chapter examined facial attractiveness judgements for several reasons: firstly, because humans find the observation of attractive faces rewarding (Aharon et al., 2001), participants were expected to prefer more attractive faces, and secondly, because attractiveness ratings are usually based upon certain specific factors e.g. facial symmetry (Langlois et al., 2000), a sense of consensus was expected in terms of which faces were considered the most and least attractive.

Based upon the literature relating to perceptual biases and the AEM, it was hypothesised that responses to moderately attractive faces would be more positive if they were paired with a cue promoting a positive state of affective readiness (attractively associated cue), rather than a cue promoting a negative state of affective readiness (unattractively associated cue). This chapter presents results from 2 experiments which were designed to test this hypothesis.

In experiment 1 a series of composite morphed face images were created by combining images of attractive and unattractive faces. At the beginning of the experiment participants were shown the unaltered attractive and unattractive faces, and face type (attractive/unattractive) was reliably predicted by 2 unique cues. Over time the composite morphs were introduced, which were paired with either the attractively associated, or the unattractively associated cues, and changes in attractiveness ratings for the morphed presentations were investigated according to cue type.

In experiment 2, attractive and unattractive faces were again associated with predictive cues, but rather than introducing composite morphs over the course of the experiment, sets of moderately attractive faces were introduced which could be paired with the attractively associated, or the unattractively associated cues. Like experiment 1, this experiment was conducted to investigate changes in affective ratings for the moderately attractive stimuli, based upon expectations about the cue/stimulus pairings.
Experiment 1

3.2 Methods

3.2.1 Participants
Sixteen participants aged between 21 and 29 (12 females, mean age = 24, \(SD = 2.75\)) were recruited for experiment 1 from the staff and post-graduate population at the University of Manchester. All participants were paid for their participation in the study.

All participants had normal or corrected-to-normal vision, and no reported history of neurological or psychiatric disorders. All testing procedures were approved by the University of Manchester ethics committee.

3.2.2 Stimuli and apparatus
Experiment 1 used a set of 126 black and white images of male and female faces with neutral expressions. Twenty-eight of the images were taken from 2 emotional face databases: The Karolinska directed emotional faces database (KDEF) (Lundqvist, Flykt, & Öhman, 1998), and the Max Planck FACES database (Ebner, Riediger, & Lindenberger, 2010). The selected images were comprised of: 7 attractive males, 7 unattractive males, 7 attractive females, and 7 unattractive females. The attractive, and unattractive faces, were selected based upon attractiveness ratings collected in a short preliminary study with a different group of participants (see Appendix A.1).

Each of the attractive faces was paired with an unattractive face of the same gender to create 14 attractive/unattractive face pairs (see table 3.1). For the remainder of this chapter the attractive face in each pair will be referred to as ‘parent face A’, and the unattractive face will be referred to as ‘parent face B’.

In addition to the attractive/unattractive parent faces, the stimulus set also included 98 composite morph images. If parent face A is thought of as 0%, and parent face B as 100%, the composite morph images were created at the following increments between the two: 30%, 40%, 45%, 50%, 55%, 60%, and 70%. The experimental manipulation was based on the assumption that, as the percentage of morphing increased, attractiveness ratings should decrease, and the increments used to create the composite morphs were based upon the findings of another short preliminary study which had been conducted beforehand. The purpose of this study had been to establish where the identity boundaries in a composite morph shifted from one face to another (see Appendix A.2).
Table 3.1  Attractive / unattractive face pairs from experiment 1

<table>
<thead>
<tr>
<th>Gender</th>
<th>ID</th>
<th>Rating of attractive face</th>
<th>Rating of unattractive face</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>M1</td>
<td>( M = 6.27, \text{SD} = 1.29 )</td>
<td>( M = 2.45, \text{SD} = .81 )</td>
</tr>
<tr>
<td>Male</td>
<td>M2</td>
<td>( M = 5.35, \text{SD} = 1.57 )</td>
<td>( M = 2.60, \text{SD} = 1.26 )</td>
</tr>
<tr>
<td>Male</td>
<td>M3</td>
<td>( M = 5.18, \text{SD} = 1.16 )</td>
<td>( M = 2.57, \text{SD} = 1.04 )</td>
</tr>
<tr>
<td>Male</td>
<td>M4</td>
<td>( M = 5.45, \text{SD} = 1.60 )</td>
<td>( M = 2.85, \text{SD} = 1.08 )</td>
</tr>
<tr>
<td>Male</td>
<td>M5</td>
<td>( M = 5.43, \text{SD} = 1.44 )</td>
<td>( M = 2.60, \text{SD} = 1.07 )</td>
</tr>
<tr>
<td>Male</td>
<td>M6</td>
<td>( M = 5.88, \text{SD} = 1.26 )</td>
<td>( M = 2.28, \text{SD} = .97 )</td>
</tr>
<tr>
<td>Male*</td>
<td>M7</td>
<td>( M = 5.53, \text{SD} = 1.14 )</td>
<td>( M = 2.80, \text{SD} = 1.15 )</td>
</tr>
<tr>
<td>Female</td>
<td>F1</td>
<td>( M = 5.40, \text{SD} = 1.54 )</td>
<td>( M = 2.63, \text{SD} = 1.50 )</td>
</tr>
<tr>
<td>Female</td>
<td>F2</td>
<td>( M = 5.98, \text{SD} = 1.59 )</td>
<td>( M = 3.30, \text{SD} = 1.27 )</td>
</tr>
<tr>
<td>Female</td>
<td>F3</td>
<td>( M = 6.80, \text{SD} = 1.33 )</td>
<td>( M = 3.90, \text{SD} = 1.19 )</td>
</tr>
<tr>
<td>Female</td>
<td>F4</td>
<td>( M = 6.55, \text{SD} = 1.12 )</td>
<td>( M = 3.50, \text{SD} = 1.19 )</td>
</tr>
<tr>
<td>Female</td>
<td>F5</td>
<td>( M = 5.60, \text{SD} = 1.39 )</td>
<td>( M = 2.68, \text{SD} = 1.16 )</td>
</tr>
<tr>
<td>Female</td>
<td>F6</td>
<td>( M = 5.95, \text{SD} = 1.46 )</td>
<td>( M = 3.55, \text{SD} = 1.12 )</td>
</tr>
<tr>
<td>Female*</td>
<td>F7</td>
<td>( M = 5.38, \text{SD} = 1.43 )</td>
<td>( M = 3.50, \text{SD} = 1.45 )</td>
</tr>
</tbody>
</table>

Note:
* Denotes face pairing used only in main experiment

Grand mean for attractive faces in preliminary study 2: 5.82, SD = 1.46
Grand mean for attractive faces in main experiment: 5.77, SD = 1.44
Grand mean for unattractive faces in preliminary study 2: 2.91, SD = 1.20
Grand mean for unattractive faces in main experiment: 2.94, SD = 1.22

The stimulus set also included images of 2 symbols taken from the Vulcan language (A language created for the American science-fiction series Star Trek). The face images were included to act as target stimuli, and the symbols were included to act as predictive cues at the beginning of each trial. Images of the symbols were cropped to the same size as the face images (562 x 762 pixels).

The experiment consisted of 2 parts: an online task, and a lab-based task. The online portion of the task was coded in JavaScript, and was designed to be run in a web browser. The lab-based task was coded in E-prime software (Psychology Software Tools, Pittsburgh, PA), and was run on a computer with a 17-inch monitor at a resolution of 1280 by 1024 pixels. A standard UK layout keyboard was used to collect responses.
3.2.3 Design and procedure

Participants completed the online task first. The task involved presentations of each of the 28 parent faces paired with one of the 2 symbols, and after each presentation, participants were asked to provide an attractiveness rating. Before a participant was invited to attend the lab-based session, their responses from the online task were examined to ensure that they had learned the learned the face/symbol pairings.

The lab-based session was conducted within 7 days of a participant completing the online task, and it primarily involved presentations of morphed versions of the face stimuli seen during the earlier task. Like the online task, each face was paired with a symbol, and participants were again asked to provide an attractiveness rating after each presentation.

Online task

The online portion of the study was accessed from a web-link distributed to participants via email. At the beginning of the task participants were informed that they would be shown faces paired with novel symbols, and that they would be asked rate those faces in terms of attractiveness. Participants were instructed to remember the face/symbol pairings since their memory of those pairings would be tested at the end of the task. Throughout the task one symbol was consistently paired with attractive faces, while the other symbol was paired with unattractive faces. Face/symbol pairings were counterbalanced across participants.

The task included 56 rating trials (Two presentations of each of the 28 parent faces), and trials consisted of: a fixation cross (500ms), a symbol presentation (800ms), a face presentation (1200ms), and a response screen. On the response screen, participants were prompted to provide an attractiveness rating for the preceding face (between 1 and 9) using on-screen buttons, and the next trial did not begin until a response was provided.

After completing the rating trials, participants completed a further 5 trials to test their memory of the face/symbol pairings. These additional trials involved a random sample of 5 of the faces shown in the rating trials. Each face was presented briefly (1200ms), and participants had to confirm which symbol the face had originally been paired with. After completing this additional task, responses were stored in a data file located on a University server, and participants were notified that they would be contacted by the experimenter within 7 days.

Lab-based task

The lab-based task was divided into 2 parts. The first part of the task was a condensed version of the rating task that participants had completed online previously. This part consisted of 28
trials (1 presentation of each parent face), presentation times were identical to those used in
the online task, and each face was paired with the same symbol used during the earlier online
task. The reason for including this task was twofold: firstly it was designed to remind
participants of the face/symbol associations, and secondly, it served as a means of measuring
the consistency of each participant’s ratings.

The second part of the task was where the morphed face presentations were introduced, but
participants were not informed about their inclusion. Instead they were told that they would be
viewing presentations of the same face/symbol pairings that they had already seen, but each
pairing would only be presented briefly. Participants were instructed to respond as quickly as
possible since the experiment was designed to assess how quickly they could acknowledge and
rate each face.

Trials consisted of: a fixation cross (500ms), a symbol presentation (500ms), a morphed
composite face (150ms), and a response screen (see Figure 3.1). The results of a preliminary
study conducted previously (see Appendix A.2) had revealed that composite morphs were
reliably identified as representing a particular parent face when: (i) the morph was based upon
60% or more elements drawn from a that face, and (ii) when the morph was presented for at
least 500ms. Elsewhere, neuroimaging research has indicated that individual faces can be
discriminated from one another approximately when they are presented for durations of 150ms
or more (Jacques & Rossion, 2006; D. I. Perrett, Rolls, & Caan, 1982). In the present study,
composite morphs were presented for only 150ms to increase the difficulty of the task, and to
promote the use of the predictive cues.

Each composite presentation contained a variable percentage of an attractive/unattractive face,
and each presentation was rated in terms of attractiveness (1-9). Responses were collected
using the keyboard, and a response was required in every trial. Breaks were provided every 84
trials, and the number of remaining trials was presented at each break point.
Composite morphs at 30% and 70% were used to anchor participant responses and were always presented with the same cues (30% morphs were presented with the attractively associated cue ‘cue A’, and 70% morphs were presented with the unattractively associated cue ‘cue B’), the remaining increments (40%, 45%, 50%, 55% and 60%) were presented with each cue an equal number of times. This resulted in 672 trials for this part of the task: 4 repeats of the 30% and 70% morph cue pairings from the 14 face pairs (112 trials), 4 repeats of the remaining increments with cue A (280 trials), and 4 repeats of the remaining increments with cue B (280 trials). The symbols associated with attractive and unattractive faces were counterbalanced across participants, but remained consistent for each participant across both the online task, and the lab-based task.

3.2.4 Data analysis

Paired samples t-tests were performed on the data from the online task, and the first part of the lab-based task, to confirm that there was an overall difference in attractiveness ratings between the attractive and unattractive parent faces.

Data from the 30% and 70% presentations were analysed using a two-way, face pair (14 levels) x morphing level (30%, 70%) repeated measures ANOVA. These presentations were analysed
separately from the rest of the morphed presentations because the 30% and 70% morphs were intended to anchor responses, and as such, they were always paired with the same cues.

The remaining morphed presentations were analysed using a three-way, face pair (14 levels) x cue (attractive, unattractive) x morphing level (40%, 45%, 50%, 55%, 60%) repeated measures ANOVA.

Reaction time data was analysed using three-way, cue congruency (congruent, incongruent) x morph (40%, 45%, 55%, 60%) x face pair (14 levels) repeated measures ANOVA. Congruent presentations were defined as the 40% and 45% morph presentations where the attractive cue was used, and the 55% and 60% morph presentations where the unattractive cue was used. The opposite was true for the incongruent presentations. Reaction times are reported in milliseconds.

In cases where Mauchly’s test revealed that the assumption of sphericity had been violated, the degrees of freedom have been corrected using Greenhouse-Geisser estimates of sphericity. An alpha level of .05 has been used for all statistical tests, and effect sizes are reported here as partial eta squared ($\eta^2_p$) where 0.05, 0.1, and 0.2 correspond to small, medium, and large effects, respectively (Cohen, 1988).

### 3.3 Results

#### 3.3.1 Attractiveness ratings for parent faces

Based upon responses from the online task and the first part of the lab based task, the attractively grouped parent faces were considered significantly more attractive than the unattractively grouped parent faces (attractive group: $M = 5.70$, $SD = 1.11$; unattractive group: $M = 3.20$, $SD = .95$), $t(15) = 10.16, p < .001$.

Further paired samples t-tests, performed on parent faces of each gender separately, revealed similar differences between the groupings for both the male faces (attractive males: $M = 5.64$, $SD = .96$; unattractive males: $M = 2.84$, $SD = 1.03$), $t(15) = 10.26, p < .001$, and the female faces (attractive females: $M = 6.04$, $SD = .89$; unattractive females: $M = 3.63$, $SD = 1.01$), $t(15) = 14.98, p < .001$.

#### 3.3.2 Attractiveness ratings for 30% and 70% anchors

A main effect of morphing level was observed from the analysis of the 30%/70% presentations indicating that participants found the 30% faces ($M = 5.43$, $SD = 1.63$) more attractive than the 70% faces ($M = 4.67$, $SD = 1.78$) across the 14 face pairs, $F(1,15) = 58.34, p < .001$, $\eta^2_p = .80$. 


There was also a significant interaction between morphing level and face pair, $F(5, 76) = 2.65, p = .029, \eta_p^2 = .15$. To break down this interaction paired samples t-tests were conducted upon the 30% and 70% presentations of each face pair, these tests revealed that 12 of the 14 face pairs were rated as significantly more attractive when presented at 30% morph (see Figure 3.2).

![Figure 3.2](image)

**Figure 3.2** Results of paired samples t-tests conducted using the attractiveness ratings of the male and female 30%/70% face pairs.

### 3.3.3 Attractiveness ratings for presentations between 40% and 60% morph

The analysis of the presentations between 40% and 60% morph level revealed a significant main effect of morphing level, $F(2, 34) = 19.47, p < .001, \eta_p^2 = .57$. Contrasts revealed that
attractiveness ratings decreased as the level of morphing increased, and when compared to the 60% morphs, faces were considered more attractive at the level of 40% morph \(F(1,15) = 27.32, p < .001\), 45% morph \(F(1,15) = 47.42, p < .001\), 50% morph \(F(1,15) = 29.12, p < .001\), and 55% morph \(F(1,15) = 9.80, p = .007\).

**The effect of face pair**

There was a significant main effect of face pair \(F(6, 83) = 21.33, p < .001, \eta_p^2 = .59\). Attractiveness ratings were significantly different across the 14 face pairs, and it appears that when faces were presented for only 150ms, participants could still successfully differentiate between the different face pairs.

There was also a significant interaction between face pair and morphing level, \(F(52, 780) = 2.07, p < .001, \eta_p^2 = .12\), indicating that whilst attractiveness ratings decreased as morphing level increased, the extent of this decrease was not consistent across all 14 face pairs.

**The effect of cue**

In the memory task which was included at the end of the online task, all participants correctly identified at least 3 of 5 symbol/face pairings \((M = 3.81, SD = .83)\). However, during the lab-based task participants did not appear to use the cues as a predictor of attractiveness, and no difference in attractiveness ratings was observed when the same faces were paired with the attractively associated \((M = 5.12, SD = 1.75)\) and unattractively associated cues \((M = 5.08, SD = 1.76)\) \(F(1,15) = 4.16, p = .059\). Furthermore, no significant interactions were observed between cue and face pair \(F(13, 195) = .35, p = .982\), cue and morphing level \(F(4, 60) = 2.14, p = .087\), or between cue, face pair, and morphing level \(F(52, 780) = .76, p = .893\).

In order to further explore whether participants were using the cue information to guide their attractiveness judgements at all, an additional face pair x cue ANOVA was conducted using only the 50% presentations. This analysis was performed because theoretically, this is the point where participant uncertainty should have been at its highest.

At the 50% morph point, there was no difference in participants’ attractiveness ratings for faces paired with the attractively associated \((M = 5.17, SD = 1.77)\) and unattractively associated \((M = 5.10, SD = 1.77)\) cues \(F(1, 15) = 2.01, p = .177\). There was also no interaction between cue and face pair for the 50% morph presentations, \(F(4, 58) = .881, p = .478\).
Reaction time data

No differences in participant reaction times were observed between congruent stimulus cue pairings ($M = 946.56, SD = 412.27$) and incongruent stimulus cue pairings ($M = 961.95, SD = 486.27$), $F(1, 15) = 1.19, p = .287$. Reaction times were also not influenced by the face pair that was used, $F(5, 195) = 1.41, p = .148$, or by the level of morphing used in a particular presentation, $F(3, 45) = .48, p = .803$, and there were no significant interactions between face pair and morphing level, face pair and cue congruency, morphing level and cue congruency, or between face pair, cue congruency and morphing level (all $p$ values above .05).

3.4 Interim discussion

The results of experiment 1 reveal the following findings: (i) during the lab based task participants appeared to have difficulty differentiating between the 2 parent faces used within each composite morph. (ii) Whilst the results from the online portion of the study indicated that participants had learned the cue/stimulus associations, participants did not appear to use the predictive cues to guide their attractiveness ratings.

In experiment 1 a significant main effect of morphing level was observed, indicating that the composite morphs which were closer to parent face A (< 50% morph) were considered more attractive than the morphs which were closer to parent face B (>50% morph). This result was not unexpected, but average difference in attractiveness ratings observed between the 40% morphs and the 60% morphs (.39) was considerably smaller than the average difference in attractiveness ratings for unaltered attractive/unattractive face pairs (2.5) taken from the same participants earlier in the study. Additionally, whilst a significant main effect of morphing level was observed in the analysis of the 70% and 30% presentations, the average rating difference (.76) was again far smaller than the difference observed for the unaltered parent faces.

The 30% and 70% increments had been included throughout the task to anchor participant responses, and the difference between the 30%/70% pairs had been expected to more closely resemble the rating difference between unaltered attractive and unattractive parent faces. Since this was not the case, it is questionable how successfully they were able to anchor participant responses. The overall main effect of face pair indicates that participants were able to successfully differentiate the 14 face pairs from one another (ratings ranged from $M = 7.04, SD = .22$, to $M = 3.47, SD = .35$), yet the relatively small difference observed in attractiveness ratings between the 30% and 70% morphs indicates that whilst participants were able to successfully differentiate each face pair from one another, they were less successful at differentiating between the individual faces contained within each face pair.
Participants had not been expected to experience such difficulties in differentiating the faces within each pair from one another at morphing levels of 30% and 70%, due to the findings of previous studies which have explored face categorisation. However, a number of previous studies have used images of either bald males (e.g. Verosky & Todorov, 2010, 2013) or close-up images of faces, omitting face shape and hairlines (Jacques & Rossion, 2006). For the present study, in order to for participants to make valid judgements based upon attractiveness, the stimulus set was intended to be more representative of the faces that participants regularly encountered (especially since bald women are a minority within Western society). This meant that perceptually similar faces needed to be used in order to obscure the morphing effect (similar face shapes, similar hair styles etc.), it therefore seems possible that participants categorised faces based upon these features, causing both faces within a particular face pair (and their associated composite morphs) to be rated similarly.

An additional issue related to the current manipulation, is that participants did not appear to be using the predictive cues to aid their ratings. Again, this may be related to the difficulties participants experienced when attempting to differentiate between the parent faces used to create the composite morphs. Whilst participants rated morphs between 40% and 60% as slightly more attractive when they were paired with the cues associated with the 30% (more attractive) faces, this effect was not significant. An examination of the responses to 50% morphs stimulus uncertainty should have been greatest, also revealed no main effect of cue, indicating that participants were not using the cues at any stage to guide their attractiveness judgements.

The associations between the cues and the attractive/unattractive faces were never made explicitly, but participants were instructed to remember which faces were presented with which cues during the online stage of the task. Participants were also instructed at the start of the lab session that the same symbols would be used to help participants with their ratings. All participants performed above chance in the symbol memory task involved in the online task, but it appears that a more stringent manipulation check may have been required, since participants did not appear to be using the cue associations to guide their ratings. This assumption is also supported by the reaction time data which indicated that there was no difference in reaction times between congruent and incongruent cues. Taken together, these results indicate that participants never formed the cue/stimulus associations that were expected when the experiment was originally conceived.
Experiment 2

3.5 Amended design

After analysing the results from experiment 1, experiment 2 was conceived to address some of the issues with the design of the first experiment. The facial morphing techniques employed in experiment 1 had initially been appealing because they offered an intuitive way of transitioning between positively rated and negatively rated stimuli. However, due to the problems that participants experienced when attempting to differentiate parent faces from one another within a composite morph, even when they were presented at the 30%/70% morph level intended to anchor the stimulus/cue associations, the facial morphing manipulation was abandoned for experiment 2.

The removal of the morphing procedure in experiment 2 meant that attractive/unattractive faces did not need to be matched in terms of hairstyle or face shape. It also meant that, in comparison to experiment 1, the attractive/unattractive groups in experiment 2 could be comprised of more/less attractive faces. Because of the wider range of available faces, and because the attractive/unattractive faces no longer needed to be perceptually similar, it was hoped that participants would distinguish between the two groups more readily. In addition to the changes to the attractive/unattractive face groups, for experiment 2 the intermediate stimuli from experiment 1 (composite morphs existing between 2 parent faces) were replaced with a number of moderately attractive faces. Both the attractive/unattractive faces, and the moderately attractive faces used in experiment 2, were grouped together based upon the ratings from the earlier preliminary studies.

In experiment 1, participants had appeared to experience some difficulties in distinguishing between the faces contained within the composite morphs. Nevertheless, the significant main effect of face pair indicated that they had still acknowledged the difference between each of the 14 sets of face pairs. The design of experiment 2 meant that, compared to experiment 1, participants should have been able to differentiate between the stimuli more easily. In experiment 2, multiple stimuli did not share elements from the same parent faces but to compensate for the expected reduction in stimulus ambiguity, and to encourage the use of cues, the presentation durations for each face were reduced to 80ms.

Finally, the cues from experiment 1 (Vulcan symbols) were replaced with basic shapes of different colours to simplify the task, and to hopefully further enhance the cue/stimulus
associations. The rating procedure was also simplified, moving from a 1-9 attractiveness rating response scale in experiment 1, to a simpler liked/disliked two-alternative forced choice (2AFC) measure in experiment 2.

3.6 Methods

3.6.1 Participants
Fifty-eight undergraduate students aged between 18 and 26 (51 females, mean age = 19.24, SD = 1.76) were recruited for the study. All participants reported normal or corrected to normal vision, and had no reported history of neurological or psychiatric disorders. Participants were paid for their time and all participants gave written informed consent to participate. All testing procedures were approved by the University of Manchester ethics committee.

3.6.2 Stimuli and apparatus
Fifty-six black and white images of faces were used for experiment 2 (28 males and 28 females), images were drawn from the KDEF (Lundqvist et al., 1998), and FACES (Ebner et al., 2010) databases. The 56 images were divided into 3 groups: a highly unattractive group consisting of 5 male and 5 female faces (all previously rated < 3.5), a highly attractive group consisting of 5 male and 5 female faces (all previously rated > 6.5), and an moderately attractive group consisting of the remaining 18 male and 18 female faces (all previously rated between 3.5 and 6.5). The moderately attractive category was divided further into six sub-groups of ascending attractiveness; each sub-group contained 3 faces of each gender (see table 3.3). Like experiment 1, all of the faces used in experiment 2 were selected on the basis of the ratings collected during an earlier preliminary study (see Appendix A.1). Attractiveness was cued by associating the face presentations with either small blue squares, or small yellow circles which were superimposed over the four corners of each face image.

Stimuli were presented using a desktop computer with a 17-inch LCD monitor running E-Prime presentation software (Psychological Software Tools, Pittsburgh, PA) at a resolution of 1024 by 768 pixels. All images used in the task were cropped to a size of 562 x 762 pixels, and all responses were recorded using a standard UK layout keyboard.
Figure 3.3  Face Classification for stimuli used in experiment 2. Images show 2 examples from each attractiveness group. Error bars show standard errors of the mean. Note: Because participants in experiment 2 were asked whether they liked each face rather than how attractive they found each face, an additional preliminary study (3c) was conducted prior to experiment 2. The purpose of this study was to confirm that liking ratings were sufficiently correlated with attractiveness ratings (see Appendix A.3).

3.6.3 Design and procedure

Experiment 2 employed a within-participants, repeated measures design. Each of the faces within the highly attractive group was consistently paired with one type of cue (squares or circles), while each of the faces in the highly unattractive group was paired with the alternative cue (counterbalanced across participants). In total, the highly attractive faces, and highly unattractive faces were presented 20 times each (400 trials). Each of the moderately attractive faces was presented ten times with the square cue and ten times with the circle cue (720 trials).

Trials were assigned to blocks in a pseudo-random order with earlier blocks being intentionally weighted in favour of highly attractive/unattractive trials. This manipulation was implemented in early trials to emphasise the relationship between the cue and the attractiveness of a paired face. Each block contained 112 trials, and presentation order for the 10 experimental blocks was random within the following constraints: block 1 (55% highly attractive/unattractive faces, 45% moderately attractive faces), blocks 3 and 6 (25% highly attractive/unattractive faces, 75% moderately attractive faces), all remaining blocks (35% highly attractive faces, 65% moderately attractive faces).
Participants attended a single 45 minute session, and after being briefed by the experimenter, they were seated in front of the keyboard monitor setup. Before beginning the main task participants completed a set of practice trials, and the format of both the practice trials and the experimental trials was the same. Each trial began with a white cross-shaped fixation presented against a black background for 1,000ms. After the fixation cross disappeared, a face/cue pairing was presented for 80ms, followed by a rating prompt requesting a like/dislike rating. Participants responded using the ‘z’ (like) and ‘m’ (dislike) keys (counterbalanced across participants), and they were instructed to respond as quickly as possible with their initial ‘gut’ response. Each trial required a response, and the rating prompt remained on-screen until a response was provided (Figure 3.4).

Figure 3.4  Trial sequence from experiment 2. Cue and face pairings were divided into 3 different conditions: Highly attractive faces paired with 1 cue, highly unattractive faces paired with a different cue, and third category featuring moderately attractive faces. In the moderately attractive condition (shown above) a face could be presented with the cue from either of the other conditions.
3.6.4 Data analysis

For each of the moderately attractive faces, participants made 10 like/dislike judgements for both attractively associated and unattractively associated cue pairings. Four individual logistic psychometric functions (see general methods for further details) were fitted to the like/dislike ratio of the 720 moderately attractive trials from each participant using Matlab (The MathWorks, Natick, MA), and the Palamedes toolbox for Matlab (Prins & Kingdom, 2009).

Each function represents a participant’s responses to moderately attractive faces paired with a particular type of cue: male faces paired with the attractively associated cue, male faces paired with the unattractively associated cue, female faces paired with the attractively associated cue, and female faces paired with the unattractively associated cue (for an example see Figure 3.5). The data points required to plot the functions were based upon the number of ‘liked’ responses to faces within each of the moderately attractive subgroups (e.g. faces previously rated 3.5+, faces previously rated 4+ etc.).

From each function a point of subjective equality (PSE) was calculated. In this experiment the PSE illustrates the stimulus level (attractiveness) required for the participant to shift from providing one response more frequently (stimulus disliked) to providing the opposite response more frequently (stimulus liked).
Figure 3.5 Example psychometric functions illustrating responses from a single participant in experiment 2. This figure illustrates responses to moderately attractive faces with PSEs across all cue/face gender pairings.

PSE data were analysed using a 2 (cue type: attractive vs. unattractive) x 2 (gender: male vs. female) repeated-measures ANOVA. Liking ratings for the highly attractive/unattractive groups were compared using paired samples t-tests. In cases where Mauchly’s test revealed that the assumption of sphericity had been violated, the degrees of freedom have been corrected using Greenhouse-Geisser estimates of sphericity. An alpha level of .05 has been used for all statistical tests, and effect sizes are reported here as partial eta squared ($\eta_p^2$) where 0.05, 0.1, and 0.2 correspond to small, medium, and large effects, respectively (Cohen, 1988).
3.7 Results

Note: Due to the nature of experiment 2 (gauging the attractiveness of same, and different gender faces), and the low number of male participants in the main study, data from the male participants were not included in the analysis. The remaining sample consisted of 51 females aged between 18 and 26 (M = 19.04, SD = 1.35). A further 2 participants were removed at the analysis stage because adequate functions could not be fitted to their data. Therefore, the results reported below are based upon a sample of 49 female participants.

3.7.1 Liking ratings for each attractiveness level

Presentations of faces from the highly attractive group (M = 92%, SD = 10%) were liked significantly more than those from the unattractive group (M = 9%, SD = 12%), t(48) = 34.32, p < .001. A broadly linear increase was observed in the percentage of liked responses across the sub-groups of the moderately attractive face set when collapsed across gender (see table 3.2).

Table 3.2  Mean percentage of liked responses for the moderately attractive sub-groups

<table>
<thead>
<tr>
<th>Grouping</th>
<th>Attractive (SD)</th>
<th>Unattractive (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate 1</td>
<td>26% (23%)</td>
<td>22% (21%)</td>
</tr>
<tr>
<td>Moderate 2</td>
<td>37% (27%)</td>
<td>33% (27%)</td>
</tr>
<tr>
<td>Moderate 3</td>
<td>48% (28%)</td>
<td>45% (29%)</td>
</tr>
<tr>
<td>Moderate 4</td>
<td>72% (26%)</td>
<td>71% (27%)</td>
</tr>
<tr>
<td>Moderate 5</td>
<td>77% (25%)</td>
<td>74% (26%)</td>
</tr>
<tr>
<td>Moderate 6</td>
<td>85% (19%)</td>
<td>81% (23%)</td>
</tr>
</tbody>
</table>

Note: Results are based on 120 trials per group (60 trials per cue type) collapsed across gender

3.7.2 The effect of gender

After plotting functions from the responses provided by each participant, and calculating their individual PSEs, the PSE data did not reveal a main effect of face gender (male faces: M = 4.44, SD = 0.83; female faces: M = 4.56, SD = 0.99), F(1, 48) = 1.66, p = .204.
3.7.3 The effect of cue

The analysis did reveal a significant main effect of cue with PSE measures for the unattractively associated cue condition being significantly higher ($M = 4.57$, $SD = 0.98$) than those in the attractively associated cue condition ($M = 4.42$, $SD = 0.87$), $F(1, 48) = 12.86$, $p < .001$, $\eta_p^2 = .21$. This finding indicates that, when compared to the attractive cue condition, when the same moderately attractive faces were paired with the unattractively associated cues, regardless of gender, they needed to be more attractive before they were classified as liked (see Figure 3.6).

No significant interaction was observed between cue and gender, $F(1, 48) = .01$, $p = .997$, indicating that PSE measures for faces paired with attractively and unattractively associated cues across the 6 moderately attractive sub-groups were similar regardless of the gender of those faces.

![Figure 3.6](image)

*Figure 3.6* Functions for moderately attractive male and female faces presented in attractive and unattractive cue conditions. Functions are based upon the average number of liked responses for each moderately attractive sub-group collapsed across gender.
3.8 Discussion

The primary aim of the experiments discussed within this chapter, was to investigate whether expectations about the affective value of an upcoming stimulus produce a state of affective readiness, where participants’ affective responses to a stimulus fall in line with their expectations. The results of the 2nd experiment show a main effect of cue, illustrating that when the same moderately attractive faces were shown in highly attractive and highly unattractive cue conditions, participant PSEs (essentially participants’ liking thresholds) were significantly higher for the unattractive cue condition. This finding indicates that expectations brought about by the presence of the unattractively associated cue, meant that faces paired with that cue needed to be more attractive before they were considered liked. This finding also supports the ideas proposed by the AEM (T. D. Wilson et al., 1989) whereby a negative or positive expectation about a stimulus may be sufficient to cause participants to assimilate the affective value of the stimulus toward their expectations.

The influence of expectations over our affective judgements has already been explored in a number of previous studies, particularly those investigating preferences for various types of consumables. However, as Kuenzel et al. (2011) note, the differences between the stimuli used in many of those studies are most often purely categorical in nature (e.g. a product is from one region or another, a product either contains a particular ingredient or it does not). Yet in more naturalistic settings, stimuli are rarely exactly defined in terms of binary categories like these. For example, in limited quantities an ingredient like salt may be considered pleasant and thus liked, but if salt concentration increases beyond a certain threshold, it may instead produce the opposite response. It is for this reason that several recent studies have turned to psychophysical methods in order to assess the ideal concentrations of such ingredients.

The present study differs from much of the existing research regarding the influence of expectations over affective judgements, because the application of psychophysical methods illustrates more directly the level of bias or ‘affective readiness’ caused purely by the cueing manipulation. Studies which have employed similar psychophysical methods have traditionally plotted participant responses against the properties of a stimulus which can be defined in terms of an objective measure e.g. concentrations of salt (Herbert et al., 2014) or the frequency of a tone (Piazza, Sweeny, Wessel, Silver, & Whitney, 2013). In the case of experiment 2, the measure (attractiveness) was far more subjective; however, the broadly linear increase in liking responses observed across the moderately attractive sub-groups, coupled with the fact that it
was possible to fit an adequate function to participant responses in the majority of cases, provides justification for the employment of this method of data analysis.

Summerfield et al. (2006) argue that people are more likely to defer to a strong expectation when a stimulus presentation is impoverished somehow, since this creates greater stimulus uncertainty. For example, Bar's (2003) model of object recognition suggests that when an object is detected by the human visual system, low spatial frequency (LSF) information is quickly extracted offering the observer an idea of the global form of the object. This information allows the observer to form some initial guesses about the object’s identity, and over time these guesses are refined when they are integrated with high spatial frequencies providing information about the finer details of the object. However, when stimuli are presented very rapidly, observers only have access to information about the global form of an object. As a result, predictive information about the identity of an object becomes more important allowing for facilitated object recognition but with the potential cost of increased rates of error (Bar, 2004; Król & El-Deredy, 2011a).

As Jacques & Rossion (2006) indicate, the time required for the human visual system to discriminate between individual faces ranges between approximately 80ms and 150ms after stimulus onset. Consequently, experiment 2 was designed based upon the belief that by presenting faces for only 80ms per trial, participants would be forced to rely upon cue based expectations rather than purely upon the stimuli themselves, and this appears to have been the case. When participants were shown moderately attractive faces presented with attractively or unattractively associated cues, their affective responses to those faces shifted and were assimilated toward the affective values suggested by the cues.

The initial aims of experiment 1 were to investigate both assimilation, and contrast effects within the same paradigm, however these aims needed to be revised during the design of experiment 2. Theoretically, contrast effects should occur when participants acknowledge the difference between that which is expected, and that which is actually received. However, the faces that could potentially produce contrast effects (those in the highly attractive and highly unattractive categories) needed to be consistently paired with congruent cues in order to anchor the cue stimulus associations (Strack, Bahník, & Mussweiler, 2016). Had stimuli been introduced which deviated considerably from the affective values expected by the participants, then the study may not have produced the assimilation effects that were observed. If participants had experienced a number of trials where the cues and stimuli were highly incongruent, then it is likely that cue/stimulus associations would weaken over time due to due
to memory reconsolidation, and participant distrust (J. L. C. Lee, 2009; Mayo, 2015). Therefore, the investigation of contrast effects was not feasible using the paradigm employed in experiment 2, but it is something that will be revisited in later chapters.

One potential limitation of the paradigm employed in experiment 2, is that the results are based upon an all-female sample, and no data was collected relating to sexual orientation. This is potentially problematic because, in terms of their affective responses to images of faces as rewarding stimuli, males and females are understandably likely to differ. For example, imaging studies have observed greater activations in the reward circuitry of heterosexual males in response to images of attractive females (Aharon et al., 2001), and behavioural studies have found that males will expend more effort and forego greater potential financial rewards for the opportunity to observe attractive opposite-sex faces (Hayden, Parikh, Deaner, & Platt, 2007; M. Wilson & Daly, 2004). For this reason later chapters have attempted to introduce stimuli which are not so susceptible to gender biases. Nevertheless, whilst the current sample of participants liked the moderately attractive female faces marginally more overall than the male faces, a main effect of gender was not observed, and there was no interaction between gender and cue. This finding is important because, since the sample produced similar assimilation effects for both male and female faces, it suggests that the effect of the cues remained consistent regardless of whether the observer had a greater overall preference for observing male or female faces.

3.9 Conclusion

The current results, obtained using a relatively simple affective judgement task based upon facial attractiveness, support the idea that affective expectations produce a state of affective readiness. When a positive, or a negative state of affective readiness was induced using cues, participants rated the same set of moderately attractive face stimuli differently depending on the nature of the cue that they were presented with. This finding is in line with the assimilation effects suggested by the AEM (Wilson et al., 1989) where, if a stimulus adheres fairly closely to an expected affective value, affective responses to that stimulus are assimilated toward the expectation. Whilst previous studies of affect have observed similar increases or decreases in liking using cues that predict the affective value of the stimuli at hand, there are few, if any, that have used psychophysical measures to provide more direct evidence that the expectation of liking/disliking actually influences the threshold by which an individual judges particular stimuli.
4 AFFECTIVE ASSIMILATION AND CONTRAST EFFECTS THROUGH THE OBSERVATION OF HAPPY AND SAD EMOTIONAL FACES

4.1 Introduction

The results reported in chapter 3 fell broadly into line with what had been expected, based upon the predictions of the AEM (T. D. Wilson et al., 1989), namely that affective expectations would produce a state of affective readiness, priming participants to accept items of a certain valence. When moderately attractive faces were paired with positive and negative predictive cues, affective reactions to those faces assimilated toward the values suggested by the cues. In cases like this, where the valence of an item is similar to an expected value, it is more efficient (in terms of processing speed and the allocation of cognitive resources) to base an affective reaction upon the expected value, rather than the available stimulus information (Geers & Lassiter, 2005; Kirsch, 1999). However, the AEM also predicts that as the gap widens between the value of a stimulus and the affective expectation, a point is eventually reached where the discrepancy between the two is consciously acknowledged. After this point individuals are instead more likely to evaluate a stimulus in the opposite direction to their affective expectations.

Compared to assimilation effects, contrast effects are less widely reported within the affect related literature. Indeed Wilson et al. (1989) found no evidence of contrast effects in the studies related to their original proposal of the AEM, and elsewhere other studies have found only limited evidence of contrast effects (e.g. Cardello & Sawyer, 1992). According to Herr, Sherman, and Fazio (1983) contrast effects emerge when the target stimulus is extremely far removed from the prime or anchor stimulus. One reason that has been suggested for the relative scarcity of contrast effects is that many studies do not meet the necessary requirements to produce contrast effects; more specifically, the paradigms used in these studies may not have been sufficient to create the required size of discrepancy between expectation and experience necessary to produce contrast effects. Another possible reason for the lack of affective contrast related evidence is that contrast effects may simply be quite rare within the affective domain (Klaaren, Hodges, & Wilson, 1994).

Contrast effects are, however, a key part of the AEM, and without evidence for contrast effects, it could be argued that assimilation effects can be explained by other theories. For example, cognitive dissonance (Festinger, 1957) would suggest that assimilation occurs when individuals
alter their interpretation of an experience to avoid the mental discomfort associated with admitting that their expectations have been disconfirmed. At the present time, the only published experiments which have provided evidence to support both the assimilation and contrast predictions of the AEM have been conducted by Geers and Lassiter (1999, 2002, 2003, 2005), and most recently Davidenko et al. (2015). In these experiments, the authors discuss several factors which could influence whether affective reactions are assimilated toward expectations or contrasted against them. Those factors include: processing style, an individual’s level of optimism/pessimism, and stimulus familiarity.

In Geers and Lassiter (1999) participants were shown a ‘not-so-funny’ film clip, and asked to rate how much they liked the clip. Before playing the clip, half of the participants were given positive expectations via written instructions and half were not. During the clip, the level of perceptual analysis was also manipulated between participants. Some participants were instructed to segment the clip into a few meaningful actions; others were instructed to segment the clip into many meaningful actions. Affective reactions for participants who had been instructed to segment the clip into only a few actions were assimilated toward the positive expectation, but affective reactions from participants who segmented the clip into many actions, instead contrasted against the positive expectation. The authors argue that participants who divided the clip into many actions processed the clip more thoroughly; consequently the discrepancy between the film clip and the prior expectation became more apparent for participants in this condition promoting contrast effects. Elsewhere, similar findings have been reported by Gendolla, Brinkmann, & Scheder (2008) who reported changes in assimilation effects for an affective evaluation task where task importance was manipulated. When a picture rating task was classified as unimportant, affective evaluations assimilated toward expectations. When the same task was instead classified as important, assimilation effects were reduced as participants assumedly attended to the task more and the discrepancy between the expectation and the actual experience became more apparent.

In a pair of later studies, Geers and Lassiter (2002) investigated the moderating role of optimism/pessimism in determining affective experience. Study 1 involved the presentation of short film clips which were viewed under either positive expectation or no expectation conditions; study 2 was similar in design but it included negative expectation, and no expectation conditions. The results of the studies (based upon optimism/pessimism measures collected during the course of each study) revealed that optimists were more likely to assimilate their affective reactions toward an expectation, while pessimists were more likely to contrast
the affective reactions against the expectation. Moreover, this finding was repeated for both positive and negative affective expectations. In other related research, optimists have been shown to acquiesce to cues more readily than pessimists (Morton et al., 2011), optimists have displayed a greater tendency to report placebo effects (Morton, Watson, El-Deredy, & Jones, 2009), and optimists have been shown to acknowledge fewer discrepancies than pessimists in proofreading tasks (Spirrison & Gordy, 1993). These findings also tie in to the previous discussion about processing style. Higher levels of optimism should produce more positive moods, and people in happier moods are thought to engage in more global, over local processing, in other words, the likelihood of assimilation is increased because people who are more optimistic, display a greater tendency to focus on the forest rather than the trees (Gasper & Clore, 2002; Hunsinger, Isbell, & Clore, 2012).

It is also possible that stimulus familiarity acts as a similar moderating factor in shaping affective reactions. In the majority of studies examining the influence of affective expectations, expectation is used to signal that an ambiguous target item belongs in a particular category, e.g. When an experimenter says, “I’m going to show you a film clip that others have found funny.” The implication is that the target clip belongs in the funny category. However, if the participant is already familiar with the target item then it is no longer ambiguous; in the mind of the participant it is already a member of a particular category. As such, the participant is more likely to contrast against the experimentally induced expectation, because they are aware that the target item is not a sufficient match for the expected category (Herr et al., 1983). In Geers and Lassiter (2005) the influence of stimulus familiarity was assessed using film clips which were shown across multiple sessions. In the first session participants saw either the target ‘unfunny’ film clip, or an unrelated clip; in the second session all participants saw the target clip. When a sub-section of participants were given a positive expectation at the start of session 2, affective responses from participants who had not seen the target clip beforehand, assimilated toward the expectation, while responses from participants who had already seen the clip were adjusted away from the expectation. Again, these findings can be related back to issues about how an item is affectively processed. Studies have shown that experts make more associative predictions and perform better in visual perception tasks, compared to novices, (Cheung & Bar, 2012; Rota & Zellner, 2007) so familiarity (and expertise) may be influential in determining whether people adopt a global or local processing style when making affective judgements.

Contextual factors could also play a role in determining whether affective responses are assimilated toward expectations, or contrasted against them. Affective judgements are never
made in complete isolation; previous experiences influence affective judgements about new ones and vice-versa. For example, Tversky and Griffin (2000) suggest that enjoying an exceptional meal at a certain restaurant, will negatively impact judgements of meals which are merely very good at similar establishments. Similarly, when there is a strong positive expectation about a stimulus, affective judgements may be negatively influenced if the stimulus is not considered to be as positive as expected. Alternatively, in other scenarios a similar positive expectation may instead cause the stimulus in question to be rated more positively than it would have been otherwise. According to Stapel and Winkelman (1998), assimilation will occur if predictive information is used as an interpretation frame, while contrast will occur if predictive information is used as a comparison standard. The likelihood of either of these options is determined by both the ambiguity of the target stimulus (if the stimulus is unambiguous there is no need for interpretation), and also the extremity of the predictive information (extreme context information is more likely to produce contrast effects than moderate information).

One final factor which might influence the extent of assimilation and contrast effects is whether expectations are positive or negative in nature. According to Ditto and Lopez (1992) people require less information to reach preference-consistent, rather than preference inconsistent conclusions. This may simply mean that participants require less information to reach a conclusion that is consistent with an existing expectation, and both Wilson et al. (1989) and Geers and Lassiter (2002) found that both negative and positive expectations operated similarly. When a target stimulus was similar in nature to an affective expectation, affective responses assimilated toward the expectation, regardless of whether that expectation was positive or negative. However, a commonality between these studies is that they both used stimuli which were purely categorical in nature i.e. comics or film clips that that were deemed funny or unfunny. Little is currently known about the size of the discrepancies required to cause assimilation and contrast effects, and it is possible that the nature of the stimulus (in terms of being positive or negative) may be a more influential factor if a broader range of discrepancy sizes were examined.

The aims of the current study were (i) to attempt to produce both assimilation and contrast effects within a single affective expectation paradigm, and (ii) to investigate the influence of the factors discussed above: optimism/Pessimism, stimulus context, and the nature of the affective expectation in terms of it being positive or negative. A further aim (iii) was to investigate
whether the size of the discrepancy between the expectation and the target stimulus interacted with any of these factors.

The experiments discussed in chapter 3 were designed to show that differing expectations cause participants to utilise different affective thresholds when appraising target stimuli. As such their design did not lend itself to the investigation of both assimilation and contrast effects. To achieve the aims of the current study, a paradigm was adapted from an experiment conducted by Morton et al. (2011). Participants were shown a series of images of happy and sad emotional faces of differing emotional intensities, each face was preceded by a descriptor cue, and cues were manipulated, in terms of how accurately they predicted the presented target emotions. In an effort to promote contrast effects, the current study also employed a wider range of discrepancies between the predictors and target stimuli, compared to the original study conducted by Morton et al. (2011).

Measures of optimism/pessimism were collected at the start of the study, and since happier faces are more often associated with positive judgements (e.g. Oosterhof & Todorov, 2008), the positive/negative element of the investigation was related to whether the descriptor cues predicted a happier or sadder emotion than the target stimulus. Stimulus familiarity (or expertise) were not manipulated directly, but since humans are considered experts in terms of facial recognition (Tanaka, 2001), and since there is evidence that the six basic categories of emotions (happiness, sadness, surprise, disgust, anger and fear) are universally recognised (Ekman, 1994; Ekman & Oster, 1979), all of the participants should have been familiar with the types of stimuli used as targets (the different levels of emotion), even if they had not seen the specific exemplars that were used beforehand.

Context was manipulated in terms of the relationship between the emotion predicted by the cue and the emotion expressed by the target face. Cues could either predict an emotion within the same category as the target face (happy and sad congruent emotion conditions - e.g. happy predictor and target face showing happy emotion), or the cues could predict an emotion from a different category (cross emotion condition - e.g. sad predictor and target face showing happy emotion). Several studies have shown that humans find it easier to discriminate between expressions from 2 different emotional categories, rather than 2 expressions from the same category, (Calder, Young, Keane, & Dean, 2000; Etcoff & Magee, 1992; Leppänen, Richmond, Vogel-Farley, Moulson, & Nelson, 2009), therefore it was assumed that the potential for contrast effects should be increased within the cross emotion condition. During data collection an additional factor was included in the form of presentation time. It was hypothesised that, by
presenting the stimuli for shorter durations (thus increasing ambiguity), participants would assimilate further toward the values suggested by the cues. Since this additional manipulation was not added until after data collection had already begun, only half of the participants completed the version of the task including the original (100ms) and shorter (50ms) presentation times.

Experiment 3

4.2 Methods

4.2.1 Participants
Ninety-eight participants aged between 18 and 45 (74 females, mean age = 24.24, $SD = 6.23$) were recruited for the study. Participants were required to attend a single session which lasted one hour, and multiple participants took part in the study concurrently. The study was conducted across 10 sessions involving 3 to 24 participants (mean participants per session = 9.8, $SD = 7.07$).

All participants reported normal or corrected to normal vision, and had no reported history of neurological or psychiatric disorders. Participants were paid £5 for their time and all gave written informed consent to participate. All testing procedures were approved by the University of Manchester ethics committee.

4.2.2 Stimuli and apparatus
Twenty-four grey-scaled images of emotional faces were taken from the Ekman faces (Ekman & Oster, 1979), categories PE and SW. The faces drawn from PE (male face) and SW (female face) were 6 images of happy and 6 images of sad expressions. These images corresponded to Ekman emotional intensities of 3, 6, 9, 12, 15 and 18, with 3 representing the least intense, and 18 representing the most intense versions of each emotion. To aid participant learning the emotional intensities were remapped as 1, 2, 3, 4, 5 and 6 (e.g. a face with a happy expression and an Ekman intensity of 3 became a ‘Happy 1’).

Images were presented on a projector screen at the front of a lecture theatre, and the order of presentations was controlled using a computer running PsychoPy presentation software (Peirce, 2007). Audio was also controlled by this software, and any accompanying audio was routed through the speakers in the lecture theatre. At the start of each session participants were provided with a pre-printed answer sheet (Appendix B.1) to record their responses, and a copy
of the Life Orientation Test Revised (LOT-R) (Scheier et al., 1994) to measure generalised optimism.

4.2.3 Design and procedure

At the start of the session participants were briefed about the experiment, and given a copy of the response sheet and LOT-R to complete. Participants were told that they would be viewing a series of emotional faces, and that they would be trained to rate the intensity of the emotions displayed by the faces. Before beginning the task proper, participants were familiarised with the faces and emotional intensities (e.g. happy 1, sad 1 etc.). The entire range of emotional levels for each face was first presented on a single slide, before each level of emotion was shown in sequence with emotional descriptors appearing beforehand (Figure 4.1).

Figure 4.1 Examples of faces shown during the training procedure. In (a) participants were first shown all 6 emotional intensities on a single slide for all emotions (happy/sad), and all faces (PE/SW). In (b) participants were shown each intensity level for each emotion individually with PE and SW presented side by side. A description of the emotional intensity was first shown for 2000ms, followed by the images displaying that emotion for 2000ms.

Participants then completed 24 practice trials which involved the presentation of a single face per trial, each level of emotional intensity was presented once per face type (PE and SW), and
trials were presented in a random order without emotional descriptors. Presentations lasted 2000ms, and participants were given 4000ms to note down the emotion, and the level of intensity that they believed was correct. Tones were used to indicate the beginning and end of each trial. The output file produced by the presentation software indicated that presentation times were consistent across participants. It is, however, possible that the projector may have influenced the accuracy of presentation times, but any artefacts introduced by the projector should have been consistent across all participants.

After completing the practice trials participants were given a short break before beginning the experimental trials. During the break they were briefed on the procedure for the main task; this brief informed them that they would see presentations of the same faces used in the practice trials, but each presentation would now include an emotional descriptor. Participants were led to believe that descriptors were based upon the average response to a given presentation provided by a previous group; in reality the descriptor presented in each trial was actually dependent upon a pre-determined presentation condition (see Presentation conditions). For the first 46 participants the task consisted of 116 trials (92 experimental trials and 24 practice trials), for the remaining 52 participants the experiment consisted of 208 trials (184 experimental trials and 24 practice trials). The reason for the difference in the number of trials is explained in the presentation conditions section. Each trial consisted of an emotional descriptor (cue) presented for 2000ms which was immediately followed by a face presentation (see Figure 4.2). Face presentations lasted for 100ms or 50ms (depending on presentation conditions), and participants were given 4000ms to record their response. The same tones from the practice trials were used to indicate when participants needed to attend to the screen, and when they could look down at the response sheet. Sessions lasted between 30 and 45 minutes, and a break was provided at the midway point of each session.
Figure 4.2  Sequence of events from experiment 3. Each trial began with a fixation cross presented for 1000ms accompanied by a tone lasting 50ms. A descriptor cue was then presented for 2000ms which either accurately or inaccurately predicted the upcoming stimulus. The descriptor cue was followed by an emotional face presented for a variable duration (either 50ms, or 100ms). The rating prompt was accompanied by a second tone and signalled that participants should provide a response. Participants were given 4000ms to provide a response in each trial. Trial numbers were presented in the top left of the screen throughout the task.

4.2.4 Presentation conditions

Presentation conditions were determined by the following factors:

1. Presentation duration – In each trial a face was shown for either 100ms or 50ms. The 100ms presentation duration was a reduction on the 500ms presentation times reported in the original study conducted by Morton et al. (2011). Presentation time was reduced to 100ms for the present study to increase stimulus uncertainty, and whilst the accuracy of emotion categorisation was expected to be lower for 100ms presentations, previous research has shown that categorisation (for sad happy and fearful emotions) should still remain similarly accurate (Calvo & Lundqvist, 2008). Presentation time was not originally going to be varied, however, the 50ms presentation time was included
half way through participant recruitment as a means of increasing stimulus uncertainty further to investigate whether presentation duration influenced participants’ use of the descriptor cues. The inclusion of this additional condition mid-way through testing is the reason for the discrepancy in the number of trials between the first 46 participants, and the remaining 52 participants.

2. Discrepancy Size – The size of the discrepancy between the emotion suggested by the descriptor cue and the emotion presented by the face could be one of 3 levels: small discrepancy (2 step difference between the descriptor and the face), medium discrepancy (3 step difference) or large discrepancy (4 step difference).

3. Discrepancy direction – If the descriptor cue did not accurately predict the upcoming emotion, it could predict an expression that was either more positive, or more negative than the emotion present in the face presentation (Sad 5 is considered the most negative expression, and Happy 5 the most positive expression in the experimental trials).

4. Emotion – Happy, sad or cross emotion. In happy trials, the descriptor cue predicted a happy expression, and a face with a happy expression was observed. In sad trials, the descriptor cue predicted a sad expression, and a face displaying a sad expression was observed. Cross emotion trials included a descriptor cue predicting an emotion of one type (happy/sad), and a face expressing an emotion of a different type (happy/sad).

During training participants were familiarised with levels 1 – 6 the happy and sad emotions for both face PE and face SW. In the experimental trials however, only levels 1 – 5 of each emotion were actually included (see Figure 4.3).

Figure 4.3  Emotional intensities used in each presentation condition
4.2.5 Trial breakdown
Twenty four trials were included for each of the emotion conditions (Happy, sad, cross emotion). The 24 trials within each emotion condition included: 4 repeats at each level of discrepancy direction (positive and negative), size (small, medium and large), and face (PE and SW). In an effort to mask the cue manipulation, 20 congruent trials were also included where the emotional descriptor was an accurate predictor of emotional intensity. The congruent trials consisted of a single repeat of each of the 5 intensities for both the happy and sad expressions, repeated for both face PE and face SW. This resulted in a total of 92 experimental trials (all emotional conditions + congruent trials) for the first 46 participants. The final 52 participants completed double the amount of experimental trials (184 trials) due to the inclusion of the additional presentation duration (50ms and 100ms) for each of the conditions.

4.2.6 Data analysis
Measures of bias and accuracy for each participant were calculated based upon their responses during the practice trials. Bias measured how positively or negatively participants rated emotions when they were presented without predictive cues, and represented the average intensity difference between presented emotions and participant responses across the practice trials. Since emotional intensities were classified on a scale ranging from -6 (sad 6) to +6 (happy 6), bias could be represented by either a positive, or a negative number. Accuracy data was collected to ensure that all participants were similarly adept at performing the rating task before the cueing manipulation was introduced. To ensure that participant accuracy was represented by a positive number, accuracy was calculated by squaring the difference between each presentation and response, and then taking the square root of the squared difference.

Results are based upon the difference between the emotion intensity actually presented during each experimental trial, and the emotional intensity reported by the participant. For all participants, the average difference at each level of the presentation conditions was calculated first, and then the results were subjected to a 3 (emotion type: happy, sad, cross emotion) x 3 (size: small, medium, large) x 2 (direction: positive, negative) repeated measures analysis of covariance (ANCOVA). Bias and LOT-R scores were included as covariates to examine whether participants’ emotion ratings were related to their initial levels of bias, or how optimistic they were. This analysis was based upon responses to only the 100ms presentations from all participants who were considered sufficiently accurate during the practice trials.

The influence of presentation duration was assessed separately with a second ANCOVA which only included data from participants who had completed the version of the task featuring the
50ms and 100ms presentations. This second ANCOVA included the factors: emotion (happy, sad, cross emotion) size (small, medium, large) direction (positive, negative) and time (50ms, 100ms). LOT-R and bias were again used as covariates during the analysis.

Two-way interactions were analysed using paired samples t-tests between presentation conditions. Three-way interactions were first analysed using sub-ANCOVAs to determine which variable(s) (i.e. emotion, discrepancy size, or discrepancy direction) produced the significant interaction, and then subjected to paired samples t-tests between the possible presentation types. Bonferroni corrected $p$ values are reported where appropriate.

In cases where Mauchly’s test revealed that the assumption of sphericity had been violated, the degrees of freedom have been corrected using Greenhouse-Geisser estimates of sphericity. An alpha level of .05 has been used for all statistical tests, and effect sizes are reported here as partial eta squared ($\eta^2_p$) where 0.05, 0.1, and 0.2 correspond to small, medium, and large effects, respectively (Cohen, 1988).

4.3 Results

4.3.1 All participants at 100ms

Measures of accuracy and bias

On average, participants’ responses during the practice trials were 1.11 ($SD = 0.33$) steps away from the emotional expression that was actually presented. Participants who were not deemed sufficiently accurate during the practice trials (those whose responses were more than 1 standard deviation above the average accuracy response) were not included in the main analysis. This resulted in the removal of 15 participants (3 males) leaving a remaining sample of 83 participants.

Bias scores for the remaining sample revealed that during the practice trials, on average, participants rated the displayed emotions slightly more positively than they actually were ($M = 0.43$ $SD = 0.34$). However, the results of the ANCOVA revealed that the relationship between participants’ bias in the practice trials, and their accuracy in the experimental trials, was not significant, $F(1, 80) = 3.17, \ p = .079$.

Emotion and discrepancy direction

The analysis revealed a significant main effect of emotion, $F(1, 112) = 13.78, \ p < .001, \ \eta^2_p = .15$, and a significant main effect of discrepancy direction $F(1, 80) = 50.43, \ p < .001, \ \eta^2_p = .39$. A
significant discrepancy direction x emotion interaction was also observed, $F(2, 160) = 44.87, p < .001, \eta^2_p = .36$ (see Figure 4.4). Paired samples $t$-tests revealed that participants reported more positive emotions when the descriptor cues were positive in nature for the sad emotion condition, $t(82) = 16.29, p < .001$, and the cross emotion condition, $t(82) = 11.54, p < .001$, but not the happy emotion condition, $t(82) = -1.96, p = .053$.

![Figure 4.4](attachment:image.png)

**Figure 4.4** The influence of discrepancy direction and emotion condition on the accuracy of participant responses for 100ms presentations. Error bars represent standard errors of the mean.

**Optimism and discrepancy size**

Optimism scores, as measured by the LOT-R ranged between 3 and 30 ($M = 18.51$ $SD = 5.73$), and were not significantly related to the accuracy of ratings provided during the experimental trials, $F(1, 80) = 2.31 p = .13$. However, a significant optimism x discrepancy size x discrepancy direction interaction was observed, $F(2, 160) = 3.13, p = 0.46, \eta^2_p = .04$ (see Figure 4.5). In the case of small discrepancies between the presented emotion and the cue, a significant positive correlation was observed between optimism scores, and ratings for positively cued trials ($r = .243 p = .027$ (two tailed)) but not negative cued trials ($r = -.017 p = .882$ (two tailed)). For medium sized discrepancies the opposite relationship was observed; optimism scores were positively correlated with ratings for negatively cued trials ($r = .251 p = .022$ (two tailed)), but
not positively cued trials ($r = .074 \ p = .0504$ (two tailed)). Finally, for the large discrepancies, no significant correlations were observed between optimism scores and ratings for either positively cued trials ($r = .034 \ p = .761$ (two tailed)) or negatively cued trials ($r = .011 \ p = .921$ (two tailed)).

\[ a. \]
c.

Figure 4.5   Optimism x discrepancy size x discrepancy direction interaction for 100ms presentations: a, shows that higher optimism scores were related to more positive ratings for positively cued trials over small cue/target discrepancies; b, shows that higher optimism scores were related to more positive ratings for negatively cued trials over medium discrepancies; c, shows no relationship between emotion ratings and optimism scores in trials featuring large discrepancies.

Elsewhere, a main effect of discrepancy size was observed, \( F(2, 134) = 4.31 \) \( p = .021 \), \( \eta_p^2 = .05 \). A significant discrepancy size x emotion interaction was also observed, \( F(4, 320) = 16.43 \) \( p < .001 \), \( \eta_p^2 = .17 \), and a significant discrepancy size x discrepancy direction x emotion interaction, \( F(4, 320) = 7.38 \) \( p < .001 \) \( \eta_p^2 = .08 \) (see Figure 4.6). In order to unravel this interaction, separate sub-ANCOVAs were conducted for the positive and negative discrepancy conditions. A significant emotion x discrepancy size interaction was observed for both the positive discrepancies, \( F(4, 289) = 12.37 \) \( p < .001 \), \( \eta_p^2 = .13 \), and negative the discrepancies, \( F(4, 320) = 11.67 \) \( p < .001 \), \( \eta_p^2 = .13 \).

Paired samples t-tests conducted for the positive cue discrepancy condition revealed significant reductions in ratings for happy faces as discrepancy size increased from small, to medium, to large (all \( p < .001 \)). A significant increase in ratings was seen for sad faces as discrepancy size increased from small to medium, \( t(82) = 3.43 \), \( p < .001 \); but ratings did not increase significantly between medium and large discrepancy sizes \( t(82) = .633 \), \( p > 0.025 \) Bonferroni corrected). No
significant differences were observed between any of the discrepancy sizes for the cross emotion faces (all $p > .05$).

Further paired t-tests conducted on responses from the negative cue discrepancy condition revealed a significant decrease in ratings for happy faces between small and medium discrepancies only, $t(82) = -4.63, p < .001$. No difference in ratings for sad faces was seen between small, medium and large discrepancies. However, a significant increase in ratings for the cross emotional faces was observed as discrepancy size increased from small to medium, then from medium to large all $p < .001$. 

a.
4.3.2 Data from participants who completed 50ms and 100ms trials

Of the 83 participants included in the analysis described above, 38 participants (4 males) had completed the version of the task which included the 50ms and 100ms presentations. Data from these participants were included in the second ANCOVA which included presentation time as an additional within-subjects factor.

Measures of accuracy and bias

In terms of accuracy, responses from this sample were 1.04 ($SD = 0.28$) steps away from the emotion which was actually presented during the practice trials. Bias scores indicated that again, on average, participants rated the displayed emotions slightly more positively than they actually were ($M = 0.42$, $SD = 0.31$). Bias was also significantly related to the accuracy of ratings collected during the experimental trials $F(1, 35) = 4.98$, $p = .032$, $\eta^2_p = .13$.

Emotion condition and discrepancy direction

No main effect of emotion condition was observed, $F(1, 46) = 3.36$, $p = .062$, but there was a main effect of discrepancy direction, $F(1, 35) = 6.13$, $p = .018$, $\eta^2_p = .15$. A significant emotion condition x discrepancy direction interaction was still observed for this sample $F(2, 70) = 32.20$,
p = .001, $\eta^2 = .48$ (see Figure 4.7). Paired samples t-tests revealed that participants reported more positive emotions when the descriptor cues were positive in nature for the sad emotion condition, $t(37) = 14.64, p < .001$, and the cross emotion condition, $t(37) = 7.68, p < .001$. However, in the happy emotion condition, participants reported more negative emotions when the cues predicted more positive emotions, $t(37) = -3.39, p = .006$.

![Figure 4.7](image)

**Figure 4.7** *The influence of discrepancy direction and emotion condition on the accuracy of participant responses at 50ms and 100ms. Error bars represent standard errors of the mean.*

**Optimism and discrepancy size**

Optimism, as measured by the LOT-R ($M = 21.74, SD = 3.69$) was not significantly related to how participants rated the faces, $F(1, 35) = 1.57, p = .218$. The direction x size x optimism interaction seen in the previous analysis, was also not present for this sample $F(2, 70) = .641, p = .530$.

There was no main effect of discrepancy size, $F(2, 53) = .70, p = .463$, but there was a significant emotion condition x discrepancy size interaction, $F(4, 140) = 17.22, p < .001, \eta^2 = .33$, and a significant discrepancy direction x discrepancy size interaction, $F(2, 60) = 4.19, p = .025, \eta^2 = .11$. Furthermore, the emotion condition x discrepancy direction x discrepancy size interaction seen in the previous analysis was also observed here, $F(4, 170) = 10.82, p < .001, \eta^2 = .24$ (see Figure 4.8).
Sub-ANCOVAs performed on the positive and negative cue discrepancy conditions separately, again revealed a significant emotion x discrepancy size interaction for both the positive discrepancies, \( F(3, 113) = 7.81, p < .001, \eta^2_p = .18 \), and the negative the discrepancies, \( F(4, 140) = 21.60, p < .001, \eta^2_p = .38 \).

Paired samples t-tests conducted for the positive cue discrepancy condition revealed significant reductions in ratings for happy faces as discrepancy size increased from small, to medium, to large (all \( p < .001 \)). Ratings for sad faces did not increase significantly between small and medium discrepancy sizes, \( t(37) = 2.12, p = .274 \), but did increase significantly between medium and large discrepancy sizes, \( t(82) = 5.18, p < .001 \). Like the previous analysis, no significant differences were observed between any of the discrepancy sizes for the cross emotion faces (all \( p > .05 \)).

Further paired t-tests conducted on responses from the negative cue discrepancy condition revealed a significant decrease in ratings for happy faces between medium and large discrepancies only, \( t(37) = -8.45, p < .001 \). A significant decrease in ratings for sad faces was seen between small the large and small discrepancies only \( t(37) = -3.27, p < .001 \). Responses for the cross emotion condition followed a similar pattern to the previous analysis with significant increases in ratings occurring between small and medium, and medium and large cue discrepancy sizes (all \( p < .001 \)).

a.
Figure 4.8  The interaction between discrepancy size, discrepancy direction, and emotion condition for 100ms and 50ms presentations; a. shows responses when the descriptor cue predicted a more positive emotion that the one that was shown; b. shows responses when the descriptor cue predicted a more negative emotion than the one which was shown. Error bars represent standard errors of the mean.

Presentation time

No main effect of presentation time was observed $F(1, 35) = 1.91, p = .176$, but there was a discrepancy size and presentation time interaction, $F(2, 57) = 4.13, p = .028, \eta_p^2 = .11$, and a significant discrepancy direction, discrepancy size, and presentation time interaction, $F(2, 70) = 3.21, p = .046, \eta_p^2 = .08$.

Sub-ANCOVAs performed on the positive and negative cue discrepancy conditions separately, revealed a significant discrepancy size $\times$ presentation time interaction for the positive discrepancies, $F(2, 70) = 9.19, p < .001, \eta_p^2 = .21$, but not the negative the discrepancies, $F(2, 70) = .23, p = .080$.

Paired samples t-tests conducted for the positive cue discrepancy condition revealed that when the cue/target discrepancy was small, responses were significantly more positive for 100ms stimulus presentations compared to 50ms presentations, $t(37) = 4.49 p < .001$. No significant differences in ratings were observed between 50ms and 100ms presentation times for either medium or large discrepancies (all $p > .05$)
There were no further significant interactions between any of the within subjects factors or covariates.

4.4 Discussion

The current study aimed to investigate the influence of several factors (participant optimism, stimulus context, and whether an expectation is positive or negative) in determining whether affective responses are assimilated toward, or contrasted against expectation. Since few studies have shown evidence of both assimilation and contrast effects in relation to affective expectations, a primary goal of the study was to attempt to promote both of these effects within the same experimental paradigm. The study was also designed specifically to examine the extent to which these effects are influenced by the size of the discrepancy between expectation and reality. The use of emotional faces meant that a range of stimuli could be presented which all existed along the same continuum, so a wide array of cue/target discrepancies could be included in the study.

Affective responses assimilated toward the values suggested by the cues to a greater extent in the sad emotion and cross emotion context conditions, when compared against the happy emotion context condition. Assimilation effects had been expected to reduce in the cross emotion condition, because the cue/target discrepancy should have been more noticeable when the cue and target were drawn from different emotional categories. However, the lack of assimilation effects for the happy emotion category suggests that there was an issue with the emotional face stimuli that had not been anticipated. The study was originally designed based upon the assumption that participants would find it equally difficult to discriminate between the individual levels of the happy and sad emotional expressions; this does not appear to have been the case. A happy face advantage has been reported (Calvo & Beltrán, 2013; Kiritā & Endo, 1995) where humans recognise happy expressions more quickly and more accurately than other types of expression. This perhaps goes some way toward explaining why the influence of discrepancy type was greater for the sad and cross emotion conditions, compared to the happy condition. In the happy condition participants presumably found it easier to identify each of the levels of emotion since they saw only happy expressions, with this being the case, it appears that they relied upon the cues to a lesser extent than in the other conditions. This also serves to highlight the difference between the measures of participant liking employed within the present experiment, and the previous experiments discussed within chapter 3. Whereas the previous experiments asked participants to rate their own liking of attractive faces, here participants were asked to judge the emotions of others, a task which is dependent upon an individual’s
ability to accurately perceive those emotions (emotional intelligence) (Salovey & Mayer, 1990). However, despite this difference, this ability can still be considered an important dimension of participant ‘liking’ since numerous studies have observed that the presence of valenced faces (presented either above or below the threshold of conscious perception) can produce an affective priming effect which can positively or negatively influence ratings of subsequently presented stimuli (e.g. Murphy & Zajonc, 1993).

The interaction between emotion, cue type, and discrepancy size revealed that when predictive cues were more positive than targets, assimilation effects were produced for faces from the sad emotion condition, and contrast effects were produced for faces from the happy emotion condition. When the discrepancy between the cue and the target was small, participants rated happy emotion condition intensities relatively accurately (an average of 0.50 levels higher than the target). By comparison when cue/target discrepancies were positive, but small, ratings for sad emotion condition intensities were less accurate (an average of 1.08 levels higher than the target). This finding supports the assertion that participants recognised happy expressions more accurately than sad expressions. Stapel and Winkielman (1998) suggest that contrast effects can occur as a result of a correction to an expected assimilative effect. When a prime is unobtrusive, assimilation usually occurs, but when a prime becomes more apparent and the participant becomes aware of the influence of the prime, correction contrast occurs (Strack, Schwarz, Bless, Kübler, & Wänke, 1993). In the current study, if it is assumed that participants could distinguish between the happy expression levels relatively successfully, then the priming manipulation would have become more apparent as the cues became less reliable predictors. This would appear to explain the pattern of results observed for the happy face condition when predictive cues were positive in nature. For small discrepancies participants assimilated toward the expectation suggested by the cues, but as the cue/target discrepancy increased correction contrast occurred.

When cue/target discrepancies were positive, sad condition faces produced the opposite pattern of results. Based on the assumption that participants found it harder to distinguish between sad expressions than happy expressions, this could explain the assimilation effects seen for the sad emotion condition. In the case of small cue/target discrepancies participants assimilated toward the affective value suggested by the cues, however, since they were poorer at recognising sad expressions, their attention was not drawn toward the influence of the prime. Consequently, participants continued to assimilate even as the cue/target discrepancy became larger. For the cross emotion condition, when cue/target discrepancies were positive
(so participants saw sad expression faces with happy expression cues) responses assimilated toward the values suggested by the cues, and this effect did not change according to the size of the cue/target discrepancy.

When cue/target discrepancies were negative, responses from the sad emotion condition assimilated toward the values suggested by the cues but did so to a lesser extent than for positive discrepancies. One explanation for this difference is that during the practice trials where there was no influence of the predictive cues, participants showed an overall bias toward rating emotions as slightly more positive than they actually were. For the happy emotion condition, when cue/target discrepancies were negative, small discrepancies did not seem to influence emotional ratings. This is illustrated by the very minor difference between ratings for happy emotion condition presentations over small discrepancies when paired with positive and negative cues (a difference of .50 and .51 respectively, between target and response). A reduction in the ratings of happy condition faces was observed according to discrepancy size, with assimilation effects increasing for medium compared to small discrepancies, but remaining similar between medium and large discrepancies. The reason for the differences in assimilation and contrast effects for the happy emotion condition between positive and negative cue differences is currently unclear, and previous research has suggested that assimilation and contrast effects should not alter if an expectation is positive or negative (Geers & Lassiter, 2002; Klaaren et al., 1994; T. D. Wilson et al., 1989). As this chapter has already noted however, there are currently few studies which have explored assimilation and contrast effects using interval scaled stimuli, so it is possible that the nature of an expectation could act as a moderating factor when expectancy/reality discrepancies are modified slowly over time.

Contrast effects were observed for the cross emotion condition when the cues predicted more negative outcomes. Responses to small discrepancies assimilated toward the values suggested by the cues, but as the size of the cue/target discrepancy increased, targets were evaluated in the opposite direction to the cues. Since both the happy and cross emotion conditions involved the observation of only happy expressions when they were presented with negative cues, it might initially seem surprising that contrast effects were observed for only the cross emotion condition. Yet this could be explained by the prime being more obtrusive for the cross emotion condition. Whilst both the cross emotion and happy conditions involved the presentation of only happy expressions (which participants were relatively good at recognising), the task was arguably easier for the cross emotion condition, since the cues and the targets in this condition were drawn from different emotional categories.
The results of the optimism/pessimism measure support the findings reported previously in this area. Morton et al. (2011) reported that the influence of small and medium sized expectancy discrepancies was greater for more optimistic individuals, but there was no difference between optimists and pessimists for large discrepancies. The present study has replicated these findings, but once again the effect appears to be moderated by the nature of the expectation. When a cue predicted a more positive emotion than the target, but the discrepancy between the two was small, optimists rated emotions more positively than pessimists. Optimists also rated emotions more positively than pessimists when there was a medium sized negative cue/target discrepancy. Like the Morton et al. (2011) study, the current results showed no difference between optimists and pessimists when the cue/target discrepancy was large, regardless of the nature of that discrepancy. It is currently unclear why these differences have been observed between positive and negative cue/target discrepancies, and it is possible that if the previous study had also recorded the direction of cue/target discrepancies, the current findings may have simply replicated those results. Unfortunately because that study looked at cue acquiescence only, the results are not directly comparable to the current study.

One issue which the current study did not address is the time course of assimilation and contrast effects. This chapter has put forward the argument that larger discrepancies between the cue and stimulus make the experimental manipulation more apparent which leads to correction contrast. There is also research which suggests that repeated exposure to a source of information deemed ‘untrustworthy’, engages a distrust mindset (see Mayo, 2015 for review). Once engaged, the distrust mindset activates alternatives to primed concepts, and impedes confirmatory biases. The mild deception was included in the present study (the claim that the emotional descriptors were based upon responses from a previous group) seems fairly plausible, but it is likely that participants began to question this assertion as the cue/target relationship became more discrepant over the course of the session. What is not clear, is whether the influence of the cues diminished as soon as cue/target discrepancies were acknowledged, or whether the effects of distrust developed slowly over the course of the experimental session.

As a final note, although there was a significant presentation time x discrepancy size x discrepancy direction interaction, the influence of presentation time appeared to be minimal. Presentation time did not influence assimilation effects for negative cues, and for positive cues presentation time was only influential over small discrepancies. Furthermore, when presentation time was influential, the effect was counter to what had been predicted. When
presentation time was reduced participants had been expected to assimilate further toward the cues, because the target stimuli had been made more ambiguous, but assimilation effects were greater for the 100ms over the 50ms presentation times. In a study investigating the influence of display duration upon reaction times, Calvo and Lundqvist (2008) observed that facial expressions were still identified reasonably accurately for exposures of 100ms and 50ms, and only minimal differences in accuracy were observed between the two presentation durations (85.4% and 84.6% accuracy respectively). Therefore, it seems possible that the presentation times employed in the current study were a little conservative, and greater assimilation effects may have been observed if presentation times of 25ms or lower had been included. It is also possible that the effects of presentation time may have been seen in other behavioural measures such as reaction time or gaze direction, but the design of the current experiment (collecting responses from multiple participants in a single session) did not lend itself to collecting data of this type.

4.5 Conclusion
The current study manipulated participant expectations and observed evidence of assimilation and contrast effects by varying the size of the discrepancy between an affective expectation (caused by a predictive cue), and a target stimulus. The results also support the findings of previous studies which have reported that optimists assimilate toward cues more readily than pessimists, and that contrast effects occur when larger discrepancies are experienced between expected outcomes and reality. Differences in assimilation and contrast effects were observed according to the positive/negative nature of the predictive cue, which was unexpected and not in line with the results of previous studies, but these findings can perhaps be attributed to the unequal nature of the happy and sad emotional stimulus categories used during the current study.
5 EXAMINING THE TIME-COURSE OF ASSIMILATION EFFECTS USING FOOD IMAGES

5.1 Introduction

In chapter 3 a series of attractive and unattractive faces were used to demonstrate that expectations cause people to adopt different affective thresholds when judging the same stimuli, and affective responses assimilate toward expectations about whether stimuli will be liked or disliked. Faces represent a ‘special’ class of stimuli which are processed differently to other stimuli by face-selective areas of the brain like the fusiform face area (Farah, 1996). Since chapters 3 and 4 investigated assimilation and contrast effects using faces, it was important to investigate whether those results could be replicated using different stimuli. Foods are a class of stimuli which have often been used as rewards in studies of human and animal behaviour, and like faces, foods vary in terms of their rewarding value (see Higgs, 2015 for review), meaning multiple different foods can be used to create a series of stimuli that vary in affective value along a continuum. In the present chapter, a rating task was employed using a set of foods which varied in pleasantness, to test whether expectations about those foods produced affective assimilation effects similar to those discussed in chapter 3. The scope of the experiment was also expanded to investigate whether these assimilation effects persisted over multiple sessions conducted over the course of a week.

Assimilation effects have been observed in several different food related studies where product liking, and product choice have been influenced by expectations about a number of different factors including: where a product was allegedly sourced (Cerjak, Karolyi, & Kovacic, 2011; Stefani et al., 2006), how a product was allegedly produced (Caporale & Monteleone, 2004; Van Wezemaet al., 2012), what ingredients were allegedly used to create a product (L. Lee et al., 2006; Tarancon, Sanz, Fiszman, & Tarrega, 2014; Torres-moreno et al., 2012), how healthy or unhealthy a product allegedly is (Norton et al., 2013), or how a product is presented visually (De Wijk, Polet, Engelen, Van Doorn, & Prinz, 2004; Morrot, Brochet, & Dubourdieu, 2001). Like the assimilation and contrast literature in general, there are fewer studies which have observed contrast effects in relation to expectations about food products. However, there are some examples such as Cardello and Sawyer (1992) or Zellner, Strickhouser, and Tornow (2004).

A commonality shared by all of the studies listed above is that they have manipulated expectations using stimuli which were either purely categorical in nature, or which had been operationalised as being categorical in order to create an expectation (e.g. rather than providing...
a precise measure of fat content, a product would instead be described as being reduced-fat or full-fat). In chapters 3 and 4, interval scaled stimuli were used in order to explore the relationship between discrepancy size and affective assimilation and contrast, and the issue of discrepancy size is of particular interest for the study of expectations about food products. Unhealthy dietary habits are a global risk to health, consequently the World Health Organisation (WHO) recommends that sugars should represent less than 10% of total energy intake, and salt intake should be restricted to less than 5g per day, but in many cases national averages are well above these levels (World Health Organisation, 2015). In line with these recommendations, WHO member states have made commitments to reduce these levels, but this poses a problem for manufacturers in terms of creating products which contain fewer unhealthy ingredients, yet still maintain an expected sensory profile. As interest in this area has grown, several studies have shown evidence of assimilation effects using interval scaled stimuli by manipulating the sugar (Kuenzel, Zandstra, et al., 2011; Woods et al., 2010), or fat content (see Zandstra, Lion, & Newson, 2016 for review), of foods and beverages.

Elsewhere, other studies involving foods have offered some support for the assimilation-contrast approach, where small differences between an expectation and a product lead to assimilation effects, and larger differences lead to contrast effects. For example, in a study by Yeomans, Chambers, Blumenthal, and Blake (2008) participants were presented with a fish product; for some participants the product was labelled as a frozen savoury mousse, and for other participants the product was labelled as an ice-cream. Higher acceptance scores were observed in participants when the product was described as a savoury mousse, presumably because the discrepancy between the expectation and the product was smaller in the mousse condition than in the ice-cream condition. Only a small number of studies involving food products have observed both assimilation and contrast effects, but one recent study conducted by Davidenko et al. (2015) has shown evidence of assimilation-contrast effects, again using interval scaled, rather than categorical stimuli. Participants sampled a number of lemon drinks which varied in sugar and flavour concentration, and learned to associate drink intensities with novel symbols. Participants’ hedonic ratings generally assimilated towards the intensities suggested by the labels, but contrast effects were observed in specific cases; where participants expected a weak drink but received a strong drink and vice-versa.

Expectations about foods have also produced placebo-like effects where participants have demonstrated beneficial effects from foods in the absence of an expected ingredient. Examples of such effects include: increased alertness and faster responding based upon the assumption
that a decaffeinated beverage contained caffeine (C. Anderson & Horne, 2008; Flaten & Blumenthal, 1999; Kirsch & Rosadino, 1993), or increased feelings of satiety from energy drinks based upon their consistency (Chambers, Ells, & Yeomans, 2013; McCrickerd, Chambers, & Yeomans, 2014). One question that naturally arises from the observation of these expectancy effects relates to their persistence. If participants’ expectations are sufficient to mask the reduction or absence of an expected ingredient over a single exposure, does this effect persist for future exposures, and how long does the effect last before it reaches the point of extinction? Research in this area is limited, but studies investigating the placebo analgesia effect, where the expectation of a therapeutic outcome causes an inert treatment to have an analgesic effect, have shown that placebo effects can last for up to 7 days (Colloca & Benedetti, 2006).

According to Tuorila, Cardello, and Lesher (1994), expectations about food products should remain relatively stable over time, even in the face of disconfirmatory experiences. In a study using a number of different foods, they investigated the persistence of expectations relating to reduced-fat and regular-fat products. At the start of the study participants reported that they expected reduced-fat products to taste poorer than the equivalent regular-fat products. This expectation was disconfirmed in a blind taste test, with no difference in liking being reported between the regular and reduced-fat product variants. However, when participants attended an additional session one month later, they still retained their initial expectations that the reduced-fat products would taste worse. The authors also found that expected liking was best predicted by participants’ existing familiarity with a product.

There are currently few studies which have investigated the persistence of affective expectations about foods, but one study which has attempted to investigate the longer-term influence of a recent positive or negative food experience, was conducted by Robinson, Blissett, and Higgs (2013). In a first experiment, participants rated a number of snack foods according to expected liking; they then tasted samples of those foods and provided a second rating based upon their actual experience. The food item which produced the largest negative difference, in terms of expected vs. actual liking, was classified as the “disappointing food”. In a second session, participants provided expected liking ratings for a number of different foods, as well as the disappointing food from the previous session. The experimenters reported that expected liking ratings for the disappointing food reduced if the second session occurred one day later, but ratings returned to baseline levels if the second session occurred one week later. Furthermore, a second experiment using a similar methodology suggested that this effect may have been moderated by product familiarity. In this second experiment, a disappointing food
experience reduced expected liking one week later for infrequently eaten foods, but not for frequently eaten foods.

Due to the lack of research in this area, the current study was conceived to investigate the persistence of assimilation effects over time, and to do so using interval scaled stimuli. In addition to recording explicit measures of preference, the study also included an implicit preference measure. Broadly speaking, explicit attitude measures more reliably predict deliberate well-thought out behaviours, whilst implicit measures better predict more spontaneous behaviours, such as when a consumer is making a rapid choice about a product in a supermarket (Richetin, Mattavelli, & Perugini, 2015; Spence & Townsend, 2006). Studies of food choice have shown that, in some cases, implicit measures can better predict consumers’ actual behaviour than explicit measures (Maison, Greenwald, & Bruin, 2004; Perugini, 2005), but when compared to explicit attitudes, implicit attitudes are more rigid and less susceptible to short-term change (Hermans, Baeyens, & Eelen, 2003). With this in mind, an implicit measure was included at the end of the current study, to see whether any change in explicitly measured assimilation effects was mirrored in participants’ implicit attitudes.

To recap, the objectives of the current study were threefold: (i) to essentially repeat experiment 2 from chapter 3 to see if those findings generalised to other interval scaled stimuli outside of faces, (ii) to investigate whether assimilation effects changed over time, and (iii) to investigate whether assimilation effects from explicit measures were replicated in implicit measures.

The general procedure was similar to the one discussed in experiment 2, but the study was divided into 3 sessions. In the first session predictive cues were associated with images of highly pleasant and highly unpleasant foods, and images of moderately pleasant foods were slowly introduced over the course of the experiment, paired with the same cues. Like the earlier experiment, it was anticipated that the affective associations learned for each cue would generalise to the moderately pleasant food items. Participants were not familiarised with any of the food items beforehand, and expected pleasantness ratings for each item were based upon ratings taken from the database where the majority of the food images used in the study were obtained (see Foroni, Pergola, Argiris, & Rumiati, 2013). The second session was conducted 3 days later, and repeated the same procedure as the first session. The final session was conducted 7 days after the first, and used the Go/No-Go association task (GNAT) (Nosek & Banaji, 2001) to examine participants’ implicit attitudes toward the moderately pleasant items which had previously been paired with the positively and negatively associated cues, during the first and second sessions.
Experiment 4

5.2 Methods

5.2.1 Participants
Thirty-three participants aged between 20 and 35 (19 females, mean age = 26, SD = 3.97) were recruited for the study from the undergraduate population at the University of Manchester. Participants were required to attend 3 sessions separated over the course of a week.

All participants reported normal or corrected to normal vision, and had no reported history of neurological or psychiatric disorders. Participants were paid £10 for their time and all gave written informed consent to participate. All testing procedures were approved by the University of Manchester ethics committee.

5.2.2 Stimuli and apparatus
A series of 88 images of foods were taken from the FoodCast research image database (FRIDa) (Foroni et al., 2013). Images were selected based upon the average valence ratings reported in the database, where each image had been rated by 73 individuals on a scale ranging from ‘very negative’ (0) to ‘very positive’ (100). For the purposes of the current study foods were considered in terms of their place along an affective continuum, determined by these ratings. The food items were placed into the following grouping along this continuum: A positive food group (predominantly consisting of sweet foods e.g. fresh fruit or chocolate), a negative food group (consisting of spoiled or rotten foods), and six intermediate food groups whose average valence ratings increased from group 1 up to group 6 (see Figure 5.1). Each of the intermediate food groups included images of 4 different food items taken from the FRIDa. Three additional exemplars of each food item were sourced online so that each intermediate group contained 16 images total (4 exemplars of each food item). The positive and negative food groups included 31, and 33 images sourced from the FRIDa. Both groups were supplemented by additional exemplars sourced online so that they each contained 50 items in total. An additional 24 images of non-food items (kitchen utensils) were also taken from the same database, and the non-food items were selected at random from the database’s artificial food objects category.
Figure 5.1 Valence ratings for food groups used in experiment 4. Ratings reported here are based upon the ratings reported in the FRIDa (Foroni et al., 2013). Images represent 2 exemplars of the items used within each group. Error bars represent standard errors of the mean.

In total the study employed a stimulus set consisting of 220 images: 24 non-food items, 96 intermediate food items (4 examples of 24 different foods), 50 positive food items, and 50 negative food items (see Appendix C.1 for a full list of the items used in the study). Each image depicted a single item presented against a white background, and all images were saved in PNG format and cropped to a size of 530 by 530 pixels. All Stimuli were presented using a desktop computer running PsychoPy software (Peirce, 2007) on a 21-inch LCD monitor at a resolution of 1920 by 1080 pixels. All responses were recorded using a standard wired mouse and keyboard placed on a desk directly in front of the participant.

5.2.3 Design and procedure
The study was divided into 3 different sessions; each session lasted approximately 25 minutes and participants were required to complete all 3 sessions. Sessions were lab-based, and all involved presentations of various items on a computer screen. In sessions 1 and 2 participants
completed a rating task involving various foodstuffs, and in session 3 they completed a version of the GNAT (Nosek & Banaji, 2001).

The rating task involved short presentations of foods paired with, and preceded by visual cues. After each presentation participants rated the pleasantness of the preceding item using a visual analogue scale (VAS). Throughout the task, to anchor participant responses, different predictive cues were paired with the positive and negative food groups. Additionally, to test whether ratings for the positive and negative items would generalise to the intermediate items, half of the intermediate items were paired with the positively associated cue, and half were paired with the negatively associated cue.

The GNAT is a response competition task derived from the Implicit association task (Greenwald et al., 1998), and it is designed to measure implicit attitudes to certain categories of items (see general methods for additional details). In the case of the current experiment the GNAT was employed during the third session, to measure implicit attitudes to intermediate food items which had previously been paired with positively or negatively associated cues. The GNAT was chosen over the implicit association task (IAT) (Greenwald et al., 1998) because unlike the IAT, it can be used to measure implicit attitudes toward a single target category. If the study had used the IAT, which requires some form of comparison category, attention would have been drawn to the experimental manipulation (namely which items were previously presented with the positively and negatively associated cues).

**Session 1: Rating Task 1**

Participants were informed that they would be observing brief presentations of a series of foods on a computer screen, and that each presentation would be preceded by one of two sets of symbols. Participants were told that after each presentation, they would be asked to provide a pleasantness rating for the preceding item as quickly as possible using the mouse and an on-screen VAS.

Each item was paired with one of the two sets of predictive cues throughout the task (either blue squares or yellow circles). Cues were assigned to food items based upon the pre-defined category from which the food item was drawn. One cue was consistently paired with the positive food group, and the other cue was consistently paired with the negative group. Within each of the intermediate food groups, 2 items were paired with the positively associated cue throughout the task, and the remaining 2 items were paired with the negatively associated cue. Cue pairings for intermediate items were also counterbalanced across participants.
The task consisted of 148 trials total, and involved presentations of the 50 positive food items, the 50 negative food items, and presentations of 2 exemplars of each of the 24 intermediate food items. The task was divided into four blocks of 37 trials, separated by breaks. In order to weight the earlier blocks in favour of the positive/negative anchors, trials were presented in a pseudo-random order which included the following constraints: block 1 (100% positive/negative items), block 2 (73% positive/negative items, 27% intermediate items), blocks 3 and 4 (48% positive/negative items, 52% intermediate items).

Each trial began with a central fixation cross presented for 2000ms, followed by the presentation of a predictive cue for 2000ms. The target stimulus was then presented for 200ms, and during this time the cue also remained on-screen. After the cue and target disappeared, a rating screen appeared asking the participant ‘How pleasant was the image?’ Participants responded using the mouse by selecting a segment on the VAS, which was divided into 10 segments ranging from 0 (Highly unpleasant) to 10 (highly pleasant). In each trial the participant was given 5000ms to respond, and the remaining time in the current trial (in seconds) was displayed by an on-screen countdown timer (see Figure 5.2). Trials where the participant failed to respond in time were not included in the analysis. During the stimulus and cue presentations the mouse cursor was occluded; the cursor only became visible once the rating screen appeared. To prevent reaction time biases caused by responses from previous trials, the mouse cursor was assigned to a random position at the start of each trial.
Session 2: Rating Task 2

Session 2 took place 3 days after the completion of session 1, and involved a rating task which was virtually identical to the one employed in session 1. Stimulus timings and trial structure remained the same, but for each of the intermediate food groups, 2 different exemplars of each item were used. To test whether affective associations still generalised to the intermediate items, the positive and negative anchor items were not included the task. For the intermediate items, cue pairings transferred to the new exemplars (e.g. If item NF095 (an egg) had been paired with the positively associated cue in session 1, the same cue and stimulus pairing was present in session 2, but a different image of the item was used).

Session 3: GNAT

The final session was conducted 4 days after session 2 and involved a 6 block version of the GNAT. Here participants were instructed to respond to images of foods, and to not respond to images of kitchen utensils (distractor items). Each of the critical blocks contained images of the intermediate foods which had been paired with positive or negative predictive cues in the previous tasks, but during the GNAT the foods were presented without the accompanying cues (see table 5.1 for details of the block design). In the critical blocks, participants were instructed...
to also respond to words of a particular type (good or bad) whilst ignoring words of a different type (good or bad) (see Appendix C.2 for a complete list of the good and bad words).

Table 5.1  

<table>
<thead>
<tr>
<th>Block</th>
<th>Trials</th>
<th>Stimuli assigned to GO</th>
<th>Stimuli assigned to NO GO</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>48b</td>
<td>Positive Intermediate + Good Words</td>
<td>Non-Food Objects + Bad Words</td>
</tr>
<tr>
<td>4</td>
<td>48</td>
<td>Positive Intermediate + Bad Words</td>
<td>Non-Food Objects + Good Words</td>
</tr>
<tr>
<td>5</td>
<td>48</td>
<td>Negative Intermediate + Good Words</td>
<td>Non-Food Objects + Bad Words</td>
</tr>
<tr>
<td>6</td>
<td>48</td>
<td>Negative Intermediate + Bad Words</td>
<td>Non-Food Objects + Good Words</td>
</tr>
</tbody>
</table>

Note:  
1. In the Food Negative + Good to Food Positive + Bad sequence, blocks 3 and 4, were swapped with blocks 5 and 6. The order of blocks 3 and 4, and blocks 5 and 6 were also counterbalanced across participants.  
2. 24 image (12 food + 12 utensil) trials and 24 word (12 good + 12 bad) trials.

At the beginning of the session the task was explained by the experimenter, and participants were given the opportunity to complete 2 practice blocks. In the practice blocks participants were only required to classify items along a single category attribute (foods or non-foods), and presentations from these blocks involved other moderately rated items taken from the FRIDa which were not seen at any other stage of the experiment. Throughout the task target category labels were presented at the top left and right of the screen. Labels were colour coded so that participants knew which category was for classifying images, and which category was for classifying words. The horizontal position of the labels was also counterbalanced across participants.

Participants were instructed during the initial briefing that they should try to respond as quickly as possible for all of the trials. In each trial, an item (either a word or an image) was presented centrally for 700ms. Whilst the item remained on-screen, participants had to decide whether to respond by pressing the space bar (GO for target items), or to not respond (NO-GO for distractor items). If no response was provided within the 700ms that the item remained on-screen, the trial was considered a no response trial, and appropriate feedback was provided. When the participant was correct to provide no response, a green circle was presented for
200ms at the end of the trial; in cases where the participant had been incorrect to provide no response, a red X was presented for 200ms.

In trials where a response was provided whilst the item remained on-screen, feedback appeared immediately, directly below the item. The centrally presented item and feedback presentation continued until 700ms after the item had first appeared. The feedback then remained on-screen in isolation for a further 200ms, and the next trial began immediately after the feedback was removed. The symbols used for feedback were identical for both the GO and NO-GO trials (see Figure 5.3 for examples).

---

**Figure 5.3** Examples of correct and incorrect responses to ‘Food + Good Word’ sequence GNAT trials. In these examples the participant is expected to respond to images of foods or presentations of ‘Good’ words, failing to respond to these items will result in negative feedback. The opposite is true for images of utensils or presentations of ‘Bad’ words.

---

### 5.2.4 Data analysis

Note: Like experiment 2, response data was originally going to be analysed by plotting psychometric functions from participants’ responses. To achieve this, participants’ ratings were converted to a proportion of liked and disliked responses at each intermediate level. Ratings above 5 were coded as liked, and ratings below 5 were coded as disliked. However, functions
could not be fitted with meaningful PSEs and slopes from the majority of participants’ data, so this method of analysis was abandoned (see Appendix C.3 for example functions).

For the rating task, an average rating was calculated for the positively and negatively associated food items belonging to each intermediate food group. Separate ratings were calculated for each participant in session 1 and session 2. The average ratings for the positively and negatively associated intermediate food items at each level were then subjected to a three-way repeated measures ANOVA, with the following within-subjects factors: food group (6 levels: intermediate food groups 1 - 6), cue (2 levels: positively or negatively associated cues), and session (2 levels: session 1, session 2). A separate two way repeated measures ANOVA was used to investigate the ratings of the positive and negative anchor items. This analysis included the following factors: anchor type (2 levels: positive, negative) and session (2 levels: session 1, session 2).

In cases where Mauchly’s test revealed that the assumption of sphericity had been violated, the degrees of freedom have been corrected using Greenhouse-Geisser estimates of sphericity. An alpha level of .05 has been used for all statistical tests, and effect sizes are reported here as partial eta squared ($\eta^2_p$) where 0.05, 0.1, and 0.2 correspond to small, medium, and large effects, respectively (Cohen, 1988).

The data from the GNAT were used to calculate a $D$ value using the methods described in chapter 2. Higher $D$ values indicated a stronger association between a group of food items, and the concept of ‘goodness’. $D$ values from the GNAT containing food items which had previously been paired with positive cues, were compared against $D$ values from the GNAT containing food items previously paired with negative cues, using paired samples t-tests.
5.3 Results

5.3.1 Food group ratings

The food group x cue x time ANOVA revealed a main effect of food group, $F(3, 106) = 20.95, p < .001, \eta^2 = .40$. A gradual increase in pleasantness ratings was observed over the 6 intermediate food groups and contrasts revealed a significant linear trend in the food group data $F(1, 32) = 46.03, p < .001, \eta^2 = .59$. However, when this was followed up with pairwise comparisons ($p$ values are Bonferroni corrected), it was revealed that the extent of this increase was not consistent across all levels (see Table 5.4).

Table 5.2

<table>
<thead>
<tr>
<th>Group ID</th>
<th>Mean Rating</th>
<th>SE</th>
<th>Comparison to previous group (Mean Diff ± SE)</th>
<th>Sig (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate 1</td>
<td>5.039</td>
<td>.20</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Intermediate 2</td>
<td>6.071</td>
<td>.20</td>
<td>+ 1.03 (.20)</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>Intermediate 3</td>
<td>6.166</td>
<td>.19</td>
<td>+ .095 (.13)</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Intermediate 4</td>
<td>6.179</td>
<td>.18</td>
<td>+ 0.13 (.17)</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Intermediate 5</td>
<td>6.627</td>
<td>.19</td>
<td>+ .448 (.13)</td>
<td>.032*</td>
</tr>
<tr>
<td>Intermediate 6</td>
<td>6.667</td>
<td>.18</td>
<td>+ .040 (.15)</td>
<td>&gt; .05</td>
</tr>
</tbody>
</table>

Note: * Denotes significant $p$ values

5.3.2 Influence of cue

Participants rated intermediate items paired with the negative cue as slightly more pleasant ($M = 6.15, SD = 1.60$) than those paired with the positive cue ($M = 6.10, SD = 1.51$), but no main effect of cue was observed $F(1, 32) = .22, p = .644$.

5.3.3 Influence of session

A main effect of time was observed with pleasantness ratings collected at time 2 ($M = 6.27, SD = 1.47$) being higher compared to those collected at time 1 ($M = 5.98, SD = 1.61$), $F(1, 32) = 6.97, p = .013, \eta^2 = .18$. No significant interactions were observed, but the session x cue interaction approached significance $F(1, 32) = 3.09, p = .089$ (see Figure 5.3).
5.3.4 Positive / Negative anchors

The results of the anchor type x time ANOVA for the positive and negative anchors revealed a main effect of anchor type with positive anchors being rated as more pleasant ($M = 7.66, SD = .90$) than negative anchors ($M = 1.23, SD = .90$), $F(1,32) = 704.79, p < .001, \eta^2_p = .96$.

No main effect of time was observed, $F(1, 32) = .62, p = .436$, and there was no interaction between anchor type and session, $F(1, 32) = .19, p = .666$, indicating that pleasantness ratings for the two types of anchor remained fairly consistent across the two sessions.

5.3.5 GNAT data

No participants had to be removed for producing unusually high error rates or unusually fast response times, consequently the results reported here are based upon GNAT data acquired from the entire sample of 33 participants. Paired samples t-tests revealed no difference in $D$ values between intermediate food items previously paired with positive cues ($M = .19, SD = .06$), and intermediate food items previously paired with negative cues ($M = .21, SD = .10$), $t(32) = -.18, p = .86$.

Further paired samples t-tests revealed that participants were significantly faster (times are reported in seconds) when responding to images of foods and good words, compared to when

Figure 5.4  Non-significant interaction between session and cue for ratings of intermediate food objects. Error bars represent standard errors of the means.
images of foods and bad words. This was true for both foods that had previously been paired with positive cues during the rating sessions (food + good: $M = 501.20\text{ms, SD} = 4.19\text{ms}$; food + bad: $M = 514.30\text{ms, SD} = 5.09\text{ms}$) $t(32) = -2.28, p = .030$, and foods which had previously been presented with negative cues (food + good: $M = 503.68\text{ms, SD} = 6.46\text{ms}$; food + bad: $M = 520.24\text{ms, SD} = 35.70\text{ms}$) $t(32) = -3.05, p = .005$.

Participants also made significantly fewer errors when foods were paired with good words, and again this finding was repeated for items previously paired with positive cues (food + good: $M = 3.03, SD = 0.46$; food + bad: $M = 4.82 SD = .57$), $t(32) = -3.26, p = .003$, and items previously paired with negative cues (food + good: $M = 3.15, SD = .48$; food + bad: $M = 4.91 SD = .66$) $t(32) = -3.05, p = .005$.

5.4 Discussion

The primary aim of the current study was to investigate the persistence of affective assimilation effects relating to food products, but in terms of achieving this aim, the study was unsuccessful. The results of the study showed no overall cueing effect, and no interaction between cue type and any other factor. Consequently it is impossible to comment on the changes in assimilation effects over the course of the study, since participants’ ratings did not appear to assimilate toward the values suggested by the cues at any stage during the experiment.

The ANOVA conducted on the liking responses for the intermediate items revealed a significant main effect of food group, and pairwise comparisons revealed that liking increased in a broadly linear fashion across the 6 intermediate food groups. Additionally, the ANOVA conducted on the positive and negative anchor items, showed that the positive anchors were liked significantly more than the negative anchors (a mean difference of 6.43 between anchor types). These results had been roughly predicted based upon the ratings taken from the FRIDa, and the literature suggests that assimilation is likely to occur when ambiguous stimuli are rated in the context of moderately extreme stimuli (Martin, Seta, & Crelia, 1990).

One of the factors which likely to have influenced the current findings is the ambiguity, or more precisely, the lack of ambiguity, associated with the intermediate food groups. Studies involving measures of visual perception have indicated that stimulus ambiguity plays a key role in modulating top-down effects. When object recognition is difficult, and an observer has reason to expect a particular type of object (e.g. due to prior probability, or contextual information), top-down processes are more influential, offering predictions about an object’s possible identity. Alternatively, when recognition is sufficiently easy, it may be completed so quickly
through bottom-up analysis alone, that top-down facilitation does not have the time to develop (Bar, 2003; Summerfield et al., 2006). Similarly, advertisements have been shown to have a greater assimilative influence on product perception, when consumers are provided with ambiguous, rather than unambiguous, information about product quality (Hoch & Ha, 1986). In a change to the earlier experiments, the target stimulus presentation times were extended from 80ms in experiment 2, to 200ms in the current experiment. The reason behind this change was that participants were expected to require more time to process images of foods, compared to images of faces, because humans are considered experts in terms of facial processing (Tanaka, 2001). However, it appears that by making this change the food images were disambiguated to the point where participants did not need to use the predictive cues to guide their ratings. If participants were never forced to attend to the predictive cues, then this may have also negatively influenced the utility of the anchor items. This is because anchors are most effective when they are considered compatible with the item under scrutiny (Paek et al., 2011). In the case of the current experiment it appears that participants never formed the expected stimulus/cue associations, and it therefore seems reasonable to assume that participants also did not view the anchor items as a useful means of judging the intermediate items.

Another possible explanation for the absence of the assimilation effects which had been observed in the earlier studies, relates to the types of stimuli participants were asked to rate. Both the faces used in the earlier study, and the foods used in the current study were examples of moderately liked images which were varied along a continuum (attractiveness in the case of faces, or pleasantness in the case of foods). Nevertheless, faces represent a class of stimuli which have high intra-class similarity (Werheid & Clare, 2007). All faces share common elements like the approximate location of features like the eyes, ears, lips etc., whereas the foods that participants were asked to rate in the current study did not share similar commonalities. As a result, participants in the earlier study would have had to process the stimuli in a more fine-grained manner, and because of the additional cognitive resources required to do this, this would have resulted in a greater reliance upon existing expectations (Dunovan, Tremel, & Wheeler, 2014). Additionally, because the face stimuli were reasonably well controlled, it seems reasonable to assume that they were being judged along the dimension of attractiveness. In the current study, although all of the items we used in the rating task could be classified as foods, the items themselves were all highly different from one another, and that they could potentially be analysed in a coarser fashion. Even over short presentation times, it was relatively easy to determine that an image contained an instance of a particular food type, and it is also impossible to determine the exact criteria that stimuli were being judged upon, since they
varied along a numbers of dimensions such as: colour, size (relative to the background), or shape. This does not, however, mean that faces represent the only class of stimuli which possess such a high degree of intra-class similarity, and studies investigating facial processing have used objects like chairs or houses as comparison categories because they have comparable levels of intra-class similarity (Ishai, Ungerleider, Martin, & Haxby, 2000). A challenge for the design of future studies in this area is to identify stimuli which have similar intra-class similarity, and which can be manipulated along a single dimension to increase or decrease liking.

Stimulus familiarity could also account for the lack of assimilation effects in the current study. Different exemplars of the intermediate foods were used during sessions 1 and 2, and participants did not observe any of the images prior to the rating tasks, but no attempt was made to control for previous experience with the types of foods used in the study. All of the foods used in the study were fairly common, and previous studies have found it difficult to influence expectations about products when there is a high degree of existing familiarity (Kuenzel et al., 2010; E. Robinson et al., 2013; Tuorila et al., 1994). This is also likely to have further diminished the effect of the positive and negative anchors since previous research has indicated that anchoring effects are diminished when observers are knowledgeable about a judgement that they are asked to make. This because they can retrieve information relating to a stimulus judgement directly from memory rather than needing to assess the properties of the stimulus at hand (Englich & Mussweiler, 2001; Englich & Soder, 2009; T. D. Wilson et al., 1996).

The current study also only used images of foods, and participants did not have access to the type of sensory information that is often used to assess the quality of food products (taste, texture, smell etc.). As this chapter has already noted, the stimuli themselves were quite unambiguous. Therefore, it seems probable, given the lack of additional sensory information, and the probability that most people had sampled the majority of the intermediate food items in the past, that participants based their pleasantness ratings on their existing knowledge and preferences for those items, rather than basing their ratings on the specific exemplar that they were shown. This issue could be resolved to some extent, if a single type of food was used instead, and that food varied in quality somehow (e.g. levels of ripeness for a fruit), because this would mean that all of the items were more similar on a feature level, and participants would be forced to attend more to the specific presented exemplar. However, this would need to be done with a rare, or possibly conceptual food type, otherwise it would still be difficult to control for the influence of past experience.
A significant increase in overall ratings was observed for the intermediate foods between session 1 and session 2, but this finding was not unexpected, and the increase in ratings between the two sessions can be explained by the mere-exposure effect (Zajonc, 1968) where a preference for a particular stimulus is developed purely by repeated exposure to that stimulus. The non-significant interaction between cue and session type suggests that there is a possibility that the cues were beginning to slowly influence participant responses, but since data was only collected from 2 sessions it is impossible to say whether the effect would have continued in the same direction in future sessions.

For the implicit attitude measure, the GNAT, no difference in implicit ratings was observed between items which had previously been paired with the positive and negative cues. This was not wholly unexpected, given the rigidity of implicit attitudes, especially for familiar items (Hermans et al., 2003). In the context of the current experiment, this finding is even less informative, due to the lack of any assimilation effects at any stage of the experiment, so it remains to be seen whether changes in affective assimilation effects over time are reflected in implicit attitudes.

Outside of issues relating to the use of cues, the affective similarity of the intermediate food groups provides further evidence that the stimuli used in the current study were not suitable for investigating the boundaries of assimilation effects. Originally, psychometric functions were going to be plotted from participants’ affective responses, and a gradual increase in pleasantness ratings was observed from intermediate group 1 up to group 6. In spite of this finding, the extremely shallow slopes that were observed from the functions (see Appendix C.3 for examples) indicate that participants were not very sensitive to changes in the stimuli, and that the intermediate groups were essentially homogenous in respect of their affective values. This finding could also be the result of using a VAS to record participant responses in the current experiment, rather than a 2AFC paradigm like the previous experiment, if some participants did not use the entirety of the scale. Yet very shallow slopes were observed for the majority of participants, so this indicates that there was an issue with the stimuli, rather than the measurement technique.
5.5 Conclusion

In summary, this study aimed to build upon the findings of chapter 3 to investigate the longer-term persistence of affective assimilation effects. However, the study was unsuccessful in this respect, and assimilation effects were not observed at any stage of the study. The absence of assimilation effects can be attributed to a number of methodological issues, particularly a lack of stimulus ambiguity, which is highly influential in determining the extent to which people rely upon existing expectations when processing affective or perceptual information. The failure of the current study, in terms of producing affective assimilation effects, means that there are still unresolved questions relating to the persistence of affective assimilation effects. Consequently, these questions remain the primary focus for the experiments discussed in chapter 6.
6 USING CURVATURE TO EXAMINE THE PERSISTENCE OF ASSIMILATION 
AND CONTRAST EFFECTS THROUGH EXPECTATIONS ABOUT AFFECTIVE 
CONSTANCY

6.1 Introduction

The aim of the previous chapter had been to use a series of food images to examine the 
persistence of affective assimilation effects brought about by predictive cues. However, 
associating positive and negative values with images of familiar food items proved challenging 
because affective predictions about foods are highly influenced by past experiences with foods 
of a similar type (E. Robinson, Blissett, & Higgs, 2011). This factor is also difficult to control for, 
since the average consumer is familiar with a wide range of different foods. Before an attempt 
was made to repeat experiment 4, an alternative set of stimuli needed to be identified where 
affective judgements would not be similarly influenced by stimulus familiarity. In order to be 
suitable for use in the same task, these stimuli also had to fulfil two additional requirements: (i) 
They need to be adjustable along a continuum where alterations to the properties of the stimuli 
allied with increases or decreases in liking, and (ii), alterations to the stimuli need to be difficult 
to perceive to encourage top-down facilitation of affective judgements.

The study of aesthetics has revealed that preferences for visual stimuli are often formed on the 
basis of a number of low-level features. Some examples of such features include: stimulus 
complexity (Berlyne, 1970), symmetry (Jacobsen & Höfel, 2002; Makin, Wilton, Pecchinenda, & 
Bertamini, 2012), perceptual fluency (Reber, Schwarz, & Winkielman, 2004), and curvature or 
angularity (Bar & Neta, 2006, 2007). Whilst the majority of the studies in this area have shown 
evidence for these effects using abstract stimuli like groups of lines or coloured patterns, these 
same effects also generalise to preferences about more complex stimuli. For example, 
judgements about the attractiveness of faces, can, to some extent, be attributed to low-level 
features like facial symmetry (Jones et al., 2001; David I Perrett et al., 1999). Bar and Neta 
(2006) argue that these low-level features act as visual primitives, allowing individuals to quickly 
form impressions about visual objects. If this is the case, we should also expect preferences 
relating to these features to remain stable over time.

From the various low-level features could potentially be manipulated, the preference for 
curvature over angularity was considered the most suitable for use in the current study. 
Curvature can be manipulated in novel abstract stimuli in order to negate issues concerning
stimulus familiarity, and there is already evidence to suggest that preferences for more curved abstract shapes should generalise to real world scenarios. For example, studies involving preference judgements have shown that humans will choose curvilinear over rectilinear versions of various stimuli including: architecture (Vartanian et al., 2013) furniture (Dazkir & Read, 2012) or car interiors (Leder & Carbon, 2005). In these studies curvature has been used as a purely categorical variable where a set of curved stimuli are compared to a set of angular stimuli, but curvature can also be operationalised as an interval scaled variable. In one recent example, Palumbo & Bertamini (2016) studied the preference for curvature using a series of polygons whose contours were manipulated by varying properties like the number of convexities, concavities or vertices.

With this in mind, experiment 5 used the preference for curvature to address the questions posed in the previous chapter, namely, can stimuli outside of faces produce affective assimilation effects similar to those discussed in chapter 3, and if so, do those effects persist over time? In addition to experiment 5, this chapter also contains details of a second experiment (experiment 6) which was conducted to further investigate the persistence of assimilation effects over time.

As chapter 5 discussed, assimilation effects have been investigated by a number of studies aimed at promoting healthier lifestyle choices. These choices may relate to food intake, for example, where there is currently pressure on the manufacturers of food products to reduce concentrations of ingredients like sugar, salt and fat. However, unhealthy or addictive ingredients are often perceived to enhance the rewarding qualities or liking of various products, and highlighting the reduction of these ingredients can reduce consumer liking irrespective of the true quality of the product. For example, numerous studies have shown that when foods are labelled as reduced fat, or fat-free, consumers expect them to taste poorer than full fat alternatives (Kähkönen, Hakanpää, & Tuorilla, 1999; Norton et al., 2013; Tuorila et al., 1994).

Labels influence how products are perceived, encouraging consumers to place items which share common labels into categories. When products exist along a continuum, where differences between product variants may be difficult to perceive (e.g. reduced fat, to full fat), labels suggest a source of expert knowledge, offering meaning beyond that which may be gleaned from the product itself (Foroni & Rothbart, 2011). Once a consumer has established that a product belongs within a particular category, future judgements of that product are influenced by prior knowledge based upon that category membership. The influence has been termed a ‘central tendency bias’ (Hollingworth, 1910) where categorised items are viewed as
being closer to the centre of their associated category than they really are, and pairs of stimuli from the same category are viewed as more similar than equidistant pairs of items drawn from opposing categories (Huttenlocher, Hedges, Lourenco, Crawford, & Corrigan, 2007; Olkkonen, McCarthy, & Allred, 2014).

Once a consumer has developed an estimate about the properties of a product (e.g. how sweet a food or beverage should taste), labels and other external cues also serve to smooth out differences between separate consumption experiences. According to the Bayesian model proposed by Huttenlocher, Hedges, and Vevea (2000), people use prior information (like previous experience or the information provided by a label) in order to make estimations about noisy sensory information. Prior information about a product and its associated category (e.g. reduced sugar, no sugar etc.), is combined with the inexact sensory information resulting from a consumption experience to provide the current most accurate estimate of how a product should be perceived. Recently, a number of studies have investigated how prior expectations mask changes to stimuli over repeated consumption experiences, to promote a sense of perceptual constancy. Woods, Poliakoff, Lloyd, Dijksterhuis, and Thomas (2010) adjusted the amount of sugar contained in pairs of drinks that appeared to be poured from either the same, or different jugs. When pairs of drinks appeared to be poured from the same jug, participants reported that drink pairs tasted more similar than when they appeared to be poured from different jugs. Furthermore, the assumption that drink pairs apparently poured from the same jug should taste consistent, masked small changes in sugar concentration between drink pairs. As the concentrations of sugar between the drink pairs became more discrepant, this assimilation effect disappeared, and the experimenters observed that these effects also diminished over the course of the experimental session. Similar findings have also been observed in other studies using food items which have adopted similar methodologies. Le Berre, Boucon, Knoop, and Dijksterhuis (2013) reported that concentrations of a bitter compound could be masked when concentrations were manipulated across multiple bites of an ice cream product, and Dijksterhuis, Boucon, and Le Berre, (2014) were able to mask changes in salt levels when they occurred across multiple bites of the same sandwich.

A question that arises from these studies is whether these assimilation effects can be maintained over time. Woods et al. (2010) reported that assimilation effects diminished over the course of their experiment, and this is understandable given that the cues (whether a drink appeared to be poured from the same or different jugs) became less reliable as time went on. A perceptual estimate in a given trial represents an individual’s current best attempt to model the
scenario at hand, and that model is being continually updated (Dunovan et al., 2014). If the predictability of an expected outcome (that two drinks will be the same if they appear to be poured from the same container) is reduced over time, then the extent to which prior expectations determine perception, should also reduce over time (Huttenlocher et al., 2007). Considering our findings up to this point, we would also expect affective assimilation effects to reduce over time, as predictive information (in the case of experiment 5, cues indicating whether a pattern will be smooth or sharp) becomes less reliable.

Another question of interest for the present research relates to whether the persistence of assimilation effects would be influenced, if the reliability of predictive information was gradually reduced without participants’ knowledge. When stimuli vary along a single dimension (e.g. width or height), Huttenlocher et al. (2007) suggest that categories of stimuli are internally represented by a prototype, essentially the best instance of that category. For example, this could be your internal representation of the ‘ideal’ concentration of an ingredient within a particular product. Prototypes are surrounded by boundaries of stimulus variation functionally similar to the latitude of acceptance in ACT (Hovland et al., 1957). As long as stimuli fall within those boundaries, they will be accepted as members of the associated category. Thus, experiment 6 was designed to test whether two prototypes which exist along the same continuum (smooth and sharp) can be shifted closer together. If so, would this shift influence assimilation effects if participants assumed that the affective gap between the two prototypes remained consistent?

To summarise, this chapter aimed to examine the persistence of affective assimilation effects over time. In experiment 5 liked (smooth patterns) and disliked (sharp patterns) anchor items were associated with predictive cues, and responses were collected to test whether ratings for the anchor items generalised to intermediate items (moderately smooth patterns) paired with the same cues. A later session tested whether any generalisations also persisted over time, by presenting the same predictive cues and intermediate items again, but without the anchor items.

In experiment 6 the same smooth and sharp patterns selected for experiment 5 were used, based upon the assumption that the affective quality of the patterns was determined by how smooth or sharp they were. Participants were recruited in cued and non-cued groups, and over the course of two sessions, both groups were presented with smooth and sharp patterns that became increasingly similar over subsequent blocks. For the cued group, different predictive cues were presented before the smoother and sharper patterns, indicating whether a pattern
belonged in a smooth or sharp category. Since these same category expectations should have been absent in the non-cued group, their responses should have been based solely upon their perception of the stimuli in each trial. It was predicted that, in session 1, compared to participants in the non-cued group, cued participants would accentuate the difference between smoother and sharper patterns and their affective responses would assimilate toward the values suggested by the cues. If participants in the cued group shifted their prototypes for the smooth and sharp categories this difference should persist into session 2, if not, then their responses should fall in line with those from the non-cued group.

Experiment 5

6.2 Methods

Note: Prior to experiment 5, two preliminary studies were conducted. The first (Appendix D.1) was conducted to confirm the preference for curvature, and the second (Appendix D.2) was conducted to determine whether participants could detect relatively small differences in curvature between smooth and sharp patterns.

6.2.1 Participants

Twenty-five participants aged between 18 and 24 (22 females, mean age = 20, SD = 1.89) were recruited for the study from the undergraduate population at the University of Manchester. Participants were required to attend 2 sessions separated by one week. Three participants failed to attend the second session, leaving a remaining sample of 22 participants (20 females, mean age = 20.09, SD = 1.97)

All participants reported normal or corrected to normal vision, and had no reported history of neurological or psychiatric disorders. Participants were paid £10 for their time, plus an additional £5 as part of a conditioning procedure, and all gave written informed consent to participate. All testing procedures were approved by the University of Manchester ethics committee.

6.2.2 Stimuli and apparatus

For the experimental trials, forty greyscale patterns (see Appendix D.3), taken from the sharp pattern set used by Bar and Neta (2006) were selected for use in the experiment. An additional 4 patterns were selected for use in the practice trials. A smoothed version of each pattern was created using the smoothing function in Paintshop Pro Photo X2 (Corel Corporation, Ottawa, Ontario, Canada), as well as 5 additional intermediate patterns of varying smoothness. If the sharp version of each pattern (sharp pattern) is thought of as being 0% smooth, and the smooth
version (smooth pattern) is thought of as being 100% smooth, the intermediate patterns correspond to levels of 16.66%, 33.33%, 50%, 66.66% and 83.33% smoothness (see Figure 6.1).

![Figure 6.1](image.png)

**Figure 6.1** Examples of the sharp, smooth and intermediate versions of one pattern used in experiment 5.

All images were exported as PNG files at a size of 256 by 256 pixels, resulting in a total experimental stimulus set comprised of 40 smooth patterns, 40 sharp patterns, and 200 intermediate patterns. Stimuli were presented using a desktop computer running PsychoPy software (Peirce, 2007) and a 21-inch LCD monitor at a resolution of 1920 by 1080 pixels. All responses were recorded using a standard keyboard and a wired mouse placed on a desk directly in front of the participant.

### 6.2.3 Design and procedure

Participants attended 2 separate sessions. In session 1, participants completed a conditioning task, a rating task, and a version of the Implicit Association Task (IAT) (Greenwald et al., 1998). In session 2, conducted at the same time exactly one week later, participants completed the rating task followed by the IAT. In each session participants sat in a dimly lit room in front of a computer, and all tasks were completed consecutively.

**Session 1**

**Conditioning Task**

A conditioning task was included at the start of the experiment for all participants. The rationale behind the inclusion of this task was to encourage participants to associate the abstract smoother shapes with the delivery of a delayed secondary reward (money). In previous chapters, the faces which were used as target stimuli represented a class of stimuli which have obvious emotional associations, and which automatically draw the attention of the observer (Öhman, 2002). In the present experiment, based upon the existing curvature literature, participants were expected to prefer the smoother abstract patterns. However, because of the
abstract nature of the patterns, the extent of the preference for curvature was expected to be diminished compared to other studies where real objects like car interiors (Leder & Carbon, 2005) have been used. To combat this, the conditioning task was included for all participants as a means of associating the smoother patterns with financial gains, and the sharper patterns with losses.

The conditioning task used only the 40 smooth, and 40 sharp versions of each pattern. Participants were informed that they would be completing a gambling task where they had the opportunity to win up to a further £5 on top of what they were being paid for participating in the experiment (Regardless of their performance in this task all participants were given the additional £5 during the debrief at the end of the experiment). Participants were told that they would be viewing a series of images of greyscale patterns, and that certain patterns would cause them to win money, and certain patterns would cause them to lose money. It was up to the participant to determine which patterns caused them to win or lose money, and to only choose those patterns during the task.

To maintain the premise that this was a gambling experiment, each trial began with an image of a fruit machine (see Figure 6.2). Upon seeing the fruit machine, participants had 3000ms to ‘activate’ the machine by pressing a key. If a participant failed to press the key within the allotted time, a screen appeared notifying them that they had incurred a financial penalty. After ‘activating’ the fruit machine, one of the patterns was selected and presented on-screen. At this point participants had 3000ms to decide whether to accept the pattern by pressing a key, or reject the pattern by withholding their response. If a pattern was accepted, participants would be notified about whether they had won or lost money. If the pattern was rejected, the experiment moved on to the next trial but participants were not informed whether the previously displayed pattern would have won or lost them money. When a smooth pattern was accepted, participants won 90% of the time and lost 10% of the time. The opposite was true of the sharp patterns. The experiment continued until all of the smooth and sharp versions each of the patterns had been displayed; a total of 80 trials. At the end of the task participants were informed that their total winnings would be calculated, and that these winnings would be paid during the debriefing at the end of the second session.
Rating Task

After completing the conditioning task, participants were informed that the next section of the experiment would involve presentations of the same set of patterns that they had seen in the previous task, and that they would be asked to state whether they liked or disliked each pattern based upon their initial ‘gut reaction’. Responses were limited to a like/dislike forced-choice decision, and responses were collected using 2 keys at opposite ends of the keyboard.

The task included presentations of the smooth and sharp versions of each pattern, as well as all of the intermediate versions. Two sets of predictive cues were also used in the task, and for each participant, patterns were presented in cued and non-cued conditions. In cued trials predictive cues (blue or yellow borders) were assigned to the smooth and sharp versions of each pattern to act as anchors. These pairings remained consistent throughout the task, all smooth patterns were presented with one colour of cue and all smooth patterns were presented with a cue of a different colour. Additionally, the intermediate versions of each pattern were presented an equal number of times with cue that preceded smooth pattern presentations (smooth cue), and the cue that preceded the sharp pattern presentations (sharp cue). Cue colour pairings for the smooth and sharp anchor items were counterbalanced across participants. Non-cued trials were identical to cued trials in most respects, but the non-cued trials did not include the anchor items.
The task was divided into 5 blocks and consisted of 1,000 trials total. The smooth and sharp versions of each pattern were presented 5 times each (400 trials), and the 5 levels for each pattern were presented once with the smooth cue (200 trials), once with the sharp cue (200 trials), and once with no predictive cue (200 trials). Presentations were ordered in a pseudo-random fashion with earlier presentations being weighed in favour of the anchor items so that participants built cue/pattern type associations. Block 1 contained 80 trials comprised of only the smooth and sharp anchor items. The remaining 4 blocks each contained: 80 smooth and sharp anchor trials, 100 cued intermediate trials, and 50 non-cued intermediate trials.

Each trial began with a fixation cross presented for 100ms. The next screen was presented for 1000ms and in the cued trials a coloured frame appeared where the target would later be presented; in the non-cued trials this screen remained blank. Next, a target pattern appeared for 84ms, and in cued trials the frame remained on screen surrounding the target pattern. Timings were based upon those reported in the original study conducted by Bar and Neta (2006). At the end of each trial a response prompt was presented for up to 3000ms depending on how quickly a participant responded (see Figure 6.3).

Figure 6.3 Example trial sequence from the rating task. This is an example from a cued trial where the cue (yellow border) appears before the target, and remains on-screen with the target.

At the start of the task participants were instructed to try to respond as quickly as possible in all trials after the prompt appeared, once a response was provided, the task moved on to the next
trial. If no response was provided within 3000ms, a ‘too slow’ prompt was presented for 100ms, and the task automatically moved on to the next trial. Any trials where a participant failed to provide a response were noted in the output file generated at the end of the task. Before beginning the task proper, participants completed 8 practice trials to familiarise themselves with the task. The practice trials involved cued smooth and sharp presentations of 4 additional patterns which were not used during the experimental trials. The format of the practice trials was identical to that of the experimental trials described earlier.

IAT

The IAT task followed the standard IAT design outlined by Greenwald, McGhee, and Schwartz (1998). Throughout the task attribute category labels were presented in the top-left, and top-right corners of the screen, and a series of words (good or bad), images (smooth or sharp objects), or both words and images were appeared in the middle of the screen. Participants were instructed to categorise the items that appeared centrally using a key press (‘e’ as the left response key, or ‘i’ as the right response key) according to the category labels presented at the top left and right of the screen in each block. The items that appeared centrally remained on-screen until a response was provided. If an item was sorted into the incorrect category a small red x appeared below the item indicating that the participant should alter their response. Participants were instructed to try to sort the items as quickly as possible, whilst attempting to make as few mistakes as possible. The IAT task was divided into 7 blocks (see table 6.1), and participants had the opportunity to take breaks between blocks.
### Table 6.1  
*Design of IAT task*

| Block | Trials | Stimuli assigned to left response key  
<sup>a</sup> | Stimuli assigned to right response key  
<sup>a</sup> |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40</td>
<td>Smooth Objects&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Sharp Objects&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>Good Words</td>
<td>Bad Words</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>Smooth + Good</td>
<td>Sharp + Bad</td>
</tr>
<tr>
<td>4</td>
<td>80</td>
<td>Smooth + Good</td>
<td>Sharp + Bad</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>Sharp Objects</td>
<td>Smooth Objects</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>Sharp + Good</td>
<td>Smooth + Bad</td>
</tr>
<tr>
<td>7</td>
<td>80</td>
<td>Sharp + Good</td>
<td>Smooth + Bad</td>
</tr>
</tbody>
</table>

**Note:**  
<sup>a</sup> The assignment of stimuli to the left and right response keys was counterbalanced across participants.  
<sup>b</sup> In the Sharp + Good to Smooth + Good sequence, blocks 1, 3 and 4, were swapped with blocks 5, 6, and 7.  
<sup>c</sup> 20 patterns were shown in their smooth forms, and the remaining 20 patterns were shown in their sharp forms. This was also counterbalanced across participants.

### Session 2

#### Rating task

In most respects, the rating task in session 2 was identical to the rating task from session 1, but it did not include the smooth and sharp anchor items. Consequently, the task consisted of only 600 trials, the 5 intermediate levels for each pattern were presented: once with smooth associated predictive cue (200 trials), once with the sharp associated predictive cue (200 trials), and once with no predictive cue (200 trials). Due to the smaller number of trials, this version of the rating task was divided into only 4 blocks, each containing: 50 non-cued intermediate, and 100 cued intermediate trials. In all other respects the task was identical to the one described in session 1.

#### IAT

The version of the IAT conducted at the end of session 2 was identical to that of session 1, and participants completed the same IAT sequence (good + smooth to good + sharp, or good + sharp to good + smooth) in both sessions.
6.2.4 Data analysis

For the 5 intermediate levels of each of the 40 patterns, participants made a like/dislike judgement when the pattern was preceded by the smooth associated cue, the sharp associated cue, and when no cue was presented beforehand. This was repeated for session 1 and session 2, and for each cue condition within each session, three logistic psychometric functions were fitted to the ratio of participants’ like/dislike responses using the Palamedes toolbox for Matlab (Prins & Kingdom, 2009).

From each function PSE for each cue condition was calculated. In the current experiment the PSE represents the point on the x-axis (intermediate pattern smoothness) at which each type of response on the y-axis (‘liked’/‘disliked’) is given 50% of the time. For further details see general methods.

To determine whether the presence of the cues influenced participants’ affective ratings of the intermediate patterns, PSE data were subjected to a 3 (cue type: smooth cue, sharp cue, no cue) x 2 (session: session 1, session 2) repeated-measures ANOVA. Liking ratings for the smooth and sharp patterns were compared using paired samples t-tests.

Before conducting the reaction time analysis, data from no response trials were removed. Reaction time data for the intermediate presentations were analysed using a second 3 (cue type: smooth cue, sharp cue, no cue) x 2 (session: session 1, session 2) repeated measures ANOVA. Reaction times for the smooth and sharp presentations were analysed using paired samples t-tests.

IAT scores were calculated using the scoring algorithm described in Greenwald, Nosek, and Banaji, (2003), for further details see general methods. IAT scores for each participant across session 1 and session 2 were compared using paired samples t-tests.

In cases where Mauchly’s test revealed that the assumption of sphericity had been violated, the degrees of freedom have been corrected using Greenhouse-Geisser estimates of sphericity. An alpha level of .05 has been used for all statistical tests, and effect sizes are reported here as partial eta squared ($\eta^2_p$) where 0.05, 0.1, and 0.2 correspond to small, medium, and large effects, respectively (Cohen, 1988).
6.3 Results

6.3.1 Liking ratings for the smooth, sharp, and intermediate patterns

On average, participants liked the smooth versions of each pattern \((M = 79.27\%, SD = 20.25\%)\) more than the sharp versions \((M = 14.52\%, SD = 14.74\%)\). A paired samples t-test revealed that the smooth versions of the patterns were liked significantly more that the sharp versions, \(t(21) = 10.79, p < .001\). As predicted, liking responses for the intermediate patterns increased in a broadly linear fashion as pattern smoothness was increased (see Figure 6.4).

![Figure 6.4](image)

Figure 6.4  Liking ratings for intermediate stimulus levels from experiment 5 collapsed across session and cue type. Error bars represent standard errors of the mean.

6.3.2 Measures of participants’ PSEs

The repeated measures ANOVA conducted upon participants’ PSEs from the intermediate pattern trials, showed a main effect of cue type, \(F(2, 42) = 3.38, p = .043, \eta^2_p = .14\). Pairwise comparisons revealed that PSEs were significantly smaller for trials where the predictive cue associated with the smooth patterns \((M = .50, SD = .61)\) was used, rather than the cue associated with the sharp patterns \((M = .63, SD = .54)\), \(p = .034\). There was no significant difference between trials where the smooth or sharp associated predictive cues were used, and the trials where no cue was used \((M = .59, SD = .64); smooth cue: p = .110; sharp cue: p = .358\).

There was no main effect of session, indicating that PSEs did not differ significantly between session 1 \((M = .48, SD = .33)\), and session 2 \((M = .67, SD = .76)\), \(F(1, 21) = 2.74, p = .113\). The session x cue interaction was also not significant, \(F(2, 42) = 2.82, p = .071\).
6.3.3 Reaction time data

The reaction time analysis revealed a significant main effect of session where participants’ responses to the intermediate patterns were significantly faster in session 2 ($M = 283.68ms$, $SD = 84.71ms$) compared to session 1 ($M = 235.62ms$, $SD = 74.30ms$), $F(1, 21) = 19.57$, $p < .001$, $\eta^2_p = .48$. The analysis also revealed a significant main effect of cue, $F(1, 27) = 12.83$, $p = .001$, $\eta^2_p = .38$. Pairwise comparisons showed that responses in the no cue condition ($M = 270.75ms$, $SD = 87.65ms$) were significantly slower than both the smooth cue ($M = 257.10ms$, $SD = 81.93ms$, $p = .001$) and sharp cue ($M = 251.09ms$, $SD = 79.75ms$, $p = .004$) conditions. The cue x session interaction was not significant, $F(2, 42) = 2.50$, $p = .094$.

Reaction time analysis for the smooth ($M = 303.86ms$, $SD = 109.81ms$) and sharp ($M = 293.43$, $SD = 88.80ms$) patterns used in session 1 revealed no significant differences, $t(21) = 1.54$, $p = .137$.

6.3.4 IAT data

All participants showed a preference for the smooth patterns over the sharp patterns, and paired samples t-tests revealed that the preference for the smooth patterns over the sharp
patterns did not change between session 1 \( (M = -0.72, SD = 0.40) \) and session 2 \( (M = -0.68, SD = 0.37) \), \( t(21) = -0.46, p = 0.648 \).

### 6.4 Interim discussion

Across the 2 sessions, participants viewed intermediate patterns in three different cue conditions: positively cued, negatively cued, and without cues. In session 1, when different predictive cues were presented before the same moderately liked items, participants adjusted their affective ratings in line with the values suggested by the cues. This result was expected and replicated the results of experiment 2. During session 1, the inclusion of the anchor items meant that if participants were basing their responses upon the prior probability of certain outcomes, more positive (smoother patterns), and more negative (sharper patterns) outcomes could be expected with each type of cue.

In session 2 these affective assimilation effects had been expected to disappear. The anchor items were not included with the cued presentations during session 2, so there was no difference between the stimuli paired with each type of cue. Previous studies have suggested that participants' estimates about stimulus categories should slowly update over time to incorporate instances where an outcome has not matched the expectation. If this was the case here, the difference in PSE between the two sets of cued presentations should have disappeared in the second session, but this did not happen. In many previous perceptual studies (Dijksterhuis et al., 2014; Le Berrre et al., 2013; Woods et al., 2010), researchers have included trials which represent major deviations from category expectations (e.g. experiencing large changes in salt concentration over multiple bites of a single sandwich). In the case of the current study, since the cued presentations did not involve discrepancies at the extreme ends of the smooth/sharp continuum, participants may not have updated their expectations. This is possible, if the range of intermediate stimuli that were used all remained within the latitude of acceptance surrounding participants' existing internal representations of the smooth and sharp categories.

### Experiment 6

#### 6.5 Methods

##### 6.5.1 Participants

A total of 63 individuals aged between 18 and 31 (50 Females, mean age = 21.37, SD = 3.17) took part in the study. All participants were recruited from the staff and student population at the University of Manchester. Participants were required to attend two sessions separated by
one week, and written informed consent was obtained prior to the first session. Participants were assigned to one of two groups a cued group, and a non-cued group.

Before beginning the analysis the results from 7 participants were excluded: 2 because they failed to respond correctly in ≥ 75% of the catch trials, 3 because they failed to attend both sessions, and 2 because their IAT scores indicated a strong preference for the sharper patterns. After these participants were excluded, both the cued group (mean age = 21.11, SD = 2.71, male to female ratio 8:20), and a non-cued group (mean age = 22.25, SD = 3.63, male to female ratio 6:22) contained 28 participants.

All participants reported normal or corrected to normal vision, and had no reported history of neurological or psychiatric disorders. Participants were paid £10 for their time, plus an additional £5 as part of a conditioning procedure, and all gave written informed consent to participate. All testing procedures were approved by the University of Manchester ethics committee.

6.5.2 Stimuli and apparatus

The experiment used the same base set of 40 smooth and sharp greyscale patterns described in experiment 5. However, for this experiment, 9 intermediate versions were created between the smooth and sharp versions of each pattern (see Figure 6.6).

Figure 6.6 Sample set of stimuli from experiment 6. Sharp and smooth stimuli were only seen during the conditioning and IAT procedures. Intermediate stimuli were shown throughout the rating task in sessions 1 and 2. For simplicity all intermediate stimuli are referred to in terms of their smoothness.
Image properties and dimensions were the same as those described in experiment 5. The task used a total of 40 smooth patterns, 40 sharp patterns, and 360 intermediate patterns. A greyscale Gaussian noise mask was used for the non-cued participants, and the mask matched the size of the target stimuli (256 by 256 pixels). Coloured noise masks were for the cued participants, which were created by overlaying semi-transparent pink (RGB value: 255, 0, 255) and green (RGB value: 0, 255, 0) filters onto the noise mask used in the non-cued trials. Stimuli were presented using the same computers and software described in experiment 5.

6.5.3 Design and procedure

The design of the experiment was similar to experiment 5. Participants attended 2 sessions separated by a week, and the experiment involved a conditioning task, a rating task, and a version of the IAT. Task order in session 1 was slightly different from experiment 5; participants completed the conditioning task followed by a rating task, followed by the IAT. In session 2, participants completed a rating task followed by the IAT. Participants in the cued group also completed an additional task at the end of session 2 to assess whether they had acknowledged the relationship between the cues and the stimuli. Like experiment 5, all tasks were completed consecutively.

Session 1

Conditioning Task

The conditioning task was identical to the one described in experiment 5.

IAT

The IAT was also identical to the one described in experiment 5.

Rating Task

At the start of the task, participants were told that they would be viewing the same stimuli that they had seen in the previous tasks, and that they would be asked to provide a rating for each pattern in terms of a like/dislike forced-choice decision.

This target images for this rating task were a subset of the intermediate versions of each pattern, and the smooth and sharp versions of each pattern were not used at any time during the task. The intermediate versions of each pattern were divided into ‘smoother’ (70%, 80%, and 90% smooth versions of each pattern) and ‘sharper’ (10%, 20% and 30% smooth versions of each pattern) categories. The task was divided into 4 blocks, in the first block the 90% and 10% levels were used as the smoother and sharper stimuli respectively, and over the course of the
task the smoother patterns became less smooth, and the sharper patterns became less sharp (see table 6.2).

<table>
<thead>
<tr>
<th>Session</th>
<th>Block</th>
<th>Smoother trials*</th>
<th>Sharper trials*</th>
<th>Catch trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>90%</td>
<td>10%</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>90%, 80%</td>
<td>10%, 20%</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>80%, 70%</td>
<td>20%, 30%</td>
<td>10</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>70%</td>
<td>30%</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>70%</td>
<td>30%</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>70%, 60%</td>
<td>30%, 40%</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>60%, 50%</td>
<td>40%, 50%</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>50%</td>
<td>50%</td>
<td>5</td>
</tr>
</tbody>
</table>

Note: *In each block the 40 patterns used in the experiment were presented once for each intermediate level present within the smoother and sharper trials. This resulted in 85 trials in blocks 1 and 4 (40 smoother, 40 sharper, and 5 catch trials), and 170 trials in blocks 2 and 3 (80 smoother, 80 sharper, and 10 catch trials).

The task was exactly the same for the cued and non-cued participants except for the colour of the noise masks used. For the cued participants, one cue colour of noise mask (green or pink) was consistently paired with sharper patterns, and the noise mask of the opposite colour was paired with smoother patterns (see Figure 6.7). Contour (smoother/sharper) and cue colour (pink/green) pairings were counterbalanced across cued participants. For the non-cued participants, the same greyscale mask was presented before all trials regardless of the contour of the presented pattern.
Figure 6.7  Example of stimulus increments and cues used in the cued condition.

Each trial began with a fixation cross presented for 100ms, followed by a noise mask for 1000ms, the target image for 84ms, a second noise mask for 1000ms, and finally a response prompt (Figure 6.8). Participants were instructed to try to respond as quickly as possible after the prompt appeared, and once a response had been provided the task moved on to the next trial. If no response was provided within 3000ms, the task automatically moved on to the next trial. Any trials where a participant failed to provide a response were noted in the output file generated at the end of the task. Approximately 6% of the total trials were ‘catch trials’ where a blank background was presented where the target image should have been displayed. In these trials participants were instructed to acknowledge the absence of the target stimulus by pressing the spacebar, rather than providing a like/dislike response.
Figure 6.8  Example trial sequence from the rating phase of experiment 6 in the non-cued condition.

The task consisted of 510 trials total. Blocks 1 and 4 contained 85 trials: each of the 40 patterns were presented for one level of smoothness and one level sharpness, and 5 catch trials were included. Blocks 2 and 3 contained 170 trials: each of the 40 patterns was shown for 2 levels of smoothness and 2 levels of sharpness, and 10 catch trials were included.

Session 2

Rating Task

The rating task used in session 2 was virtually identical to the task that was used in session 1. However, the smoother and sharper versions of each pattern used in block 1 did not revert back to the 90% and 10% levels used at the start of session 1. Instead, the first block began with the smoother / sharper levels employed at the end of session 1 (70% and 30% respectively). From here the smoother and sharper stimuli became increasingly similar over the course of the remaining blocks, until the two became identical in the final block (see table 6.2). Aside from this, all other aspects of the task were identical to task described for session 1.

IAT

The version of the IAT conducted at the end of session 2 was identical to that of session 1, and participants completed the same IAT sequence (good + smooth to good + sharp, or good + sharp to good + smooth) in both sessions.
Manipulation Check

An additional task was included at the end of session 2 for participants in the cued group in order to assess whether they had learned the cue / stimulus associations. In each trial a fixation cross was presented centrally for 100ms, followed by one of the coloured noise masks used during the rating task, presented for 1000ms. After the noise mask disappeared, the smooth and sharp versions of one of the patterns were presented on the left and right hand sides of the screen. At this point the participant had to use the left and right arrow keys to indicate, given the noise mask that they had just seen, which of version of the pattern (smooth or sharp) they would expect to appear next. The smooth and sharp versions of the pattern remained on-screen until the participant had made a decision.

In total the task consisted of 10 trials, and for each participant, the 10 patterns used in the task were randomly selected from the overall set of 40 patterns. Five of the presentations were preceded by one of the coloured noise masks, with the remaining 5 presentations being preceded by the other coloured noise mask. The order of presentations and the horizontal positions of the smooth/sharp patterns were randomised throughout.

6.5.4 Data analysis

For each participant, a logistic psychometric function was fitted to the ratio of like/dislike responses for the smoother (90% 80% and 70%) and sharper (10%, 20% and 30%) patterns presented in session 1. An additional function was fitted to each participant’s like/dislike responses from session 2 for the smoother (70%, 60%, and 50%) and sharper (30%, 40%, and 50%) patterns. In order to fit the functions, each level of the intermediate stimuli along the x-axis needed to have a unique value, so the two 50% presentations were coded as different values. Therefore, the sharper cue 50% presentations were plotted as 49% and the smoother cue 50% presentation were plotted as 51%. The slope of the function (see general methods for details) fitted to each participant’s responses was then calculated for session 1 and session 2 using the Palamedes toolbox for Matlab (Prins & Kingdom, 2009).

To determine whether the presence of cues caused participants to differentiate more between pairs of smoother and sharper patterns, slope data were subjected to a mixed design ANOVA. Session (session 1, session 2) was used as a within subjects factor, and group (cued, non-cued) was the between subjects factor.
Before conducting the reaction time analysis, data from no response trials and catch trials were removed. Reaction time data collected during the rating tasks from both sessions were analysed using a second 2 (session: session 1, session 2) x 2 (group: cued, non-cued) mixed design ANOVA.

IAT scores were calculated using the scoring algorithm described in Greenwald, Nosek, and Banaji, (2003), for further details see general methods. IAT scores for each participant across session 1 and session 2 were compared using paired samples t-tests. IAT scores between cued and non-cued participants were compared for session 1, and session 2, using independent t-tests.

In cases where Mauchly’s test revealed that the assumption of sphericity had been violated, the degrees of freedom have been corrected using Greenhouse-Geisser estimates of sphericity. An alpha level of .05 has been used for all statistical tests, and effect sizes are reported here as partial eta squared ($\eta_p^2$) where 0.05, 0.1, and 0.2 correspond to small, medium, and large effects, respectively (Cohen, 1988).

6.6 Results

Note: 6 participants were removed from the analysis because we could not fit a suitable function to their data (see Appendix E.5). This meant that the smooth and sharp groups in the analysis each contained 25 participants.

6.6.1 Liking ratings for the intermediate patterns

Liking ratings for the intermediate patterns increased in a broadly linear fashion as the smoothness of the patterns was adjusted (see Figure 6.9)
Figure 6.9  
Liking ratings for intermediate stimulus levels from experiment 6 collapsed across session (1 and 2) and group (cued and non-cued). Error bars represent standard errors of the mean.

Paired samples t-tests revealed that, for each block in session 1, both the cued and non-cued groups liked the smoother intermediate pattern variants (90%, 80% and 70% smooth) significantly more than the sharper variants (10%, 20%, and 30%) smooth (all \( p < .001 \)). In session 2, participants in both groups liked the smoother intermediate patterns (70% and 60% smooth) more than the sharper patterns (30% and 40% smooth) for the first 2 blocks (all \( p < .001 \)), but only participants in the cued group differentiated between the presentations of the 50% patterns, liking the patterns paired with the smooth associated cue significantly more than those paired with the sharp associated cue (cued group: \( p = .032 \), non-cued group: \( p = .638 \)).

### 6.6.2 Measures of slope

The mixed ANOVA conducted upon participants’ slope data revealed a significant main effect of group, \( F(1, 48) = 5.25, \ p = .026, \eta^2_p = .10 \). Overall, the slopes of functions produced by the cued participants (\( M = 5.87, \ SD = 6.73 \)) were significantly steeper than the slopes of functions from the non-cued participants (\( M = 2.97, \ SD = 2.72 \)). No main effect of session was observed indicating that the slopes of participants’ functions were similar for session 1 (\( M = 4.58, \ SD = 4.03 \)) and session 2 (\( M = 4.25, \ SD = 6.38 \)), \( F(1, 48) = .20, \ p = .658 \), and the session x group interaction was also not significant, \( F(1, 48) = .24, \ p = .630 \) (see Figure 6.10).
Figure 6.10  Group-level functions for cued and non-cued participants presented across session 1 and session 2 from experiment 6. Functions are based upon the average number of liked responses from each group in sessions 1 and 2 at each intermediate level of pattern smoothness. Group level functions are presented for illustrative purposes; analysis was performed upon slope data from individual participant functions.

6.6.3 Reaction time data
The analysis of the reaction time data revealed that participants’ responses were significantly quicker in session 2 ($M = 267.40\text{ms}$, $SD = 85.59\text{ms}$) compared to session 1 ($M = 325.20\text{ms}$, $SD = 94.20\text{ms}$) $F(1, 48) = 46.05$, $p < .001$, $\eta_\rho^2 = .10$. No difference in overall reaction time was observed between the cued participants ($M = 293.60\text{ms}$, $SD = 100.77\text{ms}$), and the non-cued ($M = 299.00\text{ms}$, $SD = 87.95\text{ms}$) participants, $F(1, 48) = .031$, $p = .860$, and the session x group interaction was also not significant $F(1, 48) = .39$, $p = .536$.

6.6.4 IAT data
Independent samples t-tests conducted upon participants’ IAT scores from session 1, $t(48) = 1.72$, $p = .094$, and session 2, $t(48) = 1.98$, $p = .056$, revealed that there was no difference between the groups in terms of the extent to which they preferred the smoother patterns over the sharper patterns in session 1 or 2 (all $p > 0.025$ Bonferroni correction). Paired samples t-tests also revealed no difference in IAT scores when comparing session 1 to session 2 for cued participants, $t(24) = .49$, $p = .633$, or non-cued participants $t(24) = .61$, $p = .552$. 

133
6.6.5 Manipulation check and catch trial data

Of the 25 participants in the cued group, 10 participants correctly identified all of the trials from the final task. All participants in the cued group correctly identified at least 6 of the trials, \(M = 8.64, SD = 1.50\).

Across both groups participants responded correctly to 97.40% (SD = 5.12%) of the catch trials. Independent t-test revealed that the two groups did not differ in the number of correct responses during catch trials in either session 1 (cued group: \(M = 98.40\%\), \(SD = 7.33\%\); non-cued group: \(M = 96.00\%\), \(SD = 2.18\%\)) or session 2 (cued group: \(M = 97.07\%\), \(SD = 5.64\%\); non-cued group: \(M = 97.87\%\), \(SD = 3.95\%\)) (All \(p > .05\)).

6.7 Interim discussion

The aim of experiment 6 had been to investigate whether affective assimilation effects relating to 2 stimulus categories (smooth and sharp) would persist, if the difference between the two categories was gradually reduced over time. Liking response data revealed that participants from both the cued and non-cued groups affectively discriminated between the pairs of smoother and sharper patterns within each block, but the steeper slopes observed in functions plotted for the cued participants; indicate that affective responses from this group were more sensitive to smaller differences in pattern smoothness.

Throughout session 1, it remained relatively easy to perceive the differences between the smoother and sharper patterns. Because of this, the cues had been expected to mask the subtle changes to the smooth and sharp patterns as they became more similar, and this seems to have been the case, as evidenced by the steeper slopes for the cued participants. In session 2, it appears that participants in the cued group shifted their affective prototypes for the smooth (good) and sharp (bad) stimuli. This is suggested by the steeper slopes produced by these participants in the second session, compared to the non-cued participants. Cued participants continued to affectively differentiate between the smoother and sharper patterns to a greater extent throughout the entire experiment (session 1 and session 2). This effect persisted as the perceptual differences between the smoother and sharper patterns in each block became harder to detect, even to the point where the two sets of patterns became perceptually indistinguishable from one another.
6.8 General discussion

The 2 experiments discussed within this chapter aimed to investigate the persistence of affective assimilation effects, and to do so whilst attempting to control for participants’ prior experience outside of the experimental sessions. In both experiments, affective assimilation effects brought about by predictive cues were observed, and these effects persisted after a period of one week. In experiment 5, participants continued to use predictive cues to judge the affective value of noisy sensory stimuli (moderately smooth patterns) even as the reliability of the cues, in terms of accurately forecasting affective outcomes, diminished over time. In experiment 6, the presence of predictive cues caused participants to accentuate the difference between 2 sets of patterns which existed along the same continuum. This enhanced discrimination persisted into the second session even as the two sets of patterns became more homogenous over time.

The linear increase in liking observed across the intermediate patterns for both experiments, as well as the results of preliminary study 6a (Appendix E.1), indicate that stimulus liking increased as patterns became more curved, replicating the preference for curvature over angularity observed in other studies (Bar & Neta, 2006; Palumbo & Bertamini, 2016). Whether this increase was driven by a liking of curvature or a disliking of angularity, is debatable (Bertamini, Palumbo, Gheorghes, & Galatsidas, 2016), but for the purposes of the experiments presented here, it was sufficient to know that increases in curvature corresponded to increases in stimulus liking.

The results from experiment 5, session 1, replicated the findings observed in experiment 2. PSEs for presentations of intermediate stimuli paired with the smooth pattern associated cue were significantly lower than for presentations of the same stimuli, paired with the sharp pattern associated cue. Participants were not explicitly told about the stimulus/cue associations, but anchor items were included as part of the cued presentations, increasing the probability of highly positive (smooth) and highly negative (sharp) outcomes when the different cues were present. The ratings for the anchor items show that participants displayed a strong preference for the smooth over the sharp anchor items. Since participants presumably had no prior information from outside of the experimental session to influence their affective expectations, the most plausible explanation for the difference in PSEs for the intermediate items, is that participants displayed a central tendency bias assimilating the values of the intermediate stimuli towards the category suggested by their associated cues. This bias resembles similar biases reported elsewhere within the anchoring literature (Furnham & Boo, 2011; Tversky &
Kahneman, 1974). Here, participants appeared to acknowledge that the intermediate patterns were different from the initially encountered smooth and sharp patterns, but they did not adjust their judgements of the intermediate sufficiently to account for this difference.

In experiment 5, session 2, PSEs remained significantly lower for intermediate presentations paired with the smooth associated cue, compared to presentations paired with the sharp associated cue. In session 2, participants had been expected to update their expectations relating to the cued presentations because of the removal of the anchor items, and there are two possible explanations for why this did not occur. Firstly, it is possible that assimilation effects persisted into session 2 because the range of intermediate stimulus values was not so discrepant that they caused participants to update their expectations. As previous chapters have discussed, contrast effects occur relatively infrequently, and in terms of processing speed and the allocation of cognitive resources, it is often more efficient to base an affective reaction upon an expected value rather than the stimulus information that is actually available (Geers & Lassiter, 2005; Kirsch, 1999). Secondly, even though the anchor items were removed during session 2, if estimations of stimulus value are updated in a Bayesian manner, and participants considered the cued presentations across both sessions as a whole, the average value of presentations was still higher for one set of cued presentations which could explain the persistence of assimilation effects.

In chapter 5, two studies were discussed which observed that the persistence of expected liking for food products was influenced by participants’ familiarity with those products. After an experience which contradicted existing beliefs about a product, expected liking for that product could be affected in the short-term for familiar and unfamiliar products, but expected liking returned to baseline levels over the longer-term for products where there was a high degree of stimulus familiarity (E. Robinson et al., 2013; Tuorila et al., 1994). The findings of experiment 5 appear to support these observations, since affective expectations about presentations paired with the smooth associated cue remained more positive during the second session without additional reinforcement, and these expectations should only have been influenced by participants’ experience during the experimental sessions.

The reaction time data from experiment 5 supports the assertion that participants were using the cues to guide their affective judgements throughout the two sessions. Recognition is facilitated when sensory information is consistent with prior expectations, and recognition is hindered when sensory information does not match with prior expectations (Król & El-Deredy, 2011a, 2011b; Puri & Wojciulik, 2008). There were no differences in reaction times between the
two types of cued presentations, but responses for cued trials were faster than those for non-cued trials across both sessions. In session 1, only 50% of the cued trials featured the smooth and sharp anchor items, and in session 2 the anchor items did not feature at all, but the decreased reaction times which were observed for the cued trials in both sessions, provide further support for the theory that the cue associations masked the variability in the pattern presentations. This theory is also supported by the observation that IAT scores did not change between the end of session 1 and the end of session 2, which indicates that the preference for curved stimuli over sharper ones remained consistent throughout the experiment.

Responses in the second session of experiment 5 were faster overall, compared to the first session, and this may be due to practice effects, but it could also be due to boredom or a lack of attention on the part of participants. If participants learned the affective associations during the first session, they may not have been attending to the stimuli sufficiently in the second session, and this could explain why participants’ liking responses continued to assimilate toward the values suggested by the cues.

In experiment 6, the predictive cues acted as more reliable predictors of stimulus value than the cues used in experiment 5. Based upon the assumption that participants viewed smoother patterns more positively than sharper ones, the experiment was conceived so that, barring the final block, the hedonic value of the intermediate stimuli was always higher when they were paired with the smooth associated cue. The magnitude of the difference in hedonic values between the smoother and sharper stimuli diminished over the course of the experiment, and liking responses from both groups reflected this change (see Appendix E.6). However, the slope data shows that participants in the cued group continued to affectively discriminate between the smoother and sharper stimuli to a greater extent than those in the non-cued group. These findings suggest that affective responses may have been influenced by the cues in one of two ways. Participants’ either shifted their prototypes for their ideal representations of smoother and sharper patterns closer together; or the gradual reduction in smoothness/sharpness, combined with expectations about affective constancy, resulted in the smoother and sharper stimuli remaining within the boundaries of variation surrounding participants’ existing prototypes. At present it is difficult to establish which of these possibilities is correct.

The IAT data from session 1 of experiment 6 indicates that both the cued and non-cued groups showed similar implicit preferences for the smooth patterns over the sharper ones. Therefore, the increased differentiation between the smooth and sharp patterns observed for the cued group cannot be attributed to stronger implicit preferences for the sharp patterns. For
experiment 6, the IAT was deliberately moved before the rating task in session 1, so that participants only saw the most curved and most angular patterns before, and after the rating tasks. The results of the second IAT showed that there was still no difference between the groups at the end of the study, in terms of their preferences for the smoother patterns over the sharper ones. If participants in the cued group had truly adapted their prototypes of smoothness/sharpness, a difference in IAT scores may have been expected when participants saw the most curved/angular patterns in the IAT. Yet the absence of this difference cannot be considered conclusive, since implicit attitudes are rigid and might not realistically be expected to change over the course of a single week (Hermans et al., 2003).

The reaction time data from experiment 6 showed no difference in reaction times between the cued and non-cued participants, but overall reaction times were faster in session 2 than they were in session 1. In the case of this experiment, catch trials were included which meant that participants had to attend to the pattern presentations, and any participants who failed to respond correctly in 75% of the catch trials were not included in the analysis. Therefore, it is likely that the faster responses observed in session 2 of experiment 5 and experiment 6, can be attributed to practice effects rather than a lack of participant attention.

Some studies aimed at promoting reductions in salt consumption have adopted an approach similar to the one used in experiment 6, by decreasing the amount of salt in a product slowly over time. Two studies using this 'stealth' approach have observed that salt concentrations may be gradually reduced without influencing liking (Girgis et al., 2003) or consumption habits (Bolhuis, Temme, Koeman, Noort, & Kremer, 2011). One suggested problem with the stealth approach is that any reduction in salt will become more apparent when comparisons are made with similar products whose salt concentrations have not been reduced (Busch, Yong, & Goh, 2013). Nonetheless, this may not be a problem, if the stealth approach causes participants to adapt values of their ideal remembered salt concentration.

The results of experiments 5 and 6 showed that even when there was no difference in the inherent quality of items from 2 different categories (in terms of patterns’ curvature), participants still affectively differentiated between items from each category based upon the assumption that one category was better than the other. However, a limitation of the studies discussed in the current chapter is that they were only designed to investigate assimilation effects up to the point where two groups of stimuli became identical. Girgis et al. (2003) showed that assumptions of perceptual constancy could mask gradual reductions in salt content up to around 25%, and other food based studies have claimed that assimilation effects will only occur
within a limited window of difference between actual, and expected flavour intensities (Dijksterhuis et al., 2014; Le Berrre et al., 2013; Woods et al., 2010). Similarly the AEM (T. D. Wilson et al., 1989) predicts that assimilation will be replaced by contrast when large discrepancies exist between expectations and stimulus values. The current studies used stimuli where participants did not have strong pre-existing expectations about the hedonic values of the stimuli, and in order to establish the cue/stimulus associations the deviations from expectation that were included were fairly conservative. Future experiments could build upon the current findings by attempting to push the boundaries of the assimilation window further; in order to test whether assimilation effects can still persist when the quality of items from a preferred or liked category, is reduced to the point where they are objectively worse than those from a competing category. Nevertheless, it may prove difficult to create such a manipulation, because the presence of highly discrepant cue/stimulus pairings is likely to promote distrust in the participant, reducing the overall effectiveness of the cues (Mayo, 2015).

6.9 Conclusion
The current study included 2 experiments intended to investigate affective assimilation effects. In both studies, evidence was observed to indicate that predictive cues promoting expectations about affective constancy encouraged participants to discriminate between stimuli associated with each cue. Both experiments showed that affective assimilation effects persisted after a period of one week, even as the differences in hedonic value between groups of cued stimuli became harder to detect. The experiments used novel abstract stimuli where prior knowledge about stimulus quality should not have been a factor, but questions still remain about whether assimilation effects would continue to persist if further stimulus/cue discrepancies were observed.
7 BIASES IN PERCEPTUAL JUDGEMENTS OVER TIME

7.1 Introduction

Predictive coding models of perception indicate that predictions are refined over time to reduce discrepancies between expectations and actual experience, and it is thought that humans adopt a Bayesian inferential approach where estimates and predictions are slowly updated on the basis of newly learned information (Brown & Brüne, 2012). It is, however, currently unclear how quickly these predictions are updated. In experiment 6, cues intended to promote affective expectations, introduced a central tendency bias; where affective judgements from cued participants toward groups of positive and negative stimuli were biased (assimilated) toward the values suggested by the cues. This bias persisted throughout the experiment, even as the stimuli associated with each cue became more homogenous over time, suggesting that affective predictions may not be updated (at least in the short term) if discrepancies are so small that they are not consciously acknowledged. According to Bar (2007, 2009b) perceptual processes operate similarly to other cognitive process involving predictions; in each case basic sensory information relating to an input is extracted quickly so that it can be linked to analogous representations of similar inputs held within memory. As such any perceptual or cognitive judgement about an ambiguous stimulus should be based primarily upon existing experience with similar items.

The experiments discussed within chapters 3 – 6 have observed that affective predictions often appear to function similarly to predictions within the perceptual domain. Just as a strong expectation about an ambiguous stimulus can cause an observer to misperceive, a strong expectation of liking/disliking appears to cause an individual to like an ambiguous stimulus more/less than they otherwise would. Furthermore, the influence of affective expectations appears to be fairly robust even if the discrepancy between the expectation and the actual experience slowly increases over time. As the previous paragraph has already stated, there is currently little research relating to how quickly predictions are updated in the face of contradictory information. Because of this, the current chapter adopted the inverse approach to previous chapters to test whether the results observed over time for affective judgements in chapter 6, would be replicated for perceptual judgements.

The central tendency bias described in the previous chapter was first defined by Hollingworth (1910). In an earlier experiment, Hollingworth (1909) had observed that such biases were introduced when participants were asked to reproduce fine movements of different amplitudes.
Blindfolded participants were first given a pencil and instructed to trace along a groove etched into a piece of card. The card was then replaced with a plain sheet of paper, and participants were instructed to perform a movement of the same amplitude as before. Participants attended multiple sessions, and in each session a different range of movement amplitudes were presented. The range of movements included in Series A varied between 10 and 70mm, in Series B between 30 and 150mm, and in Series C between 70 and 250mm. For each session, reproductions of movement amplitudes were biased toward the mean of the range of movement amplitudes for the presented series (see Figure 7.1). When participants were asked to reproduce only one amplitude repeatedly, the bias disappeared. This indicates that for each amplitude reproduction, participants incorporated both their actual estimate of movement amplitude, and their prior knowledge about all of the movement amplitudes which had occurred within the session.

Figure 7.1  Mean reproductions of movement amplitudes from Hollingworth (1909). Diagram taken from Laming (1997). Series A, B and C represent sessions where participants were asked to reproduce different movement. The filled points represent sessions where only one movement amplitude was reproduced repeatedly.
The same biases observed by Hollingworth (1909) have been reported in numerous other magnitude estimation tasks involving different types of perceptual judgements. For example, Petzschner, Maier, and Glasauer, (2012) observed central tendency biases using a VR based task for both estimates of distance travelled and estimates of turning angle. Other recent studies have reported similar biases for judgements of colour (Olkkonen et al., 2014), and comparison tasks involving judgements of different line lengths (Ashourian & Loewenstein, 2011).

The extent of central tendency biases can also be influenced by other factors, and a range effect has been reported where central tendency biases become more pronounced as the range of values on which sensory estimates are based increases (Cheng, Spetch, & Hoan, 2010; Kowal, 1993). Viewed within a Bayesian framework, range effects may be explained by an increasing reliance upon prior knowledge about the range of values, as actual estimations of values within that range become more unreliable. The prior on which an estimate is based could be broadly fixed to the centre of a range of values, but according to Petzschner, Glasauer, and Stephan (2015) this would not explain the sequential effects which are also observed in magnitude estimation tasks. Alternatively, both sequential effects and range effects could be explained by an ‘online prior’ which is updated on a trial-by-trial basis.

If instances of central tendency bias increase in line with the range of values that is being judged, then the reverse effect should be observed if that range contracts. In experiment 6, cued participants demonstrated a central tendency bias for affective judgements of rectilinear and curvilinear patterns. The presence of predictive cues before each trial indicated whether the pattern presented in a subsequent trial would be curved or angular, and based upon the assumption that cues were a reliable indicator of actual stimulus value, each pattern was judged as being more curved or angular than it actually was (Huttenlocher et al., 2007, 2000). In addition, the cues became less reliable indicators of stimulus values as the experiment continued, and the range of values that was being judged contracted (from pairs of 90% and 10% curved presentations, to 80% and 20% curved presentations etc.). Yet, compared with participants in the non-cued condition, cued participants continued to bias their affective judgements toward the values suggested by the cues. One explanation for the persistence of this bias could be that cued participants were still updating their priors, and this assertion may be supported by the fact that the difference in liking responses for each type of cue reduced over subsequent experimental blocks (see Appendix D.6). However, it is currently unclear whether this bias would have been completely extinguished if further blocks had been included.
Just as our judgements about the affective qualities of items are based upon multiple sources of noisy sensory information, and are therefore particularly susceptible to the biasing effects of affective expectancies; our sense of time, which is generated by multiple sensory organs, relies heavily upon information drawn from similar expectancies which are related to temporal context (Shi & Burr, 2016). For example, differences in interpretations of timing intervals have been reported when intervals are perceived through different sensory modalities (Wearden, Todd, & Jones, 2006; Wittmann, 2013).

Therefore, tasks involving the measurement of timing intervals, represent a useful means of testing whether the results from experiment 6 generalise to perceptual judgements. A further advantage of testing these effects using measurements of time, is that because the interpretation of time is so heavily determined by temporal context, there is no need to repeatedly present the same set of anchor items, and the same durations can be presented repeatedly in different contexts (as long as they do not rapidly deviate substantially from expectation) without drawing attention to the experimental manipulation.

One factor which is likely to influence the sense of time, and indeed any judgement of magnitude, is scalar variability, which refers to the phenomenon where, as the average length of a series of perceptual estimates increases, those estimates become less reliable measures of a true value (Rakitin et al., 1998). Jazayeri and Shadlen (2010), recently conducted an experiment to test whether people are implicitly aware of this fact, and whether this knowledge is incorporated into estimates of timing intervals. For example, if two timing intervals are measured as being broadly the same length are they actually reported as being the same? Or, do people adopt a Bayesian approach, where the intervals could be reported as different lengths, because knowledge about how long an interval should have been (prior), will influence the extent to which the measurement (likelihood) is relied upon as an accurate estimate of the true interval duration?

In their experiment, participants measured and then reproduced intervals based upon three different overlapping prior distributions. For each distribution, participants had to first complete a learning phase (session 1). Once learning was considered complete, participants proceeded to the testing phase (session 2), and then the process was repeated for the remaining distributions. Whilst the three learned distributions contained a number of overlapping intervals, the three distributions could be broadly classified as short, intermediate and long, based upon average interval length. The authors reported that for all three prior distributions, production times were systematically biased toward the mean of the distribution, and that this
effect was most pronounced for the longest intervals, where uncertainty surrounding interval measurements should have been highest. Furthermore, they found that this bias could be best accounted for by a Bayesian observer model.

In the current study, the experiment described above was adapted to incorporate symbolic cues to test whether (i) the cues would introduce biases for timing intervals similar to those observed in Jazayeri and Shadlen's (2010) original study, and (ii) to test whether those biases would adapt in a manner similar to those reported for the affective judgement task reported in experiment 6. Previously, perceptual studies have reported that symbolic cues can influence perception leading to an increased ability to detect between category differences, at the expense of discriminability for within category differences (Petzschner et al., 2012). Consequently, in the present study participants learned to reproduce timing intervals based upon 3 different overlapping prior distributions simultaneously, and each distribution was associated with a different predictive cue. The three distributions were gradually manipulated so that they became more homogenous, and the number of overlapping intervals across the 3 distributions increased over the course of the experiment.

Experiment 7

7.2 Methods

7.2.1 Participants

Fifty participants aged between 18 and 49 (33 females, mean age = 23.60, SD = 7.03) were recruited for the study from the staff, postgraduate and undergraduate populations at the University of Manchester. Participants were required to attend 2 or 3 sessions (depending on performance) over the course of one week. The second session was conducted within 4 days of the first session, and the final session was always conducted 7 days after the first session.

All participants reported normal or corrected to normal vision, and had no reported history of neurological or psychiatric disorders. Participants received course credit, or were paid at a rate of £6 per hour for their time. As an added incentive, the participant with the highest accuracy across sessions 2 and 3 won a £50 for an online retailer. All participants gave written informed consent to participate, and all testing procedures were approved by the University of Manchester ethics committee.
7.2.2 Stimuli and apparatus

Stimuli for the experiment consisted of coloured square borders, and coloured opaque squares, all drawn at a size of 300 by 300 pixels. Borders and opaque squares were presented in one of three different colours: Aqua (RGB value: 0, 255, 255), yellow (RGB Value: 255, 255, 0) and purple (RGB value: 159, 0, 197).

Stimuli were presented using a desktop computer running PsychoPy software (Peirce, 2007) and a 21-inch LCD monitor at a resolution of 1920 by 1080 pixels. All responses were recorded using a standard keyboard and a wired mouse placed on a desk directly in front of the participant.

7.2.3 Design and procedure

The three experimental sessions were lab based and they all involved variations on the same timing paradigm. Sessions 1 and 2 acted as training tasks, and all of the data included in the analysis was collected during session 3. Session 1 took between 15 and 30 minutes to complete, session 2 took 45 minutes, and sessions 3 took around an hour and a half.

Basic trial design

The timing task used in all sessions was an amended form of the Ready-Set-Go time interval reproduction task created by Jazayeri and Shadlen (2010). Participants observed various different timing intervals signalled by the appearance of coloured borders and opaque squares. In each trial participants were expected to replicate an observed interval by holding down a key on the keyboard.

To signal the start of a trial, a fixation cross was presented for 1000ms. After the fixation cross disappeared a random delay was introduced (between 250 and 850ms), before the coloured outline of a square appeared in the centre of the screen signalling the start of a time interval (Ready). The end of an interval was signalled when the coloured outline was replaced by an opaque square of the same colour which remained on screen for 100ms (Set). Once the opaque square disappeared (Go), participants were expected to replicate the interval which they had just observed by holding down the space bar for the same duration. Measurement of production times began at the moment the space bar was pressed, and ended when the space bar was released. Production times needed to begin within 500ms after the disappearance of the opaque square, and participants were notified of occasions when their responses occurred outside of this window by ‘too early’ or ‘too late’ prompts. Positive feedback was provided.
when intervals were replicated accurately, and if a response was inaccurate, participants were informed whether it was ‘too fast’ or ‘too slow’ (see Figure 7.2)

Figure 7.2   Example trial sequence from experiment 7.

In each trial, time intervals could be drawn from one of three different overlapping distributions: short, intermediate and long (see Figure 7.2). Each distribution was represented by a different colour which determined how various elements (the fixation cross, the square outline, and the opaque square) appeared in each trial. The colours associated with each distribution were counterbalanced between participants, but for the sake of clarity, here the distributions are represented in the following manner: short distribution (aqua), intermediate distribution (yellow), and long distribution (purple).
a. Figure 7.3  

Time intervals used for short, medium and long presentations. (a) Shows the distributions of the congruent short, medium and long cued trials. Each distribution contained 11 different intervals, and 3 intervals from the long and short distributions overlapped with the intermediate distribution. (b) Shows the distributions from the second round of trials conducted as part of the final session (minor discrepancy trials). The short and long distributions contained 14 intervals: the 11 original intervals from the congruent trials plus three additional intervals which had previously only been seen as part of the intermediate distribution. The intermediate distribution contained 15 intervals: the 11 intervals seen during the congruent trials, plus 2 additional intervals at either end of the distribution which had previously only been seen as part of the short and long distributions. (c) Shows the distributions used in the extreme discrepancy trials. The short, intermediate and long distributions returned to their original levels from the congruent distributions, but 4 additional extreme intervals were included with the short and long distributions.
Session 1  Training task 1

At the start of the first session participants were informed that they would be viewing and replicating time intervals throughout the study. They were also told that the visual elements present in each trial would be shown in different colours, but the meaning of the colours was not explained explicitly.

So that participants could become accustomed to measuring and replicating time intervals, the task used in session 1 was described as a game which became progressively more difficult. Participants were told that they would begin at level 1, and that progression to higher levels would occur when they reached a certain accuracy threshold. Conversely, after progressing from level 1, if their accuracy fell below a certain threshold, they could also regress to lower levels. Task difficulty increased at each level as the accuracy window for production times became smaller. At level 1, production times were considered correct when they fell within a 60% window of the actual presented time interval. At level 2, the accuracy window shrank to 40%, at level 3 it became 20%, and at level 4 it became 10%. The aim for participants during session 1 was to reach and complete level 4. Progression from one level to the next occurred when production times were correct in 5 successive trials for each of the three distributions. Level demotions occurred if incorrect production times were recorded in 5 successive trials for any of the three distributions. Nine practice trials (three intervals randomly drawn from each distribution) were included at the start of the task.

To incentivise accurate performance throughout the task, participants could leave as soon as they completed level 4, and still be compensated for 30 minutes of participation. If participants did not complete level 4 after 30 minutes had elapsed, the program auto-terminated and participants were free to leave. Intervals in each trial were drawn in a pseudo-random manner from the three different distributions (Figure 7.3a), and the nature of the task meant that the number of included trials varied between participants.

Session 2  Training Task 2

In session 2 participants were reminded that they should try to respond as accurately as possible, and that only those who scored above 75% accuracy would be invited back for the final session. So that participants could re-familiarise themselves with the task, 33 practice trials were included at the start of the session (each interval from the three distributions was presented once, in a random order). The basic trial procedure for the task was identical to session 1, except the level progression component was removed, and the difficulty of the task
remained consistent throughout the session (production times had to be within 20% of the
presented interval for each trial to be considered correct).

In total the session contained 462 trials: 33 practice trials and 429 live trials. The live trials
consisted of 13 repeats of each of the 11 intervals contained within the short, medium, and long
distributions (Figure 7.3 a) presented in a pseudo-random order (For a complete trial
breakdown see Appendix G.1). The average interval lengths for each distribution were as
follows: short 670.5ms, intermediate 952.9ms, and long 1235.3ms. The live trials were divided
into 4 blocks with breaks provided between blocks. Participant accuracy was calculated by the
presentation software on-the-fly, and overall accuracy was presented immediately once all trials
had been completed. Participants who scored above 75% accuracy were invited to attend the
final session, and those who scored below this level were debriefed and given the appropriate
compensation for their participation in sessions 1 and 2.

Session 3 Experimental manipulation

The basic trial procedure for session 3 was the same as the one described for session 2 and
session 1, but feedback in each trial was limited. In sessions 1 and 2, participants were informed
whether incorrect responses were too fast or too slow. In session 3, participants were only told
if responses were correct or incorrect. The session was effectively split into 2 halves, but all
trials occurred as part of a continuous presentation (with breaks) to hide this fact from
participants.

The first half of the session (congruent trials) was identical to session 2 and contained the same
number of practice trials (33), and live trials (429). In the second half of the session (minor
discrepancy trials), the underlying short, intermediate, and long distributions were amended
(see Figure 7.3 b). Based upon the new distributions, the second half of the session contained
140 short cued trials, 140 long cued trials, and 160 intermediate cued trials (for a complete trial
breakdown see Appendix G.2). The average durations for each of the distributions became:
short 723.45ms, intermediate 952.9ms, and long 1182.35ms. Thus, the average short duration
was longer, the average long duration was shorter, and the average intermediate duration
remained the same.

In total session 3 consisted of 902 trials (33 practice trials and 869 live trials). After completing
all trials participants were debriefed and compensated for their participation.
Session 3 Extreme discrepancy trials

After the initial round of data collection was completed, there was an opportunity to collect some data for an additional task looking at the impact of more extreme discrepancies. This task was added to the end of session 3 for the final 8 participants recruited for the experiment. The basic trial design used in this task was the same as the other sessions, but the task employed a different underlying distribution (see Figure 7.3 c). These trials (extreme discrepancy trials) were included after the minor discrepancy trials, and like the standard version of session 3, all of the trials appeared to be part of the same presentation to avoid highlighting the manipulation.

The distribution used in the extreme discrepancy trials returned to the same base distribution used for the congruent trials, but a number of extremely discrepant short and long interval presentations were included (For a complete trial breakdown see Appendix G.3). The addition of the extreme discrepancy trials added half an hour onto the standard length of session 3. The manipulation meant the inclusion of 155 further trials on top of the standard session 3 design. This resulted in a total of 1057 trials (33 practice trials and 1024 live trials).

7.2.4 Data analysis

Data analysis was only conducted upon data from session 3. For the congruent trials, an average production time was calculated for each participant at every interval within the short, intermediate, and long distributions. Regression coefficients were then calculated using the least-squares method. To establish whether participants’ responses deviated away from the line of equality toward the mean of each distribution, three one sample t-tests (one for each distribution) were used to test whether overall regression coefficients were significantly different from 1 (the line of equality). This process was repeated for the short, intermediate and long distributions from the minor discrepancy trials.

To test whether deviations from the line of equality altered as interval durations moved away from prior expectations, regression coefficients were also subjected to a two-way repeated measures ANOVA with the following within-subjects factors: interval type (3 levels: short, intermediate or long), trial type (2 levels: congruent, minor discrepancies).

For the minor discrepancy trials, paired samples t-tests were used to compare mean production times from trials using interval/cue pairings which were not present in the originally learned distributions, with the same intervals when they were paired with the cues from the learned distributions. Production times for short cued intervals which were not present within the originally learned short distribution (882.3ms, 917.6ms and 952.9ms), were compared against
the same intervals when they were presented with cue associated with the intermediate distribution. Production times for intermediate cued intervals which were not present within the originally learned intermediate distribution (short cued intervals: 758ms and 741.1ms; long cued intervals: 1164.7ms and 1200ms), were compared against the corresponding intervals when they were paired with the cues corresponding to the learned distributions. Production times for long cued intervals which were not present in the original long distribution (952.9ms, 988.2ms and 1023.5ms) were compared against the corresponding intervals when they were paired with the cue associated with the intermediate distribution.

For the participants who completed the extreme discrepancy trials, paired t-tests were used to compare average production times for extremely discrepant intervals (long intervals paired with short associated cues, or short intervals paired with long associated cues), against production times for the same intervals when they were paired with the opposite cue. The presentation times which were considered extremely discrepant when presented with the short cue were 1305.9ms, 1341.2ms, 1376.5ms and 1411.8ms; while the presentation times which were considered extremely discrepant when presented with the long cue were 484ms, 529.3ms, 564.6ms and 599.9ms.

Paired t-test were also used to compare accuracies in session 2, session 3 congruent trials, session 3 minor discrepancy trials, and session 3 extreme discrepancy trials. For all t-tests Bonferroni corrections are applied when appropriate.

In cases where Mauchly’s test revealed that the assumption of sphericity had been violated, the degrees of freedom have been corrected using Greenhouse-Geisser estimates of sphericity. An alpha level of .05 has been used for all statistical tests, and effect sizes are reported here as partial eta squared ($\eta^2_p$) where 0.05, 0.1, and 0.2 correspond to small, medium, and large effects, respectively (Cohen, 1988).

7.3 Results

Participants who did not respond correctly in >75% of trials during session 2 were not included in the analysis. Consequently, response data from 21 participants were omitted from the results leaving a remaining sample of 29 participants (21 females, mean age = 21.79, SD = 5.63).

7.3.1 Accuracy measures

Paired samples t-tests ($p < .025$ Bonferroni corrected) showed that overall accuracy (in terms of the number of correct responses) was higher in session 2 ($M = 82.72\%, SD = 4.62\%) compared to the congruent presentations from session 3 ($M = 79.07\%, SD = 9.85\%), t(28) = 2.62, p = 0.14.
There was no difference in accuracy between the congruent trials from session 2, and the minor discrepancy trials from session 2 (M = 78.55%, SD = 9.14%), \(t(28) = .38, p = .709\). There was also no difference in accuracies between the congruent trials from session 2 (M = 75%, SD = 13.78%) and the highly discrepant trials (M = 75.13%, SD = 11.62%), \(t(7) = -.037, p = .972\).

### 7.3.2 Regression coefficients

One sample t-tests revealed that participants’ responses for all three interval durations (short, intermediate, and long) showed a significant deviation away from the line of equality. This finding was replicated in both the congruent trials, and the minor discrepancy trials (see table 7.1)

<table>
<thead>
<tr>
<th>Trial type</th>
<th>Interval</th>
<th>M</th>
<th>SD</th>
<th>Df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congruent</td>
<td>Short</td>
<td>.62</td>
<td>.23</td>
<td>28</td>
<td>-8.80</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Congruent</td>
<td>Intermediate</td>
<td>.63</td>
<td>.24</td>
<td>28</td>
<td>-8.04</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Congruent</td>
<td>Long</td>
<td>.55</td>
<td>.27</td>
<td>28</td>
<td>-8.77</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Discrepant</td>
<td>Short</td>
<td>.62</td>
<td>.25</td>
<td>28</td>
<td>-8.08</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Discrepant</td>
<td>Intermediate</td>
<td>.58</td>
<td>.24</td>
<td>28</td>
<td>-9.43</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Discrepant</td>
<td>Long</td>
<td>.55</td>
<td>.21</td>
<td>28</td>
<td>-11.25</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

The repeated measures ANOVA showed no main effect of interval type indicating that the extent to which production times deviated away from the observed intervals was similar for the short, medium and long intervals \(F(1, 40) = 2.49, p = .112\). There was also no main effect of trial type \(F(1, 28) = .60, p = .445\, and the interval type x trial type interaction was also not significant, \(F(2, 44) = .92, p = .384\).
Figure 7.4  
*Response data from congruent trials in session 3. Individual production times from each trial are represented by the small dots, and average production times for each interval are represented by the larger connected dots. Average production times from each cue condition deviated away from the line of equality (diagonal line) toward the mean of each cued distribution (horizontal lines).*

7.3.3 Production times for discrepant intervals (minor discrepancy trials)

Analysis of the minor discrepancy trials revealed that production times for intervals which were not originally part of the short distribution were reduced significantly ($p < 0.0125$ Bonferroni corrected) when they were paired with the short duration associated cue ($M = 868.9ms, SD = 89.93ms$), rather than the intermediate duration associated cue ($M = 931.3ms, SD = 68.42ms$), $t(28) = -4.04, p < .001$. Similarly, production times for intervals which were not originally part of long distribution, were significantly increased when they were paired with the long duration associated cue ($M = 1032.8ms, SD = 77.35ms$), rather than the intermediate associated cue ($M = 971.1ms, SD = 67.11ms$), $t(28) = 4.28, p < .001$. Production times for intervals which were not originally included at the shorter end of the intermediate distribution were significantly longer when they were presented with the intermediate duration associated cue ($M = 797.6ms, SD = 73.12ms$), instead of the short duration associated cue ($M = 752.1ms, SD = 66.59ms$), $t(29) = 2.93, p = .007$. Finally, production times for intervals which were not originally associated with the longer end of the intermediate distribution were significantly shorter when they were
presented with the intermediate duration associated cue ($M = 1071.9\text{ms}, SD = 90.44\text{ms}$), rather than long duration associated cue ($M = 1154\text{ms}, SD = 80.41\text{ms}$), $t(28) = -4.89, p < .001$.

**Figure 7.5**  *Response data from minor discrepancy trials in session 3. Individual production times from each trial are represented by the small dots, and average production times for each interval within the modified distributions are represented by the larger connected dots. Production times for intervals which were not part of the original distributions are represented by coloured squares. For clarity individual production times are only shown for intervals which differed in the new distribution. Average production times from each cue condition continued to deviate away from the line of equality (diagonal line) toward the mean of each cued distribution (horizontal lines).*

### 7.3.4 Production times for discrepant intervals (extreme discrepancy trials)

Production times for intervals which were previously only present in the short distribution, did not differ significantly when they were presented with the long duration associated cue ($M = 621.47, SD = 52.44\text{ms}$), rather than the short associated cue ($M = 643.37\text{ms}, SD = 64.77\text{ms}$), $t(7) = 1.34, p = .222$. Furthermore, for intervals which were previously only present as part of the long distribution, no significant difference in production times was observed between presentations with the longer duration associated cue ($M = 1195.96\text{ms}, SD = 85.87\text{ms}$) and
presentations with the short duration associated cue ($M = 1221.52\text{ms}, SD = 71.92\text{ms}$), $t(7) = -1.98, p = .088$.

![Graph showing response data from extreme discrepancy trials in session 3. Individual production times from each trial are represented by the small dots, and average production times for each interval are represented by the larger connected dots. Extreme discrepancy trials are represented by coloured squares. For clarity individual production times are only shown for intervals which differed in the new distribution. Means for highly discrepant trials are shown in addition to means for each distribution.]

**Figure 7.6** Response data from extreme discrepancy trials in session 3. Individual production times from each trial are represented by the small dots, and average production times for each interval are represented by the larger connected dots. Extreme discrepancy trials are represented by coloured squares. For clarity individual production times are only shown for intervals which differed in the new distribution. Means for highly discrepant trials are shown in addition to means for each distribution.

### 7.4 Discussion

The current study was conducted to test whether the results of the affective judgement task described in experiment 6, would be replicated in a perceptual judgement task. Firstly the study aimed to test whether the central tendency biases for timing intervals that had previously been observed by Jazayeri and Shadlen (2010), would persist if the 3 prior interval distributions were learned simultaneously using sensory cues. If similar biases were observed, the second aim of the study was to investigate whether the extent of those biases altered, as the average value and range of the learned prior distributions was manipulated.
After participants had learned the interval duration and cue associations from the training task, central tendency biases were observed for all three prior distributions during the congruent trial phase of the testing session. When compared against the actually presented interval durations, regression coefficients revealed that participants’ responses deviated toward the mean of each distribution. During the minor discrepancy trial phase, these biases remained consistent for the short, intermediate and long intervals, even as the average duration of the three interval types was brought closer together. These biases were confirmed through comparisons of responses for the overlapping intervals within each distribution. Overlapping short and intermediate cued interval durations were judged as significantly longer when presented with the intermediate associated cue, and significantly shorter when presented with the short associated cue. The opposite pattern of results was observed for overlapping intermediate and long cued intervals, and even when the same interval (952.9ms) had occurred within all three prior distributions, judgements of its length were assimilated toward the duration suggested by the cue that it was presented with. This indicates that here, like previous experiments discussed within this thesis, predictive cues (colours associated with each distribution) were sufficient to promote anchoring effects where stimulus judgements were biased toward those anchors.

The influence of participant expectations based upon prior knowledge, appeared strongest for the long cued intervals, but the extent of this bias did not differ significantly between the three types of cued intervals. According to Jazayeri and Shadlen (2010) people should be implicitly aware that their estimates of longer intervals are less reliable, and their judgements of these durations should be more heavily biased toward existing expectations. Consistent with this hypothesis, they reported that bias was significantly greater for long intervals when compared to short intervals. The changes made to the original timing paradigm could provide one explanation for the absence of these effects in the present research. For example, in the current study production times were measured using a press, hold, and release procedure. In the original study, the start of each production time coincided with the appearance of the go signal, and ended when a key was pressed. Nevertheless, the interval reproduction task has also been replicated by Cicchini, Arrighi, Cecchetti, Giusti, and Burr (2012) who reported comparable effects to the original study, using a similar press, hold and release procedure to the one described here.

A more plausible explanation is that the similar biases observed across the three distributions used in the present study, could be accounted for by the accuracy window that was used. In the original study, an adaptive accuracy window was employed and data collection did not begin
until the variance in participants’ production times had stabilised. In the current study, because participants learned all three distributions simultaneously, a static accuracy window was employed, meaning that, compared to the original study production times could be less accurate and still be considered correct. As a result, the variance in production times for the current study was far greater overall, across the three distributions.

In spite of the increased variances, the consistent production time biases observed across the congruent, and minor discrepancy phases of the current study, indicate that participants’ prior knowledge about how long each interval should have been, remained similarly influential throughout. Overall participant accuracy was significantly lower during the testing session, compared to the earlier practice session, and this can be explained by the removal of the “too fast” and “too slow” feedback in the testing session (Rakitin et al., 1998); but no significant changes in accuracy were observed between the congruent and minor discrepancy testing phases. Therefore, even though prior expectations remained influential throughout the testing session, participants’ production time estimates appear to have been updated with reference to the actually observed intervals. Again, these findings are in line with the existing accounts of the anchoring effect where individuals are expected to make insufficient adjustments when stimuli are judged to be different from anchor items (Furnham & Boo, 2011; Tversky & Kahneman, 1974)

In the minor discrepancy phase of the study, the range of each of the prior distributions was expanded to incorporate a number of additional overlapping intervals. When the range of a series of values increases, sensory estimates of those values become less reliable, and as previous chapters in this thesis have discussed, when estimates become less reliable, peoples’ judgements are biased more heavily toward their expectations. For example, Malapani et al. (1998) reported differences in timing biases for Parkinson’s disease patients, with and without their dopaminergic medication. Studies have shown that dopamine plays a role in the modulation of internal timekeeping (Meck, 1996, 2005), and timing biases were reduced when patients were tested with their medication, presumably because patients have greater confidence in the reliability of their sensory estimates when they are medicated.

In the current study, as the range of intervals across each distribution increased, their associated means were also drawn closer together, but the extent to which participants differentiated between overlapping cued intervals remained similar throughout. If participants’ expectations about intervals were being updated in an online manner as suggested by Petzschner et al. (2015), differentiation between the three cued interval types should, in theory,
have reduced, as the average duration associated with each cue became more similar. However, as Cicchini et al. (2012) suggest, prior knowledge does not correspond exactly to the physical distribution of stimuli, and should instead be thought of as a simplified neural representation of it. Therefore, it appears quite possible that, since only minor discrepancies were included in the second phase of the study, those discrepancies fell within an acceptable boundary of error, and the existing expectancies about the stimuli continued to provide the most reliable means of estimating them. In chapter 6, similar results were reported for affective responses to stimuli, and prior knowledge about the affective value of stimuli is also thought to be represented inexactly because emotional experiences are episodic in nature (M. D. Robinson & Clore, 2002).

The observation that similar effects were observed in the both perceptual, and affective domains also lends support to Bar’s (2007, 2009b) assertion that the influence of predictions should be similar for perceptual processes and other cognitive processes.

One limitation of the current study, is that, in terms of range effects, it is difficult to directly compare the observed results with the results from the previous chapter. During the affective judgement study, the range of values to which participants were exposed, contracted over subsequent blocks (90% and 10% block 1, 80% and 20% block 2 etc.). In the current perceptual study, the entire range of values (across short, intermediate and long presentations) remained the same throughout (all presented intervals were between 494ms and 1411.8ms). Yet for both studies, the mean values associated with the groups of stimuli that were being affectively or perceptually judged were drawn closer together as the experiment progressed. Consequently, it appears that for both studies at least, participants’ prior expectations about the groups of stimuli that they were estimating, remained highly influential in shaping their perceptual or affective judgements. However, it remains unclear if these expectancy biases would persist, if the reliability of the cues, in terms of accurately predicting stimulus values, continued to deteriorate in the same slow, but steady manner.

The additional manipulation involving the extreme discrepancy trials, included at the end of the study for a small sample of participants, was intended to investigate the effect of a sudden large shift in experienced interval duration, away from expectation. The central tendency biases observed throughout this study support the predictions of ACT, which state that as long as a stimulus is considered ‘close enough’ to an expected value, judgements of that stimulus will be assimilated toward the expectation. Alternatively, contrast should occur when the discrepancy between stimulus and expectation becomes sufficiently large, and as previous chapters have discussed, contrast effects are comparatively rare. In the extreme discrepancy phase of the
current study, when the longest intervals were presented with the short associated cue, and vice versa, average production times moved in the opposite direction to the cues, but this effect was not significant. Still, this manipulation only included a small sample of the total number of participants, and it may be worthwhile to attempt a similar manipulation in future, using a larger number of participants.

7.5 Conclusion
This study aimed to investigate the influence of cues in a timing interval reproduction task, and to test whether the effects of a previous affective judgement task would be replicated when the timing cues became less reliable. Consistent with previous timing studies, cues promoted a central tendency bias where interval reproduction times deviated away from actual observed times, toward the average durations of learned distributions suggested by the cues. As the cues became slightly less representative of actual interval times, in a replication of the findings from the affective judgement study, participants continued to assimilate their production times toward the durations suggested by the cues. Contrast effects were investigated by creating larger discrepancies between the cues, and the actually presented intervals, but significant contrast effects were not observed. Taken as a whole, these findings suggest that in both perceptual and affective domains, expectations will continue to influence judgements, even as they become less reliable over time. The extent to which this effect will persist if expectations continue to gradually become less reliable is, however, currently unclear.
8 GENERAL DISCUSSION

8.1 Chapter Outline
The studies conducted in this thesis have explored the influence of expectations in determining affective experience, as well as investigating whether the extent of this influence changes over time. This final chapter is divided into several sections which discuss the findings of those studies. A summary of the main findings is presented first, followed by the interpretation of these findings within the context of the related literature, and a brief discussion of the practical applications for the present research. The primary methodological challenges and limitations associated with the studies are then discussed, as well as some suggestions for future work in this area.

8.2 Summary

8.2.1 Theme 1: Affective assimilation and contrast effects (Chapters 3 and 4)
In chapters 3 and 4, experiments were designed to provide evidence of both affective assimilation (chapter 3), and contrast effects (chapter 4) using interval scaled stimuli. After the initial issues experienced in experiment 1 (see methodological challenges), the results of experiment 2 showed that participants exhibited different PSEs when they were shown the same set of affectively ambiguous stimuli in negatively associated, and positively associated cue conditions. This finding indicates that, in line with the predictions of the AEM (T. D. Wilson et al., 1989), the cues produced a state of affective readiness which influenced participants’ affective thresholds for judging stimuli.

In chapter 4, differing assimilation and contrast effects were observed in response to cued presentations of happy and sad emotional faces. Overall, assimilation toward the cues was greater for the sad faces, but this difference could be attributed to the happy face advantage (Calvo & Beltrán, 2013; Calvo & Lundqvist, 2008). Assimilation also appeared to be greater for more optimistic participants when the discrepancy between the cue and the target was small or medium sized. Contrast effects were observed as the cue/target discrepancies grew larger, but only for the presentations involving happy faces. This appears to suggest that the larger target/cue discrepancies were more noticeable when participants were judging happier emotional expressions.
8.2.2 Theme 2: Maintaining affective assimilation effects over time (Chapters 5 and 6)

In chapter 5, responses to images of food items were used to explore the persistence of assimilation effects over time. However, the results from this study indicated that the chosen stimuli were not sufficiently ambiguous, and participants were not motivated to use the predictive cues to guide their affective ratings (see again methodological challenges). A revised version of the study was conducted in experiment 5, using curved and angular abstract patterns. The main findings replicated the results of experiment 6, where the presence of predictive cues influenced participants to adopt different affective thresholds for judging the same set of moderately smooth/sharp patterns. Furthermore, the cue/target associations persisted into a second session, conducted 1 week later, even without additional reinforcement of the positive and negative associations during the second session.

8.2.3 Theme 3: The boundaries of assimilation effects (Chapters 6 and 7)

Both experiment 6 and experiment 7 were conceived to investigate the boundaries of assimilation effects. Experiment 6 showed that once participants had learned a set of cue/target associations, those cues promoted a central tendency bias, where cued participants differentiated between groups of stimuli to a greater extent than non-cued participants. Even as the two groups of stimuli became more homogenous over time, cued participants continued to assimilate toward the values suggested by the cues, indicating that cued participants expected the difference in affective values between the two groups to remain consistent.

In chapter 7, to further investigate the boundaries of assimilation effects, participants experienced three overlapping timing distributions signalled by predictive cues. In the case of all three distributions, participants’ estimates of intervals deviated toward the mean of the distribution associated with each cue. Over time, the three distributions were made more homogenous so that an increasing number of time intervals could exist within multiple distributions, but participants’ estimations of overlapping intervals continued to deviate toward the average duration (short, intermediate or long) suggested by the presented cue. The inclusion of an extreme discrepancy condition at the end of the experiment for a small sample of participants, indicated that extreme cue/interval discrepancies may be sufficient to promote contrast effects, but this is only speculation at present, and no significant differences in interval estimations were observed between the differently cued short and long trials for this group.
8.3 Interpretation of findings

8.3.1 Theme 1: Affective assimilation and contrast effects (Chapters 3 and 4)

Predictive coding accounts of perception indicate that the brain does not attempt to model a given situation through bottom-up information alone, instead sensory information is interpreted through rules and predictions based upon statistical regularities (Barlow, 2001). Similarly, the AEM (T. D. Wilson et al., 1989) indicates that the affective interpretation of a stimulus is, in a large part, determined by how a person expects to feel about it. As long as the valence of a stimulus broadly adheres to an expected level, any judgement of that stimulus will be biased by the affective expectation (assimilation). The aim of experiment 2, was to provide a quantifiable measure of that bias.

The highly positive (attractive faces) and highly negative (unattractive faces) items predominantly used within the earlier blocks of the experiment, were intended to promote the expectation that the accompanying cues were associated with positive and negative categories of items. This expectation appears to have generalised to the intermediate items introduced later on in the experiment, as evidenced by the difference in observed PSE measures when the same intermediate items were paired with different cues. This finding indicates that participants’ affective ratings for the intermediate items assimilated toward the values suggested by the cues, yet the observation of affective assimilation effects is in itself, nothing new. What differentiates this research from other studies of affective assimilation effects is the application of psychophysical methods, which provided evidence of a systematic bias across a range of stimuli. This indicates that when items are broadly considered as part of a positive or negative category, similar affective biases will be applied to judgements of all of the stimuli within that category, even if those stimuli are judged as being affectively better or worse than one another.

The aims of the study conducted in chapter 4 were to provide evidence of both the previously discussed assimilation effect, and the contrast effect, where the discrepancy between an affective expectation and a stimulus is exaggerated if that discrepancy is sufficiently large and is noticed. The results of the study showed that, in line with the predictions of the AEM, when the discrepancies between the predictive cues and the subsequently presented emotional faces were small, participants’ judgements aligned with the values suggested by the cues. However, while judgements of presentations involving sad emotional faces continued to assimilate toward the predictive cues throughout the experiment, judgements involving happy emotional faces contrasted against the values suggested by the cues, as the cue/target discrepancy grew
larger. The observation of both assimilation and contrast effects is a fairly novel finding since within the affective judgement literature contrast effects are fairly rare. This is potentially due to the need for a high degree of expectancy discrepant information before contrast occurs; as Ditto and Lopez (1992) observed, there is evidence to suggest that more discrepant information is required to reach a non-preferred conclusion (expectation inconsistent), rather than a preferred conclusion (expectation consistent). There are also difficulties in designing experiments to promote both assimilation and contrast effects (see methodological challenges).

Following a review of the related literature, it appears that only a very limited number of studies such as Geers and Lassiter (1999, 2002, 2003, 2005), or more recently Davidenko et al. (2015), have reported evidence of both affective assimilation and contrast effects within the same experiment. These studies indicate that contrast effects are more likely to occur when participants have had prior exposure to an expectation-inconsistent stimulus (Geers & Lassiter, 2005), or when stimuli are judged in a more fine-grained manner (Geers & Lassiter, 1999, 2003). Neither of these factors was manipulated directly in the study conducted in chapter 4, but the finding that contrast effects were observed only in presentations involving happy faces, suggests that happy and sad faces were processed differently by participants. According to Schwarz and Bless (1992), assimilation occurs when two items are judged to be from the same category, and contrast occurs when those items are judged to be from different categories. They also suggest that contrast effects are more likely to be observed for experts rather than novices, because experts’ definitions of stimulus categories are more specific. When the present experiment was originally designed, it was thought that participants would find the task of judging happy and sad emotional faces similarly difficult, but the results showed that overall, participants rated the happy faces more accurately, presumably due to the happy face advantage (Calvo & Beltrán, 2013; Calvo & Lundqvist, 2008). This advantage or ‘expertise’ in terms of judging faces of one type, suggests that cue/target discrepancies were more noticeable for one set of emotional faces, and provides an explanation as to why contrast effects were only observed for happy face presentations.

One final observation from the study conducted in chapter 4, was that assimilation and contrast effects differed depending on whether a cue predicted a more positive or more negative emotion than the target. It is difficult to determine whether this finding was simply an artefact caused by the happy face advantage, and previous studies have indicated that affective expectations should have a similar influence regardless of whether they are positive or negative (Geers & Lassiter, 2002; Klaaren et al., 1994; T. D. Wilson et al., 1989). One potential
explanation is that this difference may be associated with the differing consequences of positive
and negative expectations. Negative expectations may lead to a sense of dread, but they reduce
the impact of a negative experience through affective attenuation, e.g. ‘I always knew that this
would be bad, so I am not surprised’. Furthermore, they lead to affective amplification when
events are more positive than expected, ‘I thought this was going to be bad, but it was actually
good, and I am delighted’. In contrast, the opposite is experienced when expectations are
positive. Positive expectations mean that the pleasant state of savouring an anticipated
outcome can be experienced, but delight from positive outcomes is reduced when they are
expected, and disappointment from negative experiences is amplified (Golub, Gilbert, & Wilson,
2009).

8.3.2 Theme 2: Maintaining affective assimilation effects over time (Chapters 5 and 6)
Based upon the earlier finding that affective expectations produced a systematic bias across a
range of stimuli which were believed to be part of the same category, the primary aim of the
studies described in chapter 5 and chapter 6 (experiment 5) was to examine whether that bias
persisted over time. In chapter 5, while assimilation effects were not observed at any stage
during the study, the null findings highlight the importance of ambiguity in encouraging the top-
down facilitation of perceptual (Bar, 2003; Summerfield et al., 2006) and affective judgements
(Deliza & MacFie, 1996; Hoch & Ha, 1986).

Participants did not appear to associate the predictive cues with positive or negative stimulus
presentations, because those stimuli were never presented ambiguously enough to necessitate
use of the cues. In the other experiments described in this thesis, the stimuli that participants
were asked to judge in a given experiment shared a basic category membership, e.g. in chapters
3 and 4 the target stimuli were all faces, and in chapter 6 the stimuli were abstract shapes.
However, participants were asked to judge those stimuli on the basis of their subordinate
category membership, e.g. attractive or unattractive faces, or faces displaying positive or
negative emotions. In chapter 5, the food images used as target stimuli in each trial could be
identified relatively easily, so even if the details of a specific target were hard to perceive, it was
clear that the target was an instance of a particular food. Herr, Sherman, and Fazio (1983)
suggest that a determinant of how stimuli are judged, is the accessibility of a relevant category.
Since the stimuli from chapter 5 were instances of fairly common food items, it appears that the
most accessible means of judging those stimuli was through their broad category membership,
and these judgements were influenced by prior experience accumulated outside of the
experimental sessions.
This issue was resolved in experiment 5, and the results from session 1 of the experiment, replicated the findings from experiment 2. Anchor items (smooth and sharp patterns) which were consistently paired with cues, promoted the expectation that visual cues were reliable predictors of positive and negative outcomes (Strack et al., 2016). Participants’ PSE measures, based upon intermediate presentations (moderately smooth or sharp patterns) paired with the positively and negatively associated cues, showed that judgements of the intermediate items assimilated toward the values indicated by the cues.

When participants returned for the second session of the experiment, a week later, their ratings for the same intermediate stimuli continued to differ according to the cues that preceded them, even though the anchor items were not part of the presentations from the second session. In session 2, there was no overall difference (in terms of curvature) between the patterns paired with the positively and negatively associated predictive cues, and from a predictive coding perspective, participants should have updated their affective expectations about the stimuli in light of this new evidence (Dunovan et al., 2014). Nevertheless, if the presentations from session 1 and session 2 are taken as a whole, the probability of highly positive outcomes was still greater for one set of cues, and the opposite was true for highly negative outcomes. Furthermore, because the second session contained only the intermediate items, it is possible that assimilation effects persisted because the discrepancy between expectation and reality was never so large that it was consciously acknowledged. These findings are consistent with those reported in the placebo literature by Colloca and Benedetti (2006), who found that prior experience of an analgesic effect from a treatment, produced a placebo effect when an inert version of the same treatment was administered up to 7 days later. The persistence of affective expectations is not a widely researched area, and when the present study was designed it was unclear if affective expectations would persist in the short to medium term. The results of the study indicate that recently acquired affective expectations persist in the medium term, and these expectations will bias judgements of similar stimuli. However, it is currently unclear how long the affective assimilation effect from the present research would have lasted before extinction, or how much discrepant information would have been required before affective expectations were updated.
8.3.3 Theme 3: The boundaries of assimilation effects (Chapters 6 and 7)

Having observed evidence that expectations could produce assimilation effects which persisted over time, the goal of the studies described in chapter 6 (experiment 6) and chapter 7, was to test the boundaries of assimilation effects. In experiment 6, expectation was manipulated so that one group of participants (cued) had access to predictive information about the affective value of an upcoming target stimulus, and another group did not (non-cued). At the start of the experiment cued participants learned to associate predictive cues with positive and negative outcomes. The presence of these cues increased their sensitivity, indicated by the slope of their psychometric functions, promoting them to affectively discriminate between smoother and sharper versions of the same patterns to a greater extent than the non-cued participants.

ACT (Hovland et al., 1957) indicates that the size of a discrepancy determines whether a stimulus will be assimilated toward prior expectations, or contrasted against them. During the early blocks of the experiment, when the differences between the smoother and sharper versions of each pattern were more easily perceptible, assimilation effects had been expected, but it had been assumed that as these differences grew smaller, assimilation effects would diminish, potentially leading to contrast. The results of the study indicate that this was not the case, and cued participants continued to assimilate toward the values suggested by the cues, even to the point where there was no difference between the positively and negatively cued patterns.

One explanation for the persistence of assimilation effects throughout the experiment, is that the cue/target discrepancies were never so large that they were consciously acknowledged (see also limitations). Consequently all of the stimulus presentations may have remained within the latitude of acceptance surrounding a participants’ internal reference point for the positive and negative stimuli. However, the proportion of liked responses for the smoother intermediate patterns diminished over each subsequent block as curvature was decreased, and this finding was observed for both the cued and the non-cued participant groups. This indicates that even in the cued group, participants acknowledged that the difference between the smoother and sharper patterns was decreasing. According to ACT and the AEM, once a deviation from expectation is acknowledged, contrast should occur. ACT claims that the latitude of acceptance is a finite boundary surrounding a participant’s internal reference for a stimulus, therefore according to ACT these boundaries should have remained consistent throughout the experiment. Instead, it appears that over the course of the experiment, participants updated their prior expectation, or their reference points for the ideal prototypes representing each
category, in response to the new evidence acquired throughout the study (Huttenlocher et al., 2007). As a result a central tendency bias persisted throughout the experiment, where participants continued to view stimuli paired with each cue as more representative of their associated category (positive/negative) than they really were.

The aim of chapter 7 was to investigate whether the effects observed in the previous study, central tendency biases that persist even as category boundaries are blurred over time, would also be observed within the perceptual domain. Previously, Jazayeri and Shadlen (2010) reported that temporal context influenced interval timing when participants learned, and then replicated, intervals from 3 different prior distributions across different sessions. In chapter 7, similar biases were observed when the three distributions were learned simultaneously, and temporal context was manipulated by pairing each distribution with a different coloured cue. For all three distributions, central tendency biases were observed as responses deviating toward the mean of each distribution (short, intermediate and long).

When the three distributions were manipulated so that they became more homogenous, participants continued to base their responses upon inferences drawn from prior knowledge, as evidenced by the different response times observed for the overlapping intervals within each distribution. Like the previous study, this suggests that when small changes were made to stimuli (in this case timing distributions), people may slowly update their internal representations of those stimuli, but if those changes are not sufficiently jarring they will still be biased by the influence of their prior expectations (assuming that an interval is representative of a short, intermediate, or long category). Taken as a whole these findings indicate that assimilation will generally occur when stimuli are manipulated so that deviations from expectation remain small, and expectations can be updated slowly. Whether this slow manipulation will occur indefinitely, or whether a point will be reached where the assimilation effect eventually breaks, remains unclear. The results from the extreme discrepancy trials included at the end of the study from chapter 7, hinted that when cue discrepancies are so large they become jarringly obvious, contrast may occur, but these results were not significant, and more research is required in this area.

8.4 Practical Implications

The primary driver of research in this area is the development of products which contain fewer harmful ingredients. On average, the daily intake of ingredients like sugar and salt is higher than the recommendations suggested by the World Health organisation, which can lead to issues like cardiovascular disease (World Health Organisation, 2015; Zandstra et al., 2016). One potential
obstacle preventing people from shifting to healthier diets is the assumption that healthier alternatives automatically equate to being poorer tasting products (Kähkönen et al., 1999; Norton et al., 2013; Tuorila et al., 1994). The findings from the current research suggest that, if consumers harbour such expectations, their judgements of healthier products are likely to be negatively biased, even if the rewarding properties of those products can be matched to their less healthy alternatives. The fact that these biases persist in the short to medium-term, could make dietary change more difficult, since people are unlikely to purchase a product in future, if they have had multiple experiences where a product has failed to disconfirm their existing prejudices.

On the other hand, if consumers have a favoured brand, expectations relating to the quality of that brand may be sufficient to mask changes to the rewarding qualities of a product, specifically in cases where a healthier version is less rewarding than the full-fat/sugar alternative. The results of the research presented in this thesis also suggest that if properties of a product are changed gradually over time, assimilation toward expectation will generally occur. This research could have implications for other unhealthy behaviours such as smoking, but since such habits also rely upon chemical dependencies, it appears unlikely that a reduction in the concentration of harmful ingredients could be masked through expectation persistence alone.

8.5 Methodological challenges

8.5.1 How much uncertainty is too much uncertainty?
Uncertainty plays a key role in determining the influence of expectations. When stimuli are presented in a way that is sufficiently ambiguous, top-down information in terms of prior expectations and predictions will be relied upon more heavily to make perceptual or affective judgements. Conversely, if a stimulus is presented in such a way that its value or identity is easily discernible, then bottom-up analysis of that stimulus may be completed so quickly that judgements are not influenced by top-down processes (Bar, 2003). This was a concern during the design of all of the studies discussed within this thesis, and evidence of this phenomenon was observed in chapter 5, where stimuli were clearly recognised as examples of familiar items, and it was not necessary for participants to use the predictive cues to determine their affective values.

Alternatively, in order to investigate assimilation effects, it is necessary to first encourage participants to develop affective expectations based upon the rules of an experimental paradigm. Consequently, the presentation of items intended to anchor affective expectations,
must not be so ambiguous that it is difficult to infer affective associations between those items and a set of predictive cues. This problem was experienced in experiment 1. The results of preliminary studies conducted before the main experiment had indicated that participants could easily differentiate between attractive/unattractive composite morphs presented at levels of 30% and 70%. However, during the main experiment, it appears that the difference between these presentations was not substantial enough for them to serve as effective anchors.

8.5.2 Testing affective expectations in psychophysical experiments.

Psychophysical experiments can be used as an effective means of testing bias (the primary goal of this thesis), but each experimental condition that is being tested requires a considerable number of repeat trials. For example, Kingdom and Prins (2010) suggest that around 400 trials is a reasonable aim when estimating the slope and threshold of a psychometric function. As the previous section has discussed, before assimilation or contrast effects can be investigated, it is necessary to first promote the expectation that cues are a reliable predictor of stimulus value. When attempting to test the boundaries of assimilation effects, the cues used to promote expectation must, by necessity, become less reliable. This serves to highlight the experimental manipulation, and an awareness of a priming manipulation has previously been shown to influence assimilation effects (Strack et al., 1993). This makes the study of contrast effects particularly difficult as more obvious discrepancies mean that participants will begin to distrust the source of their expectations, and as the number of expectancy discrepant trials increases, the influence of top-down expectations will diminish (Mayo, 2015; Woods et al., 2010).

8.6 Limitations and Future Work

8.6.1 Lack of consequence

All of the studies reported in this thesis have investigated affective responses to visual stimuli, and in the majority of cases, participants have been forced to make a binary like/dislike choice about a target stimulus. Controls were present within the studies to ensure that participants were attending to the stimuli and not responding randomly, but participants were in no way rewarded or punished for their choices. In most real-world scenarios the choices that people make between stimuli have definitive consequences (Appleton, Gentry, & Shepherd, 2006). For example, when a person makes a choice between 2 competing products at the supermarket, they have to pay for the product that they choose, and that product is bought in the knowledge that it will likely be used in the future. For the present research, preliminary studies were conducted before each of the main experiments in chapters 3 and 5 (See Appendix A.2 and B.2) to try to ensure that participants could actually perceive the differences between the levels of...
the stimuli that they were asked to judge. This was done with the aim of ensuring that any changes in response bias were based upon affective rather than perceptual judgements. However, as Witt, Taylor, Sugovic, and Wixted (2015) argue, when measuring bias and sensitivity through techniques like signal detection, because a factor like bias may reflect a perceptual bias, a memory bias or a response bias, it is often difficult to separate responses reflecting a decisional component from those reflecting a perceptual component. Therefore, some questions still remain about whether the same assimilation effects would reported here would also be seen in real-world scenarios where affective decisions carry greater consequences. These questions could be addressed in a follow up study if participants’ decisions involved delivery of actual rewards.

8.6.2 Boundaries of assimilation

Throughout this thesis, affective assimilation effects have been reliably observed in response to liked/disliked category expectations. These expectations have persisted as the liked/disliked groups of stimuli have become progressively better/worse, even to the point where stimuli that are expected to be liked or disliked, are functionally identical. Yet, none of the studies have tested whether assimilation effects would persist beyond this point, and it is currently difficult to identify a means by which this issue could be resolved (see methodological challenges). Even if this challenge cannot be overcome in future experiments, the design of the current studies has been reasonably conservative so as not to highlight the experimental manipulation.

The anchor items used within experiments 2, 5, 6 and 7 from this thesis appear to have promoted anchoring effects similar to those reported elsewhere (see Furnham & Boo, 2011 for review). In the case of the current research, participants adjusted their responses over time as the stimuli slowly became more discrepant with their expectations, but these adjustments were typically insufficient because of the assimilation bias introduced by the positive / negative anchors. Previous research has indicated that perceptual biases based upon existing expectations are increased when bottom-up sensory information is weak or ambiguous (Summerfield et al., 2006). Therefore, even with the difficulties associated with testing the boundaries of assimilation effects and increasing the overall size of the discrepancy between expectation and actual affective value; if the present research were adapted so that the rate of stimulus change occurred more rapidly, or in larger increments, this should still serve to disambiguate the stimuli somewhat. Consequently, one means of testing the persistence and boundaries of assimilation effects further, could be to modify the affective value of stimuli in a
manner that is broadly similar to the present studies, but to do so using between groups designs to adjust the rate of stimulus change for different participants.

8.6.3 Persistence over time

One of the primary goals of this research was to test whether hedonic cues would serve to bias stimulus liking over time. Predictive coding frameworks such as those suggested by Huttenlocher et al. (2000), Bar (2003) and Friston (2005), indicate that, at least in terms of stimulus perception, humans adopt a Bayesian inferential approach where predictions and expectations are updated gradually to incorporate new information and statistics. Based upon the limited research relating to time-course within the affective domain, the present research investigated stimulus liking over the course of one week, and replicated the findings from previous salt reduction studies which had shown that consumption was not affected over time if stimulus quality was manipulated using the “stealth” approach (Bolhuis et al., 2011; Girgis et al., 2003).

What is not clear, however, is whether the cues’ ability to adjust liking for better/worse items toward expectations, remained stable after the one week period. For example, one interpretation for the present results could be that since human knowledge of probabilities is represented inexactly, participants did not have sufficient opportunity to update their existing expectations. Consequently, if participants were exposed to stimulus discrepancies over a longer period the observed effects could disappear. When considering this issue in terms of an ideal category exemplar or “prototype” as described by Huttenlocher et al. (2007); it is currently unclear whether the cues encouraged participants to update their ideal category representations, or whether those representations remained stable, and participants instead adjusted the boundaries of tolerable stimulus variation which surrounded those prototypes. Therefore, in addition to exploring whether affective expectations persist over longer periods, future work should also look to investigate whether stimulus liking for superior/inferior items persists when the original unadjusted item is resampled. This is of particular importance when considering the real world applications for this type of study, because consumers will often be faced with, and may have the opportunity to sample, competing products at the point of sale.

8.6.4 Controlling for familiarity

In order to promote stimulus associations which could be manipulated over the course of 1 or more experimental sessions, wherever possible, the studies described in this thesis have attempted to control for experiences outside of those sessions. Previous studies have observed that it is often difficult to promote new stimulus associations when stimuli are already highly
familiar (Cacioppo, Marshall-Goodell, Tassinary, & Petty, 1992; Kuenzel et al., 2010; E. Robinson et al., 2013). Since the current research was primarily concerned with manipulating those associations, it can be argued that this was a sensible approach to adopt, and similar approaches have been adopted in many previous studies investigating affective assimilation effects. Geers and Lassiter (2005) have been critical of this approach, arguing that we rarely encounter stimuli which are truly novel, and that the majority of our experiences are not devoid of prior exposure. Therefore, another potential avenue for future research could be to test whether affective persistence is similar for more familiar stimuli.

8.6.5 Mood Manipulations

In chapter 4 it was observed that assimilation and contrast effects differed for positive and negative primes. The AEM indicates that positive and negative expectations should be treated equally, and the size of the discrepancy necessary to produce assimilation or contrast should remain similar in both cases. However, the current research has provided further evidence that personality traits like optimism also influence the extent of assimilation effects. Recent research has indicated that assimilation is more prevalent when moods are manipulated positively, and greater contrast is observed when moods are manipulated in the opposite direction (Huntsinger, 2014). It is also possible that the moods of participants may have influenced the effectiveness of the anchor items used within the current research, since other studies have shown that the extent of anchoring effects may be modulated by through factors like mood and participant expertise (Englich & Soder, 2009). Elsewhere, other studies involving binocular rivalry have shown that the perception of different emotional faces can be dominated by positive stimuli (smiling faces) when positive moods are induced, and that the opposite is true (scowling faces dominate) when negative moods are induced (E. Anderson, Siegel, & Barrett, 2011). Therefore, combining the current paradigms with mood manipulations could be another direction for future study, to test whether such manipulations influence the boundaries of affective assimilation effects, or whether these boundaries are resistant to changes in mood.

8.6.6 Imaging paradigms

Another, avenue for future research could be to adapt one of the current paradigms to incorporate the use of neuroimaging techniques. fMRI would probably not be suitable for this purpose since it lacks the necessary temporal resolution, but EEG has previously been employed in numerous experiments involving the study of prediction errors. If, as the current findings indicate, people adapt their internal representations for affective categories over time in
response to relatively minor deviations from expectation, it would be interesting to investigate whether this adaptation is accompanied by changes in the prediction error signal.

8.7 Conclusion
The primary aims of this thesis were as follows: (i) to investigate whether hedonic cues bias affective judgements in the same way that sensory cues bias perceptual judgements, (ii) to test whether assimilation effects persist over time, and (iii) to investigate the boundaries of assimilation effects. The findings indicate that just as prior expectations bias perception in favour of interpretations based upon previously encountered stimuli, affective expectations bias judgements toward those of an associated category when a stimulus is thought to belong to that category. As long as affective expectations are not challenged by substantial contradictory evidence, and stimulus value remains open to interpretation, the same stimuli will be judged differently in the short to medium term, based upon the assumption that they belong to different affective categories.

Furthermore, if the values of items from 2 affectively distinct categories are slowly updated over time so that they become less distinct, people will adapt to the new values, but still distinguish between the items due to priors based on their original category membership. From a commercial perspective, this suggests that once a preference for a product has been developed, even if the rewarding properties of that product deteriorate, the expectation of liking will smooth over these changes as long as they are not sufficiently jarring. However, questions still remain about whether these category based assimilation effects will persist if the rate of change is increased, or what will happen if the consumer is given the opportunity to compare the modified product against its original level of performance.
REFERENCES


Golub, S. a, Gilbert, D. T., & Wilson, T. D. (2009). Anticipating one’s troubles: the costs and


Raichle, M. E., MacLeod, a M., Snyder, a Z., Powers, W. J., Gusnard, D. a, & Shulman, G. L.


Tuorila, H., Cardello, A. V., & Lesher, L. L. (1994). Antecedents and Consequences of
Expectations Related to Fat-free and Regular-fat Foods. *Appetite*, 23, 247–263.


A.1 Preliminary Study 3a: Identifying attractive and unattractive faces

Rationale

Preliminary study 3a was conducted to collect attractiveness data for a set of male and female faces. The results of this study were then used to select a series of attractive and unattractive faces for use in preliminary study 3b (Appendix A.2), and experiments 1 and 2 in chapter 3.

Methods

Participants

Twenty participants aged between 18 and 24 (16 females, mean age = 19.35, SD = 1.72) were recruited for preliminary study 1 from the undergraduate population at the University of Manchester. All participants received course credit for their participation in the study.

Stimuli

A series of 122 images of male and female faces (63 females) were selected from the neutral sub-sets of the Karolinska directed emotional faces (KDEF) database (Lundqvist et al., 1998) and the Max Planck FACES database (Ebner et al., 2010). Each image was grey-scaled, resized, and cropped to a size of 562 x 762 pixels preserving the aspect ratio of the original image. All images displayed the head and shoulders of a single male or female wearing a grey t-shirt with no facial hair, jewellery or eye-glasses present. All persons depicted were aged between 19 and 31 (both KDEF and FACES).

Design and procedure

A within-participants, repeated measures design was employed where participants rated the entire stimulus set of 122 faces twice. Faces were presented individually in random order, and after all 122 faces had been viewed and rated, the presentation order was randomised again for the next iteration. Constraints built into the program ensured that no three faces of the same gender could appear consecutively at any time.

The image set was accessed via a web link that was distributed via email. This link directed participants to a web-based presentation, coded in JavaScript. Each trial began with a fixation cross (500ms), followed by a presentation of one of the faces (1000ms). After viewing a face, participants were prompted to rate that face. Responses were collected using on-screen
response buttons and the available responses were based on a 9-point scale ranging from: 1 - ‘Highly unattractive’, to 9 - ‘Highly attractive’. This rating prompt remained on-screen until the participant had provided a response.

Each participant completed 244 trials (across all both repetitions), and all trials had to be completed in a single session. The task took approximately 30 minutes to complete, and at the end of the task response data was stored on a server to be retrieved by the experimenter at a later date.

Data Analysis

The average rating of each face for each presentation (presentation 1, presentation 2) was calculated using the responses from all twenty participants. Separate paired samples t-tests were used for the male and female faces, to compare ratings from presentation 1, with those from presentation 2.

Attractive and unattractive face were selected for use in preliminary study 3b and experiment 1, based upon the responses from all participants averaged across the two presentations. Once these attractive/unattractive faces were selected, paired samples t-tests were used to compare the average ratings of the unattractive and attractive face sets. An alpha level of .05 was used for all statistical tests.

Results

Attractiveness ratings remained broadly consistent across the two presentations of each face. Ratings for the female faces were not significantly different between the two presentations, t(62) = -1.55, p = .127 (presentation 1 females: M = 4.32, SD = 1.00; presentation 2 females: M = 4.36, SD = .94). Similar results were also observed for the ratings of the male faces, with no significant difference in overall attractiveness ratings between the two presentations, t(58) = -1.62, p = .110 (presentation 1 males: M = 4.00, SD = 1.04; presentation 2 males: M = 4.08, SD = 1.13).

After establishing that attractiveness ratings remained consistent across both rating times, attractiveness ratings from each participant were then averaged for each face across both presentation times. The associated variances were then examined to investigate whether participants agreed about which faces were the most, and least attractive. This revealed several outliers: One female face, with a variance of 3.83 (variance of female face set: M = 2.01, SD =
.56), and 2 male faces with variances of 3.27 and 2.90 respectively (variance of male face set: \( M = 1.50, SD = .57 \)). These faces were not considered for use in any further part of the study.

Providing they could be matched with a similar item to create a convincing composite morph, the most attractive and unattractive faces, were then selected for use in preliminary study 3b, and later experiment 1. The 12 faces chosen to represent the attractive group in preliminary study 3b were considered significantly more attractive than the 12 faces which represented the unattractive group, \( t(19) = 11.75, p < .001 \) (attractive group: \( M = 5.81, SD = .79 \); unattractive group: \( M = 2.91, SD = .79 \)). The 14 faces chosen for the attractive group in experiment 1 were also considered significantly more attractive than the 14 faces in experiment 1’s unattractive group, \( t(19) = 11.12, p < .001 \) (attractive group: \( M = 5.76, SD = .84 \); unattractive group: \( M = 2.94, SD = .80 \)).
A.2 Preliminary Study 3b: Establishing identity boundaries for composite morphs

Rationale

Previous research (Jacques & Rossion, 2006; Levin & Beale, 2000; Verosky & Todorov, 2010) has observed that when 2 faces are combined in a composite morph image, if that image contains >65% of a given face, the composite morph image will be reliably identified as that face.

Preliminary study 3b was conducted prior to experiment 1 in chapter 3, to test whether the number of face identifications for a particular parent face, differed across a series of different morphing increments.

Methods

Participants

Twenty-one participants aged between 21 and 33 (13 females, mean age = 27.37, SD = 2.97) were recruited for preliminary study 2 from the staff and post-graduate population at the University of Manchester. Participants were not compensated for their participation in the study.

Stimuli

Twenty-four images of faces (12 males and 12 females) were taken from the stimulus set described in preliminary study 3a. Faces were selected based upon the attractiveness data collected in the earlier study, and whether a face could be matched with another face of the same gender to create a convincing composite morph. For both the male and female face sets, 6 faces were classified as attractive ($M = 5.82$, $SD = 1.46$), 6 faces were classified as unattractive ($M = 2.91$, $SD = 1.20$), and each attractive face was paired with an unattractive face of the same gender to create 12 attractive/unattractive face pairs.

Morphed composite images were generated from each face pair using the techniques outlined by Steyvers (1999). Within each face pair, the attractive face was classified as parent face A, and the unattractive face was classified as parent face B. To create the composite images, control points were placed around the key feature locations of each of the parent faces. These locations were defined as: around the eyes, the nose, the mouth, the eyebrows, and around the circumference of the face, neck and hair. Control points formed a grid across each parent face and meshed together during the morphing process to create the composite images (see Figure A2.1). Approximately 92 control points were placed around each parent face, and all composite
morphs were created using SmartMorph picture morphing software (http://logicnet.dk/SmartMorph/). If the attractive face (parent face A) is considered 0% and the unattractive face (parent face B) is considered 100%, morphed composite images between each face pair were created at increments of: 30%, 40%, 45%, 50%, 55%, 60% and 70%. This resulted in the creation of a total stimulus set consisting of 108 images: 24 parent face images, and 84 composite morph images.

![Example of control points and mesh for one pair of parent faces](image)

**Figure A2.1** Example of control points and mesh for one pair of parent faces

**Design and procedure**

Preliminary study 3b was designed as an online study and coded in JavaScript, so that a web-link could be distributed to participants via email. After opening the link, and filling in their details, participants completed a series of trials featuring only the composite morphed face images. At the start of each trial, 30% and 70% morphs generated from one pair of parent faces, were presented on the left and right hand sides of the screen. After 1000 milliseconds, a third ‘transitional’ morphed face, generated from the same pair of parent faces, appeared in the centre of the screen. The transitional morphed face could be presented at any one of the following levels: 40%, 45%, 50%, 55% or 60% (Figure A2.2). The 30% morph, the 70% morph, and the transitional morph were then presented together for 1000 milliseconds, after which time the 30% and 70% morphs disappeared, and the transitional morph remained on-screen for a further 500 milliseconds. At the end of each trial participants were prompted to state which of
the initial 30%/70% faces the transitional morph more closely resembled using on-screen buttons. The prompt remained on-screen until participants had provided a response. Participants were not given feedback relating to the accuracy of their responses.

In total the study consisted of 240 trials; the twelve face pairs at 30% and 70% morph were presented once for each of the five levels of the transitional morph (40%, 45%, 50%, 55%, 60%), and this process was repeated four times per participant. The order of face presentations was randomised within each repetition, and also randomised across participants. The lateral position of the 30%/70% presentations was counterbalanced across the task.

![Morphing increments used in preliminary study 2 and experiment 1](image)

**Figure A2.2** Morphing increments used in preliminary study 2 and experiment 1

**Data Analysis**

For the 6 male, and 6 female face pairs, the number of 70% (parent face B) identifications per participant was calculated at each level of the transitional morph presentations (40%, 45%, 50%, 55% and 60%). The number of 70% identifications provided by each participant for each level was then averaged across each face pair. The data were analysed using a two-way repeated measures ANOVA, with the factors ‘morphing level’ (5 levels) and ‘gender’ (2 levels).

In cases where Mauchly’s test revealed that the assumption of sphericity had been violated, the degrees of freedom have been corrected using Greenhouse-Geisser estimates of sphericity. An alpha level of .05 has been used for all statistical tests, and effect sizes are reported here as
partial eta squared ($\eta^2_\rho$) where 0.05, 0.1, and 0.2 correspond to small, medium, and large effects, respectively (Cohen, 1988).

**Results**

A significant main effect of morphing level was observed indicating that, as the level of morphing in the transitional morph image increased the number of 70% (parent face B) identifications increased, $F(2, 34) = 99.98, p < .001, \eta^2_\rho = .83$. Pairwise comparisons ($p$ values are Bonferroni corrected) revealed a significant increase in the number of 70% identifications at each level of the transitional morph over the previous level (see Figure A2.3; all $p < .001$).

![Graph showing identity boundaries of facial morphs used in preliminary study 2. The graph shows the percentage of 70% morph (parent face B) identifications for each level of the transitional morph image. Error bars represent standard errors of mean responses observed at each level of morphing.](image)

**Figure A2.3**  *Identity boundaries of facial morphs used in preliminary study 2. The graph shows the percentage of 70% morph (parent face B) identifications for each level of the transitional morph image. Error bars represent standard errors of mean responses observed at each level of morphing.*
A significant main effect of gender was observed, $F(1, 20) = 10.45, p = .004, \eta^2_p = .34$ (males: $M = 52.18\%$, $SD = 24.60\%$; females: $M = 44.56\%$, $SD = 23.16\%$), but there was not a significant interaction between gender and morphing level $F(4, 80) = 2.06, p = .287$. Whilst male faces produced a greater number of 70\% (parent face B) identifications overall, the number of 70\% identifications at each level of the transitional morph did not significantly differ according to the gender of the presented composite morph.
A.3 Preliminary Study 3c: Testing the assumption that attractive faces are more liked

Rationale
In chapter 3 experiment 1 participants were asked to rate a series of morphed face images in terms of attractiveness using a 0-9 scale. In chapter 3 experiment 2, participants were instead asked to classify faces as either ‘liked’ or ‘disliked’ using a 2 AFC paradigm. It was assumed that more attractive faces would be classified as liked, and that the opposite would be true for the unattractive faces. Preliminary study 3c was conducted prior to experiment 2 to test this assumption.

Additionally, presentation durations for experiment 2 were reduced compared to those from experiment 1. Preliminary study 3c also served as an opportunity to test whether responses to the attractive/liked, and unattractive/disliked faces remained similar when they were presented for shorter durations than those that had been employed previously.

Methods

Participants
A total of 20 participants aged between 23 and 34 (10 females, $M = 27$ $SD = 3.69$) were recruited for the study from the postgraduate population at the university of Manchester. Participants were divided into 2 groups; 10 participants (5 females, $M = 28$ $SD = 3.77$) were asked to rate the attractiveness of a set of faces (attractiveness group), and ten participants (5 females, $M = 27$ $SD = 3.60$) were asked to state whether they liked or disliked each of the faces (liking group).

Stimuli and Apparatus
The stimulus set for this study consisted of 80 faces. These 80 faces were split into 2 categories, attractive/liked faces, and unattractive/disliked faces. The attractive/liked face category consisted of the 20 faces of each gender rated as the most attractive during preliminary study 3a. The unattractive/disliked category consisted of the 20 faces of each gender rated as the least attractive during preliminary study 3a.

The task was coded in E-prime software (Psychology Software Tools, Pittsburgh, PA), and was run on a computer with a 17-inch monitor at a resolution of 1280 by 1024 pixels. A standard UK layout keyboard was used to collect responses.
Design and Procedure

Participants in both groups attended a single 15 minute session where they were asked to provide multiple attractiveness/liking ratings for the entire series of 80 faces. Each trial began with a fixation cross (500ms) followed by a face presentation (1000ms for the attractiveness group, or 250ms for the liking group). After each presentation participants were asked to provide either an attractiveness (1-9) rating using the numeric keyboard, or a liking (liked/disliked) rating using two keys at opposite ends of the keyboard.

The reason for the difference in presentation times between the two groups is because the liking group required more presentations, and an increased viewing time would have made the study prohibitively long. Furthermore, the shorter durations more accurately represented the presentation conditions that were intended for use in experiment 2, so they offered an opportunity to test whether the stimulus groupings (attractive/unattractive or liked/disliked) persisted when presentation times were reduced.

Participants in the liking group completed 640 trials. These 640 trials consisted of 8 presentations of each of the 80 faces, and participants provided a like/dislike response for each presentation. Participants in the attractiveness group completed 320 trials. Each of the 80 faces was presented 4 times and participants provided an attractiveness rating for each presentation. In both the attractiveness group and the liking group, stimuli were presented in a pseudo random order. All 80 faces were presented once per response block and presentation order within each block was randomised.

Data Analysis

For the participants in the attractiveness group, an average rating for each face was calculated based upon the four stimulus presentations. For the liking group, the percentage of liked responses to the 8 presentations of each face for each participant was calculated instead. Following these initial calculations, correlations between the average ratings from each group (average liking (%) and average attractiveness (0-10)) were examined for each face. The 7 most attractive and 7 most unattractive faces from the present study were also compared, with their counterparts from experiment 1 using paired samples t-tests. Attractiveness ratings from experiment 1 were based upon the average of the 2 online presentations and the single
presentation at the beginning of the lab session. An alpha level of .05 was used for all statistical tests.

Results

The relationship between attractiveness and liking

A positive correlation was observed between attractiveness ratings and liking ratings, \( r = .950, p \) (one tailed) < .001. The R-Squared value of .90 indicates that approximately 90% of the variance in the liking ratings for faces within this set can be accounted for by the attractiveness ratings (see Figure A3.1).

![Figure A3.1](image)

**Figure A3.1**  The relationship between liking and attractiveness for faces presented in preliminary study 3.

An examination of the attractive/unattractive face groupings originally based upon the attractiveness scores from preliminary study 1 (see Figure A3.2) revealed that in the vast majority of cases, these groupings persisted in preliminary study 3. The fact that the variance in liking ratings can be largely accounted for by attractiveness ratings, was encouraging in terms of
the protocol for experiment 2. It was hoped therefore, that if, as expected, attractiveness ratings remain fairly stable across responders, then the selection of a series of faces which are broadly grouped in terms of attractiveness, should produce liking ratings which are grouped similarly.

![Figure A3.2](image.png)

Figure A3.2  Distribution of attractiveness / liking ratings for attractive and unattractive face groupings.

Attractiveness ratings

In addition to assessing whether attractiveness ratings and liking ratings were significantly correlated, this study was conducted to address a problem with the stimuli from the experiment 1. In experiment 1, because faces had to be perceptually similar due to the morphing manipulation, the difference between the attractive and unattractive stimuli was not as large as it could have been otherwise. The morphing manipulation was removed for experiment 2 allowing for the selection of a broader range of faces for the attractive and unattractive sets. When compared to the 7 most attractive faces of each gender used in experiment 1, the most
attractive faces of each gender taken from the present study were significantly more attractive, 
$t(26) = -5.14, p < .001$, (experiment 1 faces: $M = 5.77, SD = .49$; new face set: $M = 6.57, SD = .32$). The 7 most unattractive male and female faces from the new face set were also significantly less attractive than those used in the previous study $t(26) = 2.54$, $p = .009$, (experiment 1 faces: $M = 2.58, SD = .20$; new face set: $M = 2.94, SD = .50$).
APPENDIX B

B.1 Example page from participant response sheet

<table>
<thead>
<tr>
<th>Participant Number:</th>
<th>Date:</th>
</tr>
</thead>
</table>

Please use this sheet to record your responses during the study. Please record your response by placing an X in the box which you think matches the emotional intensity of the presented face.

1.

<table>
<thead>
<tr>
<th>SAD 6</th>
<th>NEUTRAL</th>
<th>HAPPY 6</th>
</tr>
</thead>
</table>

2.

<table>
<thead>
<tr>
<th>SAD 6</th>
<th>NEUTRAL</th>
<th>HAPPY 6</th>
</tr>
</thead>
</table>

3.

<table>
<thead>
<tr>
<th>SAD 6</th>
<th>NEUTRAL</th>
<th>HAPPY 6</th>
</tr>
</thead>
</table>

4.

<table>
<thead>
<tr>
<th>SAD 6</th>
<th>NEUTRAL</th>
<th>HAPPY 6</th>
</tr>
</thead>
</table>

5.

<table>
<thead>
<tr>
<th>SAD 6</th>
<th>NEUTRAL</th>
<th>HAPPY 6</th>
</tr>
</thead>
</table>

6.

<table>
<thead>
<tr>
<th>SAD 6</th>
<th>NEUTRAL</th>
<th>HAPPY 6</th>
</tr>
</thead>
</table>

7.

<table>
<thead>
<tr>
<th>SAD 6</th>
<th>NEUTRAL</th>
<th>HAPPY 6</th>
</tr>
</thead>
</table>

8.

<table>
<thead>
<tr>
<th>SAD 6</th>
<th>NEUTRAL</th>
<th>HAPPY 6</th>
</tr>
</thead>
</table>
B.2  Questionnaire item: Life orientation test revised (LOT-R)

Participant Number:  Age:  Gender:

Please be as honest and accurate as you can throughout. Try not to let your response to one statement influence your responses to other statements. There are no "correct" or "incorrect" answers. Answer according to your own feelings, rather than how you think "most people" would answer.

A = I agree a lot
B = I agree a little
C = I neither agree nor disagree
D = I DISagree a little
E = I DISagree a lot

Please circle the appropriate response for each statement.

1. In uncertain times, I usually expect the best.

A         B         C         D         E

2. It's easy for me to relax.

A         B         C         D         E

3. If something can go wrong for me, it will.

A         B         C         D         E

4. I'm always optimistic about my future.

A         B         C         D         E

5. I enjoy my friends a lot.

A         B         C         D         E
6. It's important for me to keep busy.

A B C D E

7. I hardly ever expect things to go my way.

A B C D E

8. I don't get upset too easily.

A B C D E

9. I rarely count on good things happening to me.

A B C D E

10. Overall, I expect more good things to happen to me than bad.

A B C D E
### APPENDIX C

#### C.1 List of images used in the GNAT and rating tasks

*Images taken from the FoodCast research image database* (Foroni et al., 2013)

<table>
<thead>
<tr>
<th>Item</th>
<th>FRIDA ID</th>
<th>FRIDA Rating</th>
<th>FRIDA Group</th>
<th>Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cake Slice</td>
<td>AFO034</td>
<td>63.92</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Brass Spoon</td>
<td>AFO039</td>
<td>38.90</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Salad Server</td>
<td>AFO045</td>
<td>59.74</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Hand Blender</td>
<td>AFO052</td>
<td>59.96</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Electric Whisk</td>
<td>AFO053</td>
<td>60.96</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Zester</td>
<td>AFO056</td>
<td>54.46</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Cheese Grater</td>
<td>AFO057</td>
<td>65.30</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Rolling Pin</td>
<td>AFO065</td>
<td>59.04</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Nut Cracker</td>
<td>AFO090</td>
<td>54.37</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Skimmer</td>
<td>AFO091</td>
<td>61.65</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Pastry Cutter</td>
<td>AFO098</td>
<td>56.41</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Oyster Knife</td>
<td>AFO032</td>
<td>42.62</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Tea Strainer</td>
<td>AFO030</td>
<td>64.55</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Fork</td>
<td>AFO048</td>
<td>66.19</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Whisk</td>
<td>AFO055</td>
<td>61.33</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Cheese Slicer</td>
<td>AFO057</td>
<td>65.30</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Spoon Set</td>
<td>AFO068</td>
<td>51.48</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Spatula</td>
<td>AFO095</td>
<td>59.36</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Wooden Spoon</td>
<td>AFO041</td>
<td>67.70</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Can Opener</td>
<td>AFO003</td>
<td>59.22</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Tenderiser</td>
<td>AFO006</td>
<td>48.09</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Item</td>
<td>Code</td>
<td>Price</td>
<td>Category</td>
<td>Type</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------</td>
<td>--------</td>
<td>-------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Baking Tray</td>
<td>AFO051</td>
<td>64.81</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Corkscrew</td>
<td>AFO027</td>
<td>64.23</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Silver Spoon</td>
<td>AFO040</td>
<td>59.54</td>
<td>Artificial Food Objects</td>
<td>Non-Food</td>
</tr>
<tr>
<td>Frankfurter</td>
<td>TF156</td>
<td>48.22</td>
<td>Transformed Food</td>
<td>Intermediate 1</td>
</tr>
<tr>
<td>Dates</td>
<td>NF032</td>
<td>46.30</td>
<td>Natural Food</td>
<td>Intermediate 1</td>
</tr>
<tr>
<td>Rice Cakes</td>
<td>TF059</td>
<td>47.87</td>
<td>Transformed Food</td>
<td>Intermediate 1</td>
</tr>
<tr>
<td>Blue Cheese</td>
<td>TF066</td>
<td>44.27</td>
<td>Transformed Food</td>
<td>Intermediate 1</td>
</tr>
<tr>
<td>Cereal</td>
<td>TF028</td>
<td>52.64</td>
<td>Transformed Food</td>
<td>Intermediate 2</td>
</tr>
<tr>
<td>Avocado</td>
<td>NF012</td>
<td>52.83</td>
<td>Natural Food</td>
<td>Intermediate 2</td>
</tr>
<tr>
<td>Kidney Beans</td>
<td>NF033</td>
<td>52.41</td>
<td>Natural Food</td>
<td>Intermediate 2</td>
</tr>
<tr>
<td>Croquette</td>
<td>TF047</td>
<td>54.81</td>
<td>Transformed Food</td>
<td>Intermediate 2</td>
</tr>
<tr>
<td>Green Onion</td>
<td>NF086</td>
<td>57.67</td>
<td>Natural Food</td>
<td>Intermediate 3</td>
</tr>
<tr>
<td>Mushroom</td>
<td>NF039</td>
<td>59.29</td>
<td>Natural Food</td>
<td>Intermediate 3</td>
</tr>
<tr>
<td>Peas</td>
<td>NF082</td>
<td>56.24</td>
<td>Natural Food</td>
<td>Intermediate 3</td>
</tr>
<tr>
<td>Cracker</td>
<td>TF043</td>
<td>57.30</td>
<td>Transformed Food</td>
<td>Intermediate 3</td>
</tr>
<tr>
<td>Celery</td>
<td>NF092</td>
<td>62.90</td>
<td>Natural Food</td>
<td>Intermediate 4</td>
</tr>
<tr>
<td>Figs</td>
<td>NF035</td>
<td>63.29</td>
<td>Natural Food</td>
<td>Intermediate 4</td>
</tr>
<tr>
<td>Lettuce</td>
<td>NF024</td>
<td>64.21</td>
<td>Natural Food</td>
<td>Intermediate 4</td>
</tr>
<tr>
<td>Radish</td>
<td>NF090</td>
<td>60.36</td>
<td>Natural Food</td>
<td>Intermediate 4</td>
</tr>
<tr>
<td>Asparagus</td>
<td>NF011</td>
<td>66.60</td>
<td>Natural Food</td>
<td>Intermediate 5</td>
</tr>
<tr>
<td>Cucumber</td>
<td>NF026</td>
<td>66.95</td>
<td>Natural Food</td>
<td>Intermediate 5</td>
</tr>
<tr>
<td>Egg</td>
<td>NF095</td>
<td>66.30</td>
<td>Natural Food</td>
<td>Intermediate 5</td>
</tr>
<tr>
<td>Sliced Loaf</td>
<td>TF089</td>
<td>66.67</td>
<td>Transformed Food</td>
<td>Intermediate 5</td>
</tr>
<tr>
<td>Carrot</td>
<td>NF019</td>
<td>71.04</td>
<td>Natural Food</td>
<td>Intermediate 6</td>
</tr>
<tr>
<td>Pomegranate</td>
<td>NF058</td>
<td>70.84</td>
<td>Natural Food</td>
<td>Intermediate 6</td>
</tr>
<tr>
<td>Blackberry</td>
<td>NF063</td>
<td>73.27</td>
<td>Natural Food</td>
<td>Intermediate 6</td>
</tr>
<tr>
<td>Food</td>
<td>Code</td>
<td>Rottenness</td>
<td>Status</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>------------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>Olives</td>
<td>TF084/TF085</td>
<td>71.86</td>
<td>Transformed Food</td>
<td>Intermediate 6</td>
</tr>
<tr>
<td>Watermelon</td>
<td>RF002</td>
<td>10.67</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Orange1</td>
<td>RF003</td>
<td>9.10</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Orange2</td>
<td>RF004</td>
<td>10.73</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Bananas</td>
<td>RF005</td>
<td>4.05</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Meat</td>
<td>RF007</td>
<td>4.23</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Onion</td>
<td>RF010</td>
<td>3.65</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Green Beans</td>
<td>RF011</td>
<td>10.19</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Hamburger</td>
<td>RF015</td>
<td>2.57</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Lemon</td>
<td>RF016</td>
<td>6.89</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Corn</td>
<td>RF017</td>
<td>6.40</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Apple1</td>
<td>RF018</td>
<td>2.81</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Apple2</td>
<td>RF019</td>
<td>7.80</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Apple3</td>
<td>RF020</td>
<td>3.35</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Melon</td>
<td>RF022</td>
<td>8.37</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Bread1</td>
<td>RF024</td>
<td>3.15</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Bread2</td>
<td>RF026</td>
<td>7.55</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Chilli1</td>
<td>RF027</td>
<td>5.85</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Chilli2</td>
<td>RF028</td>
<td>6.26</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Pear</td>
<td>RF029</td>
<td>4.50</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Peach1</td>
<td>RF030</td>
<td>6.58</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Peach2</td>
<td>RF031</td>
<td>2.08</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Peach3</td>
<td>RF032</td>
<td>4.13</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Fish</td>
<td>RF033</td>
<td>2.95</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Pizza</td>
<td>RF034</td>
<td>4.90</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Chicken</td>
<td>RF035</td>
<td>4.57</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Tomato</td>
<td>RF036</td>
<td>6.85</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Tomato2</td>
<td>RF038</td>
<td>17.15</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Tomato3</td>
<td>RF039</td>
<td>4.90</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Item</td>
<td>Code</td>
<td>Score</td>
<td>Category</td>
<td>Status</td>
</tr>
<tr>
<td>-------------</td>
<td>-------</td>
<td>-------</td>
<td>----------------</td>
<td>---------</td>
</tr>
<tr>
<td>Egg</td>
<td>RF040</td>
<td>5.79</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Plum</td>
<td>RF041</td>
<td>10.07</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Vegetables</td>
<td>RF042</td>
<td>2.95</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Sweet Potato</td>
<td>RF043</td>
<td>5.80</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Pumpkin</td>
<td>RF044</td>
<td>4.08</td>
<td>Rotten Food</td>
<td>Negative</td>
</tr>
<tr>
<td>Watermelon</td>
<td>NF005</td>
<td>80.05</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Orange1</td>
<td>NF008</td>
<td>74.09</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Orange2</td>
<td>NF009</td>
<td>75.09</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Banana</td>
<td>NF013</td>
<td>78.93</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Cherry1</td>
<td>NF027</td>
<td>79.93</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Cherry2</td>
<td>NF028</td>
<td>81.36</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Coconut</td>
<td>NF030</td>
<td>76.24</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Strawberry1</td>
<td>NF037</td>
<td>77.85</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Strawberry2</td>
<td>NF038</td>
<td>86.04</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Lime</td>
<td>NF048</td>
<td>77.45</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Lemon</td>
<td>NF049</td>
<td>77.81</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Apple1</td>
<td>NF055</td>
<td>75.38</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Apple2</td>
<td>NF056</td>
<td>80.58</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Apple3</td>
<td>NF057</td>
<td>76.04</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Melon</td>
<td>NF061</td>
<td>77.82</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Honey</td>
<td>NF062</td>
<td>78.00</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Peach1</td>
<td>NF079</td>
<td>81.31</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Peach2</td>
<td>NF080</td>
<td>73.45</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Grapefruit</td>
<td>NF085</td>
<td>78.70</td>
<td>Natural Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Baguette1</td>
<td>TF006</td>
<td>78.59</td>
<td>Transformed Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Baguette2</td>
<td>TF007</td>
<td>79.21</td>
<td>Transformed Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Cookie</td>
<td>TF012</td>
<td>79.21</td>
<td>Transformed Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Donut</td>
<td>TF014</td>
<td>78.29</td>
<td>Transformed Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Chocolate1</td>
<td>TF033</td>
<td>77.13</td>
<td>Transformed Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Chocolate2</td>
<td>TF034</td>
<td>84.68</td>
<td>Transformed Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Food</td>
<td>Code</td>
<td>Score</td>
<td>Classification</td>
<td>Result</td>
</tr>
<tr>
<td>----------------</td>
<td>-------</td>
<td>-------</td>
<td>---------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Chocolate3</td>
<td>TF035</td>
<td>79.92</td>
<td>Transformed Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Croissant</td>
<td>TF037</td>
<td>75.50</td>
<td>Transformed Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Cupcake</td>
<td>TF050</td>
<td>79.80</td>
<td>Transformed Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Ice Cream1</td>
<td>TF062</td>
<td>81.59</td>
<td>Transformed Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Ice Cream2</td>
<td>TF063</td>
<td>74.43</td>
<td>Transformed Food</td>
<td>Positive</td>
</tr>
<tr>
<td>Fruit Pastry</td>
<td>TF093</td>
<td>82.96</td>
<td>Transformed Food</td>
<td>Positive</td>
</tr>
</tbody>
</table>
C.2 List of words used in the GNAT task

Good words
Rainbow
Glorious
Sunshine
Laughter
Peace
Heaven
Excellent
Happy
Glee
Joyful
Terrific
Superb

Bad words
Agony
Hell
Ugly
Angry
Destroy
Awful
Nasty
Sad
Evil
Dirty
Hate
Miserable
C.3 Example Functions from intermediate food responses

Below are some examples of participants’ functions plotted for the intermediate food items. The X-axis represents the intermediate food level, and the Y-axis represents reported pleasantness. Measures of both slope and PSE suggest that participants could not differentiate particularly well between the intermediate levels, and as such function data was not used during the analysis.

<table>
<thead>
<tr>
<th>Session 1</th>
<th>Session 2</th>
<th>Combined Plot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pleasant Cue</td>
<td>Unpleasant Cue</td>
<td>Pleasant Cue</td>
</tr>
<tr>
<td>PSE</td>
<td>Slope</td>
<td>PSE</td>
</tr>
<tr>
<td>2.852594795</td>
<td>0.183161377</td>
<td>-1.4290.2339</td>
</tr>
<tr>
<td>-0.295298194</td>
<td>0.089476725</td>
<td>-3.700.170637</td>
</tr>
<tr>
<td>-1.991346598</td>
<td>0.050067292</td>
<td>-9.747.525999</td>
</tr>
<tr>
<td>0.682407062</td>
<td>0.14898342</td>
<td>0.682407062</td>
</tr>
</tbody>
</table>
APPENDIX D

D.1 Preliminary Study 6a: Confirming the preference for curvature

Rationale

The studies reported in chapter 6 were conceived based upon the findings of previous studies which have indicated that humans prefer curvature over angularity. The stimuli used in the studies in chapter 6 were adapted from a series of abstract patterns used by Bar and Neta (2006). In the original study participants had shown a preference for smoother abstract patterns over sharper abstract patterns. Preliminary study 6a was conducted prior to the studies in chapter 6 to first investigate whether this preference for curvature could be replicated.

Methods

Participants

Fifteen participants aged between 22 and 37 (10 females, mean age = 29.87, SD = 4.87) were recruited for the study from the staff and post-graduate population at the University of Manchester. Participants were not compensated for their participation in the study.

Stimuli and apparatus

The stimuli for the task were the 120 of the abstract patterns originally created by Bar and Neta (2006). Each pattern had a smooth and a sharp variant (240 images total), and the smooth and sharp variants had been matched for all features except contour. All patterns were presented against a grey background and were scaled to a size of 256 by 256 pixels. Stimuli were presented using a desktop computer running PsychoPy software (Peirce, 2007) and a 21-inch LCD monitor at a resolution of 1920 by 1080 pixels. All responses were recorded using a standard keyboard placed on a desk directly in front of the participant.

Design and Procedure

Participants were informed that they would be viewing a series of abstract patterns, and that they should state whether they liked or disliked each pattern based on their initial ‘gut’ reaction. Each participant viewed the sharp variants of 60 patterns and the smooth variants of the remaining 60 patterns. Whether a pattern was shown in its smooth or sharp form was counterbalanced across participants. Each pattern was shown for 84ms, after each presentation a ‘like or dislike?’ prompt appeared, and participants were given 2000ms to provide a response.
using two keys on the keyboard. In total the task consisted of 120 trials and smooth and sharp patterns were presented in random order.

**Data Analysis**

For each pattern type (smooth or sharp), liking was calculated as the proportion of ‘liked’ responses from the total number of presentations of that type. Trials where no response was provided were not counted as liked or disliked. The proportion of liked and disliked responses for each pattern type was compared using paired samples t-tests. Average reaction times for each pattern type were calculated for each participant, and were compared using paired samples t-tests. An alpha level of .05 has been used for all statistical tests.

**Results**

Participants liked the smooth patterns significantly more than the sharp patterns $t(14) = 2.44$, $p = .029$ (Liking for smooth objects: $M = 50.44\%$, $SD = 14.98$; liking for sharp objects: $M = 38.13\%$, $SD = 18.94$). No significant difference in overall reaction times was observed between the smooth and sharp objects $t(14) = .18$, $p = .858$ (Response time for smooth objects: $M = 529.80$ms, $SD = 148.19$ms; response time for sharp objects: $M = 524.80$ms, $SD = 134.11$ms).
D.2 Preliminary Study 6b: Pattern differentiation task

Rationale

The tasks reported in chapter 6 all used an expanded set of smooth and sharp patterns based upon those used by Bar and Neta (2006). The sharp patterns were adapted using a smoothing procedure to create a series of intermediate patterns of increasing smoothness. The purpose of preliminary study 6b was to ascertain whether participants could detect the relatively small difference in smoothness between each intermediate level.

Methods

Participants

Eighteen participants aged between 22 and 35 (11 females, mean age = 28.38, SD = 3.73) were recruited for the study from the staff and post-graduate population at the University of Manchester. Participants were not compensated for their participation in the study.

Stimuli and apparatus

Thirty greyscale patterns were selected from the sharp pattern set used by Bar and Neta (2006). For each pattern, 7 different variants were used in the task: The original unaltered sharp pattern, a smoothed version of the pattern, and 5 intermediate versions of increasing smoothness which existed at several increments (16.66%, 33.33%, 50%, 66.66% and 83.33%) between the smooth and sharp variants. The smoothed and intermediate patterns were created using the smoothing function in Paintshop Pro Photo X2 (Corel Corporation, Ottawa, Ontario, Canada), and all patterns were scaled to a size of 256 by 256 pixels. Stimuli were presented using a desktop computer running Psychopy software (Peirce, 2007) and a 21-inch LCD monitor at a resolution of 1920 by 1080 pixels. All responses were recorded using a standard keyboard and a wired mouse placed on a desk directly in front of the participant.

Design and Procedure

Participants were told that they would be shown a series of intermediate patterns of varying smoothness, and that their task would be to state whether the intermediate pattern (labelled pattern C) more closely resembled a ‘very sharp’ or a ‘very smooth’ (labelled pattern A and B respectively) version of that same pattern. In each trial pattern C was presented at the top of screen, and pattern A and pattern B were presented on the left and right hand sides of the screen. Responses were collected using a VAS situated in the centre of the screen, and participants selected the appropriate response using the mouse (see Figure D2.1). The VAS
offered 5 possible responses to participants: The same as pattern A, very similar to pattern A, equally similar to pattern A and pattern B, very similar to pattern B, the same as pattern B.

Each trial began with a fixation cross presented for 2000ms. When the response screen appeared participants were given 10000ms to respond, and the time remaining in a trial was indicated by a countdown timer. The task consisted of 150 trials with each of the 30 patterns being shown once for each of the 5 intermediate versions.

**Figure D2.1**  *Typical trial design*

**Data Analysis**

Data were analysed using a one way repeated-measures ANOVA with intermediate presentation type (15%, 30%, 50%, 65% and 80%) as the within subjects factor. The dependant variable was the average reported similarity to pattern B. Responses from each of the trials were coded relative to pattern B in the following manner: The same as pattern A (0% similarity to B), very similar to pattern A (25% similarity to B), equally like pattern A and pattern B (50% similarity to B), very similar to pattern B (75% similarity to B), the same as pattern B (100% similarity to B). Trials where no response was provided were excluded from the analysis.

In cases where Mauchly’s test revealed that the assumption of sphericity had been violated, the degrees of freedom have been corrected using Greenhouse-Geisser estimates of sphericity. An alpha level of .05 has been used for all statistical tests, and effect sizes are reported here as
partial eta squared ($\eta^2$) where 0.05, 0.1, and 0.2 correspond to small, medium, and large effects, respectively (Cohen, 1988).

**Results**

A main effect of presentation type was observed; intermediate patterns appeared increasingly similar to the very smooth pattern as the intermediate pattern became smoother $F(2, 32) = 295.74, p < .001, \eta^2 = .946$. Pairwise comparisons revealed a significant increase in the reported similarity to B at each level of presentation type over the preceding level (see Figure F3.2; all $p < .001$). This result suggests that, given enough time, participants could successfully differentiate between the intermediate levels that were included in the task.

![Figure D2.2](image)

*Figure D2.2  Reponses to intermediate levels in pattern differentiation task. Error bars represent standard errors of the mean.*
D.3 Sharp Patterns used in pattern judgement tasks

<table>
<thead>
<tr>
<th>PATTERN NUMBER</th>
<th>PATTERN NUMBER</th>
<th>PATTERN NUMBER</th>
<th>PATTERN NUMBER</th>
<th>PATTERN NUMBER</th>
<th>PATTERN NUMBER</th>
<th>PATTERN NUMBER</th>
<th>PATTERN NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>16</td>
<td>20</td>
<td>21</td>
<td>32</td>
<td>36</td>
<td>37</td>
</tr>
<tr>
<td><img src="image1.png" alt="Pattern1" /></td>
<td><img src="image4.png" alt="Pattern4" /></td>
<td><img src="image16.png" alt="Pattern16" /></td>
<td><img src="image20.png" alt="Pattern20" /></td>
<td><img src="image21.png" alt="Pattern21" /></td>
<td><img src="image32.png" alt="Pattern32" /></td>
<td><img src="image36.png" alt="Pattern36" /></td>
<td><img src="image37.png" alt="Pattern37" /></td>
</tr>
<tr>
<td>38</td>
<td>39</td>
<td>40</td>
<td>41</td>
<td>42</td>
<td>43</td>
<td>44</td>
<td>50</td>
</tr>
<tr>
<td><img src="image38.png" alt="Pattern38" /></td>
<td><img src="image39.png" alt="Pattern39" /></td>
<td><img src="image40.png" alt="Pattern40" /></td>
<td><img src="image41.png" alt="Pattern41" /></td>
<td><img src="image42.png" alt="Pattern42" /></td>
<td><img src="image43.png" alt="Pattern43" /></td>
<td><img src="image44.png" alt="Pattern44" /></td>
<td><img src="image50.png" alt="Pattern50" /></td>
</tr>
<tr>
<td>63</td>
<td>66</td>
<td>67</td>
<td>84</td>
<td>89</td>
<td>95</td>
<td>97</td>
<td>103</td>
</tr>
<tr>
<td><img src="image63.png" alt="Pattern63" /></td>
<td><img src="image66.png" alt="Pattern66" /></td>
<td><img src="image67.png" alt="Pattern67" /></td>
<td><img src="image84.png" alt="Pattern84" /></td>
<td><img src="image89.png" alt="Pattern89" /></td>
<td><img src="image95.png" alt="Pattern95" /></td>
<td><img src="image97.png" alt="Pattern97" /></td>
<td><img src="image103.png" alt="Pattern103" /></td>
</tr>
<tr>
<td>113</td>
<td>115</td>
<td>120</td>
<td>122</td>
<td>123</td>
<td>124</td>
<td>125</td>
<td>126</td>
</tr>
<tr>
<td><img src="image113.png" alt="Pattern113" /></td>
<td><img src="image115.png" alt="Pattern115" /></td>
<td><img src="image120.png" alt="Pattern120" /></td>
<td><img src="image122.png" alt="Pattern122" /></td>
<td><img src="image123.png" alt="Pattern123" /></td>
<td><img src="image124.png" alt="Pattern124" /></td>
<td><img src="image125.png" alt="Pattern125" /></td>
<td><img src="image126.png" alt="Pattern126" /></td>
</tr>
<tr>
<td>127</td>
<td>128</td>
<td>129</td>
<td>130</td>
<td>131</td>
<td>132</td>
<td>133</td>
<td>134</td>
</tr>
<tr>
<td><img src="image127.png" alt="Pattern127" /></td>
<td><img src="image128.png" alt="Pattern128" /></td>
<td><img src="image129.png" alt="Pattern129" /></td>
<td><img src="image130.png" alt="Pattern130" /></td>
<td><img src="image131.png" alt="Pattern131" /></td>
<td><img src="image132.png" alt="Pattern132" /></td>
<td><img src="image133.png" alt="Pattern133" /></td>
<td><img src="image134.png" alt="Pattern134" /></td>
</tr>
</tbody>
</table>

Note: Numbers denote unique identifiers of sharp patterns reported in Bar and Neta (2006).
D.4 List of words used in the IAT task portion of pattern judgement task

**Good words**

Rainbow
Glorious
Sunshine
Laughter
Peace
Heaven
Excellent
Happy
Glee
Joyful

**Bad words**

Agony
Hell
Ugly
Angry
Destroy
Awful
Nasty
Sad
Evil
Dirty
D.5 Excluded participants from Experiment 6

We could not fit adequate functions to the responses from 6 participants during experiment 6. In each case, the plot on the left shows responses from session 1, the centre plot shows responses from session 2, and the plot on the right shows the responses from both sessions on a single plot.

Cued participants
Non-cued participants
D.6 Graphs of average liking ratings for each group in experiment 6

Session 1 - Cued

Session 1 - Non Cued
### E.1 Trial breakdown for session 2 and 3 (congruent trials)

*Note: Colours associated with each distribution are changed for readability purposes.*

<table>
<thead>
<tr>
<th>Short Interval Duration</th>
<th>trials congruent Cue</th>
<th>Intermid Interval Duration</th>
<th>trials congruent cue</th>
<th>long</th>
<th>Duration Congruent</th>
<th>trials congruent Cue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>494</td>
<td>13</td>
<td>1</td>
<td>776.4</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>529.3</td>
<td>13</td>
<td>2</td>
<td>811.7</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>564.6</td>
<td>13</td>
<td>3</td>
<td>847</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>599.9</td>
<td>13</td>
<td>4</td>
<td>882.3</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>635.2</td>
<td>13</td>
<td>5</td>
<td>917.6</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>670.5</td>
<td>13</td>
<td>6</td>
<td>952.9</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>705.8</td>
<td>13</td>
<td>7</td>
<td>988.2</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>741.1</td>
<td>13</td>
<td>8</td>
<td>1023.5</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>776.4</td>
<td>13</td>
<td>9</td>
<td>1058.8</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>811.7</td>
<td>13</td>
<td>10</td>
<td>1094.1</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>847</td>
<td>13</td>
<td>11</td>
<td>1129.4</td>
<td>13</td>
<td></td>
</tr>
</tbody>
</table>

143 trials

| 1                      | 1058.8              | 13                          | 1                    | 1058.8| 13                  |
| 2                      | 1094.1              | 13                          | 2                    | 1094.1| 13                  |
| 3                      | 1129.4              | 13                          | 3                    | 1129.4| 13                  |
| 4                      | 1164.7              | 13                          | 4                    | 1164.7| 13                  |
| 5                      | 1200                | 13                          | 5                    | 1200  | 13                  |
| 6                      | 1235.3              | 13                          | 6                    | 1235.3| 13                  |

143 trials

| 7                      | 1270.6              | 13                          | 7                    | 1270.6| 13                  |
| 8                      | 1305.9              | 13                          | 8                    | 1305.9| 13                  |
| 9                      | 1341.2              | 13                          | 9                    | 1341.2| 13                  |
| 10                     | 1376.5              | 13                          | 10                   | 1376.5| 13                  |
| 11                     | 1411.8              | 13                          | 11                   | 1411.8| 13                  |

143 trials
### E.2 Trial breakdown for session 3 (minor discrepancy trials)

*Note: Colours associated with each distribution are changed for readability purposes.*

<table>
<thead>
<tr>
<th>Short Interv</th>
<th>trials congruent Cue</th>
<th>trials incongruent cue</th>
<th>Intermid Interv</th>
<th>trials congruent cue</th>
<th>trials incongruent cue</th>
<th>long Interv</th>
<th>Duration</th>
<th>trials congruent Cue</th>
<th>trials incongruent cue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>494</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>529.3</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>564.6</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>589.9</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>635.2</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>670.5</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>705.8</td>
<td>10</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>741.1</td>
<td>10</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>776.4</td>
<td>10</td>
<td>10</td>
<td>776.4</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>811.7</td>
<td>10</td>
<td>10</td>
<td>811.7</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>847</td>
<td>10</td>
<td></td>
<td>847</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>882.3</td>
<td>10</td>
<td>10</td>
<td>882.3</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>912.6</td>
<td>10</td>
<td>10</td>
<td>912.6</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>952.9</td>
<td>20</td>
<td>10</td>
<td>952.9</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>988.2</td>
<td>10</td>
<td>10</td>
<td>988.2</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1023.5</td>
<td>10</td>
<td>10</td>
<td>1023.5</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1058.8</td>
<td>10</td>
<td>10</td>
<td>1058.8</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1094.1</td>
<td>10</td>
<td>10</td>
<td>1094.1</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1129.4</td>
<td>10</td>
<td>10</td>
<td>1129.4</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1164.7</td>
<td>10</td>
<td>10</td>
<td>1164.7</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1200</td>
<td>10</td>
<td>10</td>
<td>1200</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1235.3</td>
<td>10</td>
<td>10</td>
<td>1235.3</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>180 trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1270.6</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1305.9</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1341.2</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1376.5</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1411.8</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

180 trials
E.3 Trial breakdown for session 3 (extreme discrepancy trials)

Note: Colours associated with each distribution are changed for readability purposes.

<table>
<thead>
<tr>
<th>Short</th>
<th>Interval Duration</th>
<th>trials congruent cue</th>
<th>trials incongruent cue</th>
<th>Intermed Interval Duration</th>
<th>trials congruent cue</th>
<th>long Duration Congruent</th>
<th>trials congruent cue</th>
<th>trials incongruent cue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>494</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>776.4</td>
<td>3</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>529.3</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>811.7</td>
<td>3</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>564.6</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>847</td>
<td>3</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>599.9</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>882.3</td>
<td>3</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>635.2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>670.5</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>705.8</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>741.1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>776.4</td>
<td>3</td>
<td></td>
<td>1</td>
<td>776.4</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>811.7</td>
<td>3</td>
<td></td>
<td>2</td>
<td>811.7</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>847</td>
<td>3</td>
<td></td>
<td>3</td>
<td>847</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>61 trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>917.6</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>952.9</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>988.2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1023.5</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1058.8</td>
<td>3</td>
<td></td>
<td>1</td>
<td>1058.8</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1094.1</td>
<td>3</td>
<td></td>
<td>2</td>
<td>1094.1</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>33 trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1129.4</td>
<td>3</td>
<td></td>
<td>3</td>
<td>1129.4</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1164.7</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1200</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1235.3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1270.6</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1305.9</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1341.2</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1376.5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>61 trials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

61 trials