' Habit-like' attention: The effect of outcome revaluation on attentional prioritisation of reward-related stimuli

Author: Le, Jenny

Publication Date: 2023

DOI: https://doi.org/10.26190/unsworks/25116

License: https://creativecommons.org/licenses/by/4.0/

Link to license to see what you are allowed to do with this resource.

Downloaded from http://hdl.handle.net/1959.4/101409 in https://unsworks.unsw.edu.au on 2024-03-08
‘Habit-like’ attention: The effect of outcome revaluation on attentional prioritisation of reward-related stimuli

Jenny Le

Primary Supervisor: Professor Mike E. Le Pelley
Joint Supervisor: Dr Poppy Watson
Secondary Supervisor: Professor Thomas J. Whitford

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy/Master of Psychology (Clinical)

School of Psychology
Faculty of Science
2023
Thesis submission for the degree of Doctor of Philosophy

ORIGINALITY STATEMENT

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project’s design and conception or in style, presentation and linguistic expression is acknowledged.

COPYRIGHT STATEMENT

I hereby grant the University of New South Wales or its agents a non-exclusive licence to archive and to make available (including to members of the public) my thesis or dissertation in whole or part in the University libraries in all forms of media, now or hereafter known. I acknowledge that I retain all intellectual property rights which subsist in my thesis or dissertation, such as copyright and patent rights, subject to applicable law. I also retain the right to use all or part of my thesis or dissertation in future works (such as articles or books).

For any substantial portions of copyright material used in this thesis, written permission for use has been obtained, or the copyright material is removed from the final public version of the thesis.

AUTHENTICITY STATEMENT

I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis.
**Thesis submission for the degree of Doctor of Philosophy**

<table>
<thead>
<tr>
<th>Thesis Title and Abstract</th>
<th>Declarations</th>
<th>Inclusion of Publications Statement</th>
<th>Corrected Thesis and Responses</th>
</tr>
</thead>
</table>

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in the candidate’s thesis in lieu of a Chapter provided:

- The candidate contributed **greater than 50%** of the content in the publication and are the "primary author", i.e. they were responsible primarily for the planning, execution and preparation of the work for publication.
- The candidate has obtained approval to include the publication in their thesis in lieu of a Chapter from their Supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis.

The candidate has declared that **their thesis contains no publications, either published or submitted for publication.**

**Candidate’s Declaration**

I declare that I have complied with the Thesis Examination Procedure.
Acknowledgements

When finalising my thesis, I did not think that the acknowledgements would be one of the hardest sections to write. It is bittersweet to accept that the whirlwind of my PhD candidature almost over, and as I write this, it is dawns on me how many people have been a part of my journey.

First and foremost, I want to thank my primary supervisor, Mike Le Pelley. When I started working with you as a third-year internship student, I knew next to nothing about research. Looking back now over the last six years, I feel so much gratitude for how much you have helped me grow. Thank you for teaching me how to research: from reading and thinking, to coding and writing. More importantly, thank you for showing me why we research: you have instilled in me a passion for discovery and a love of learning that will stay with me for a lifetime. Thank you for your unwavering support, for always lending a listening ear, for keeping your door open for me no matter how busy you were, for the countless hours you spent talking through the hard questions with me, for your many delicious recipes, and for all the laughs we have shared. You have been endlessly patient, kind, and generous with me, and I feel incredibly fortunate to have had you as my supervisor.

Thank you to my joint supervisor, Poppy Watson. You have been such a huge inspiration to me since I met you many years ago. Thank you for all the talks and wisdom on my thesis, for always being a comforting face in the crowd of many nerve-racking talks, for always checking in on me, for refining my music taste, and for simply being an amazing person at your core. You are such a warm-hearted and thoughtful supervisor, and it has been such a privilege to share so many fun memories with you over these last few years. I know that many students will be so lucky to work with you in the years to come.

Thank you to my friends and colleagues from the Human Learning/Whitford Lab, past and present, particularly David Ng, Sandra Lagator, Daniel Pearson, Peter Lovibond, Jess Lee, Julie Chow, Saoirse Connor Desai, and Tom Whitford. Thank you for the support, advice, and companionship over these last few years. Thank you to Kellie Williamson, Wendy Chen, and Emma Bryant, who have always been a shoulder to lean on during the Clinical/PhD program.
I also want to express my gratitude for my closest friends and family who have always offered so much support and encouragement, not only in the last few years, but throughout my whole academic life.

Last but not least, I want to thank Richie. Thank you for your endless cheerleading, understanding, and love. I am so blessed to have shared the highest and lowest moments of the last few years with you, and I am so excited for our next chapter.
Abstract

Stimuli associated with large rewards are more likely to be prioritised by our attention than stimuli associated with small rewards – a phenomenon known as value-modulated attentional capture (VMAC). Importantly, VMAC appears to occur independently of, and indeed contrary to, top-down goals and intentions. Viewed another way, reward-related stimuli can be said to automatically elicit a conditioned attentional prioritisation response. Conceptualising attention as a response that can be conditioned through learned associations with reward raises a question that is central to this thesis: does attentional prioritisation of reward-related stimuli exhibit features of a ‘habit’? To answer this question, this thesis investigated: (1) the conditions under which attentional prioritisation of reward-related stimuli is sensitive to changes in the values of outcomes; (2) whether the amount of training has an influence on the extent to which attentional prioritisation will be ‘habit-like’ in its persistence; and (3) whether attentional prioritisation of reward-related stimuli results in ‘habit-like’ delayed disengagement. Findings in Chapter 2 revealed that attentional prioritisation of high- and low-value stimuli was insensitive to devaluation (of a previously high-value outcome) and super-valuation (of a previously low-value outcome), but partially sensitive to a value-reversal manipulation. These findings suggest that the particular revaluation procedure used will influence the extent to which participants allow learned patterns of attention to persist, or instead engage in effortful updating of attentional control settings. Given this evidence of a habit-like component of attention, Chapter 3 investigated whether the persistence of attentional patterns would vary as a function of moderate versus extended training. Finally, Chapter 4 revealed that habit-like attentional prioritisation of reward-related stimuli may not only have an influence on the extent to which a stimulus captures attention, but also the extent to which it holds attention following devaluation and value-reversal. Taken together, the findings of this thesis suggest that attentional prioritisation may indeed show many features of a habit-like response. In turn, our findings provide a novel framework to further understand the nature of the process by which stimuli elicit attentional prioritisation, and how ‘habit-like’ outcome-independent prioritisation may play a role in shaping overt behaviour.
Contents

1 General Introduction ............................................................................................................. 1

1.1 Influences on Attentional Priority ..................................................................................... 2

1.1.1 Top-Down versus Bottom-Up Influences on Attentional Priority ......................... 2

1.1.2 Influence of Previous Experiences on Attentional Priority ......................................... 3

1.2 Value-Modulated Attentional Capture (VMAC) ............................................................. 3

1.2.1 The Two-Phase VMAC Procedure .............................................................................. 4

1.2.2 The Single-Phase VMAC Procedure .......................................................................... 6

1.2.3 Clinical Implications of VMAC .................................................................................. 9

1.2.4 Section Summary ....................................................................................................... 10

1.3 Attentional Prioritisation as a ‘Habit-Like’ Response ...................................................... 11

1.3.1 Defining Habitual versus Goal-Directed Behaviours ................................................. 12

1.3.2 The Habit Theory of Addiction .................................................................................. 14

1.3.3 Translating Outcome Devaluation in Humans ........................................................... 14

1.3.4 ‘Habit-like’ Attention ............................................................................................. 16

1.4 Research Presented in this Thesis .................................................................................... 19

2 The Effect of Outcome Revaluation on Attentional Prioritisation of Reward-Related
Stimuli ........................................................................................................................................ 21

2.1 Introduction ..................................................................................................................... 21

2.2 Experiment 1 ................................................................................................................... 24

2.2.1 Methods and Materials ............................................................................................ 28

2.2.2 Results ..................................................................................................................... 34

2.2.3 Discussion ............................................................................................................... 39

2.3 Experiment 2 ................................................................................................................... 41

2.3.1 Methods and Materials ............................................................................................ 42

2.3.2 Results ..................................................................................................................... 43
2.3.3 Discussion ................................................................. 47

2.4 Experiment 3 ................................................................... 49
  2.4.1 Method .................................................................... 50
  2.4.2 Results .................................................................... 52
  2.4.3 Discussion ................................................................. 58

2.5 Experiment 4 ................................................................... 60
  2.5.1 Method .................................................................... 64
  2.5.2 Results .................................................................... 68
  2.5.3 Discussion ................................................................. 73

2.6 Chapter Discussion ........................................................... 76

2.7 Conclusions .................................................................. 79

3 The Effect of Moderate Versus Extended Training on Attentional Prioritisation of Reward-Related Stimuli ................................................................. 80

3.1 Introduction .................................................................. 80

3.2 Experiment 5 ................................................................... 83
  3.2.1 Method and Materials .................................................. 86
  3.2.2 Results ..................................................................... 93
  3.2.3 Discussion ................................................................. 100

3.3 Chapter Conclusions .......................................................... 105

4 The Effect of Outcome Revaluation on Attentional Disengagement from Reward-Related Stimuli ................................................................. 106

4.1 Introduction .................................................................. 106

4.2 Experiment 6 ................................................................... 111
  4.2.1 Method and Materials .................................................. 114
  4.2.2 Results ..................................................................... 120
  4.2.3 Discussion ................................................................. 124
5 General Discussion

5.1 Summary of Findings

5.1.1 Under what conditions is attentional prioritisation of reward-related stimuli sensitive to changes in the value of the outcome?

5.1.2 Does the amount of training have an influence on the extent to which attentional prioritisation will be ‘habit-like’ in its persistence?

5.1.3 Does attentional prioritisation of reward-related stimuli result in ‘habit-like’ delayed disengagement?

5.1.4 ‘Habit-like’ Attentional Prioritisation and Sign-Tracking

5.1.5 Overall Summary

5.2 Implications of the Current Findings

5.2.1 Attentional Economics

5.2.2 Clinical Implications

5.3 Limitations and Future Directions

5.3.1 Habits as Null Effects

5.3.2 Individual Differences in Sensitivity to Outcome Revaluation

5.3.3 Knowledge Tests versus Consumption Tests

5.3.4 Instrumental Attentional Habits

5.3.5 Assessing Habitual Attention Using the Belief Criterion

5.3.6 Impact of ‘Habit-Like’ Attention on Choice

5.3.7 ‘Habit-Like’ Attention in Clinical Populations

5.4 Concluding Remarks

References
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.1</td>
<td>Design of each phase in Experiment 1</td>
<td>31</td>
</tr>
<tr>
<td>Table 2.2</td>
<td>Design of each phase in Experiment 2</td>
<td>43</td>
</tr>
<tr>
<td>Table 2.3</td>
<td>Design of each phase in Experiment 3</td>
<td>51</td>
</tr>
<tr>
<td>Table 2.4</td>
<td>Design of each phase in Experiment 4</td>
<td>67</td>
</tr>
<tr>
<td>Table 3.1</td>
<td>Design of each phase in Experiment 5</td>
<td>90</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>Results in each condition during the training phase of Experiment 5</td>
<td>95</td>
</tr>
<tr>
<td>Table 3.3</td>
<td>Results in each condition during the test phase of Experiment 5</td>
<td>98</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Design of each phase in Experiment 6</td>
<td>118</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1.1  Examples of the stimulus presentations in the two-phase procedure by Anderson et al. (2011b) ................................................................. 5
Figure 1.2  Example trials of a single-phase visual search task .................. 7
Figure 2.1  Examples of stimulus presentation in Experiment 1 ............... 29
Figure 2.2  Distraction trials and saccade latency data in the training phase of Experiment 1 .......................................................... 35
Figure 2.3  Distraction trials and saccade latency data in the test phase of Experiment 1 .......................................................... 37
Figure 2.4  Distraction trials and saccade latency data in the training phase of Experiment 2 .......................................................... 44
Figure 2.5  Distraction trials and saccade latency data in the test phase of Experiment 2 .......................................................... 46
Figure 2.6  Distraction trials and saccade latency data in the training phase of Experiment 3 .......................................................... 54
Figure 2.7  Distraction trials and saccade latency data in the test phase of Experiment 3 .......................................................... 56
Figure 2.8  Example of trial events in the RT version of the attentional revaluation task .......................................................... 65
Figure 2.9  Response time and error data in the training phase of Experiment 4 ...... 69
Figure 2.10 Response time and error data in the test phase of Experiment 4 ........... 72
Figure 3.1  Examples of stimulus presentation in Experiment 5 .................. 87
Figure 3.2  Response time and error data in the training phase of Experiment 5 .... 94
Figure 3.3  Response time and error data in the test phase of Experiment 5 ........ 97
Figure 4.1  Illustration of the attentional disengagement procedure used by Watson et al. (2020) .......................................................... 109
Figure 4.2  Examples of stimulus presentation in Experiment 6 .................. 115
Figure 4.3  Saccade latency data in the training phase of Experiment 6 ........... 121
Figure 4.4  Saccade latency data in the test phase of Experiment 6 ............... 123
Chapter 1

General Introduction

The world we live in is full of sensory information; bright lights, loud noises, and rapid movements surround us at our every turn. However, our brain does not have a limitless processing capacity, meaning that we could not bring our full perceptual and cognitive analytics to bear on every element of information impinging on our senses. Instead, our cognitive system enacts processes of *selective attention* to filter incoming information, with the aim of prioritising stimuli that are important for further processing, and ignoring those that are deemed irrelevant or distracting (Desimone & Duncan, 1995).

Critically, by acting as a central gateway through which we perceive the world, what we attend to will ultimately have knock-on effects for our behaviours. Almost every step of our lives can be seen as a decision — from the banal (What should I wear to work? What should I have for lunch?) to those with far-reaching consequences (Should I quit smoking?). But these overt choices are simply the end point of a decision-making process — a process that begins with paying attention to what our options are (Gluth et al., 2018; Pearson et al., 2022).

For example, when given a day off work, it often feels as if there are endless possibilities for how the day may take shape. However, when looking around our home, stimuli that are prioritised by our attention (e.g., our running shoes, or a cigarette) are likely to feature in our set of behavioural options, and hence become more likely to be the chosen as targets of our behaviour (to run, or to smoke). Conversely, stimuli that are not prioritised by our attention are less likely to be targets of action. By shaping the process by which information is gathered, attention has the potential to play a critical role in the first steps in many of our decisions — and hence to mould our overt behaviours, the end points of these decision-making cascades (Krajbich, 2019; Krajbich & Rangel, 2011).

We typically strive to make consequential decisions in a *goal-directed* way — to choose options and perform behaviours that are targeted at achieving our current goals and obtaining valued outcomes (praise from others, satiety, health and longevity; Balleine & O’Doherty, 2010). But despite being highly goal-directed creatures, we still sometimes feel that we have behaved in ways we did not mean to. I may walk into my kitchen and open the
fridge even if I am not hungry, or an abstaining smoker may accept a cigarette when offered and light it up. An influential view posits that these ‘slips’ in behaviour represent habits that are potentiated by stimuli in the environment, despite diminished desire for the outcomes at hand (i.e., food in my fridge, or the cigarette; Dezfuli & Balleine, 2012). Interestingly, recent evidence suggests that habits are not restricted to overt behavioural choices: attention also seems to show some features of ‘habits’. In this thesis I will present evidence to suggest that evidence of habit-like processes may indeed be found early on in the decision-making process – at the level of attention. In the current chapter, I will introduce previous research on the modulators of attentional priority, and discuss how viewing attention as a conditionable ‘habit’ can offer an alternative way of understanding habitual processes.

1.1 Influences on Attentional Priority

1.1.1 Top-Down versus Bottom-Up Influences on Attentional Priority

I begin by introducing some key concepts in the area of attentional priority. Until recently, traditional models of attentional control tended to distinguish between two processes that operate to influence attentional selection. First, our attention can be influenced by top-down control, which acts to prioritise stimuli that are relevant to achieving our current goals and obtaining valued outcomes (also known as goal-directed attention; Posner, 2016; Theeuwes, 2010; Wolfe & Horowitz, 2017). For example, when driving on our way to work, top-down attentional control would prioritise road signs and other cars that are important to our goals of arriving at our destination safely, on time, without getting lost. Likewise, while writing this thesis, it was imperative that I ignored distracting objects in my environment and instead strategically oriented my attention to my work, in line with my goal of finishing my thesis and obtaining my doctorate.

In other instances, however, our attention may also operate in a bottom-up manner when it is ‘captured’ by physically salient stimuli in our environment (also known as stimulus-driven attention). Physical salience here refers to the distinctiveness of an item’s physical properties. For example, a stimulus may be physically salient based on its dissimilarity from other items (e.g., a single red flower among green leaves) or its intensity (brightness, abruptness, loudness; Jonides & Yantis, 1988; Theeuwes, 1993). Importantly, bottom-up attentional prioritisation of physically salient stimuli operates rapidly, and
independently of our goals. Consequently, when physically salient stimuli are inconsistent with our current goals, stimulus-driven attentional capture can have negative consequences. For example, while driving a bright flash in the rear-vision mirror may distract us from our goal of focusing on the road ahead. On the other hand, knowledge of these features of bottom-up attentional capture has been harnessed in the design of signals that promote adaptive behaviours: bright ‘School Zone’ signs remind us to slow down, and seeing the flashing lights of an ambulance allows us to pull over early.

1.1.2 Influence of Previous Experiences on Attentional Priority

Recently, a growing body of work has brought this traditional dichotomy of top-down versus bottom-up influences on attentional priority into question, by demonstrating a further influence on attentional prioritisation based on selection history that operates independently of both physical salience and an observer’s current goals (Anderson et al., 2021; Awh et al., 2012; Failing & Theeuwes, 2018; Le Pelley et al., 2016). Selection history refers to a range of factors that relate to an observer’s prior experience with stimuli, and with the consequences of selecting those stimuli. For example, prior experience of attending to a stimulus as a target of search can lead to persistent prioritisation of that stimulus even if it subsequently becomes task-irrelevant (Kyllingsbæk et al., 2001; Maljkovic & Nakayama, 1994; Sha & Jiang, 2016); and prior experience that a target item or a distractor item is likely to appear in a particular location can promote or suppress (respectively) the prioritisation of items that subsequently appear in that location (e.g., Chun, 2000; Colagiuri & Livesey, 2016; Geng & Behrmann, 2005; Wang et al., 2019). Most importantly for the purposes of the current thesis, a substantial literature in the selection history field has shown that prioritisation is also influenced by prior learning about the relationship between stimuli and motivationally significant outcomes, such as reward and punishment (for reviews, see: Failing & Theeuwes, 2018; Le Pelley et al., 2016; Watson, Pearson, Wiers, et al., 2019). Here I will focus on the influence of reward learning on attention, since this forms the basis of my subsequent studies.

1.2 Value-Modulated Attentional Capture (VMAC)

The ability to predict when and where rewarding events (food, money, sex etc.) will occur is clearly of value for our survival and flourishing in an ever-changing environment. Consequently, stimuli that signal the availability of reward have an important motivational
status. For example, in our hunter-gatherer past, the presence of animal tracks or colourful berries might indicate the availability of a nearby meal; likewise an attractive face was a ‘reward of beauty’ that signalled a reproductively fit mate (Chehokova et al., 2014). Reward learning in the modern world may take somewhat different forms – such as learning where to find the best coffee each day, or what different notifications from social media mean about interactions with our friends. Given that rewards are often the targets of motivated behaviour, it makes sense that we would often prioritise processing of stimuli associated with reward in a goal-directed way. For example, if I am informed that every time I detect a target on the left-hand side of a display I will receive $100, whereas every time I detect a target on the right I will receive 10c, I should clearly direct my attention to the left in order to optimise the outcome I receive (cf. Maunsell, 2004; Störmer et al., 2014).

Notably, however, recent research suggests that attentional prioritisation of reward-signalling stimuli may be mediated instead by processes of selection history. A stimulus consistently paired with a high-value outcome is more likely to ‘capture’ attention than a stimulus paired with a low-value outcome – an effect termed value-modulated attentional capture (VMAC; Le Pelley et al., 2017) – and this modulating effect of reward on attention has been found to occur independently of the observer’s goals. In this next section, I review procedures that have been developed to study the influence of reward learning on attentional priority.

1.2.1 The Two-Phase VMAC Procedure

A common procedure used to study this relationship is the two-phase procedure, introduced in a series of studies by Anderson et al. (2011a, 2011b). On each trial of an initial training phase, participants were presented with an array of six shapes which featured a target circle (either a red or green circle) among non-target circles (rendered in other colours). Participants were required to respond to the orientation (horizontal or vertical) of a line segment inside the target circle by making a speeded keypress (Figure 1.1A). On each trial, participants could earn monetary rewards for making rapid, correct responses, with the magnitude of reward dependent on the colour of the target on that trial. For example, a particular participant could have typically earned a large reward (5 cents) when the target was red, and a small reward (1 cent) when the target was green – hence, red was the high-value colour, and green was the low-value colour (the assignment of colours to conditions was
counterbalanced across participants – so for other participants green was the high-value colour and red was low-value).

Figure 1.1. Examples of the stimulus presentations on a trial in the A) training phase, and B) test phase of the two-phase procedure by Anderson et al. (2011b). In the training phase, participants were required to respond to the line segment inside a red- or green-coloured target to earn a reward. The colour of the target signalled whether a high- or low-value reward could be earned on that trial for a correct, rapid response. Then, during the test phase, participants responded to a shape-singleton target, and the red and green coloured circles served as distractors.

Following extensive training on this task, participants completed an unrewarded test phase. In this phase, the target on each trial was now defined by shape (a diamond among circles, or a circle among diamonds; Figure 1.1B). Critically, on half of the trials in the test phase, one of the non-target shapes was rendered in either red or green: that is, the search display could contain a *distractor* item that appeared in either the high- or low-value colour from the preceding training phase. Participants knew that: (1) colour was no longer relevant during the test phase of the experiment; (2) the target would never appear in red or green; and (3) that rewards were no longer available. Consequently, there was no goal-directed reason to attend to red or green items during the test phase, since doing so would only impair search for the shape-singleton target. Nevertheless, the data suggested that the reward history of the distractor modulated how likely it was to capture participants’ attention. Specifically, participants were slower to respond to the target shape when the trial featured a distractor in the high-value colour than in the low-value colour. As noted above, there was no top-down reason to prioritise the high-value colour in the test phase, and the physical salience of high- and low-value colours was equivalent (ensured by counterbalancing of stimulus assignments across participants).
Anderson et al. (2011a, 2011b) interpreted these findings as demonstrating that selection history had modulated the likelihood that the colours would automatically capture attention, independent of goal-directed and stimulus-driven processes – i.e., that participants’ previous experience of the high-value colour as a target associated with a large reward (in the training phase) meant it received greater attentional priority in the test phase than the low-value distractor. These findings have been replicated in other studies using two-phase procedures using reaction time measures (Anderson & Halpern, 2017; Failing & Theeuwes, 2014; Infanti et al., 2015; MacLean & Giesbrecht, 2015). Together, these findings have been interpreted to suggest that automatic, selection-history-based attentional biases toward stimuli can be formed through learning about the relationships between the stimuli and rewarding outcomes. Notably, all of these studies using the two-phase procedure had one common feature – in the initial training phase, participants were required to direct their responses toward the reward-related stimuli since these colours defined the targets of search. It is perhaps unsurprising then that being previously reinforced for attending to a stimulus will increase the likelihood that the stimulus continues to be attended to – as in the well-established Law of Effect (Thorndike, 1911) – even if the response is no longer task-relevant.

1.2.2 The Single-Phase VMAC Procedure

Some subsequent studies of value-modulated attentional capture have instead used single-phase procedures, wherein participants are never required to attend to the reward-related stimuli to earn points; that is, the reward-signalling stimuli are never targets of search. Hence, by contrast to the above two-phase approach which first examines the instrumental relationship between attention and rewarding outcomes, the single-phase procedure examines whether reward learning can exert an influence over attention when stimuli merely signal reward – that is, even when attending to a stimulus is not directly reinforced with a rewarding outcome.

One study by Pearson et al. (2016) used eye-tracking as a measure of spatial attention, since eye-movements (saccades) to a given location are closely and necessarily preceded by shifts in attention to that location (Deubel & Schneider, 1996) – such that gaze is taken as providing a measure of ‘overt attention’. On each trial of this task, participants were presented with an array of shapes; one of these shapes was a diamond, and the rest were circles (see Figure 1.2). Participants were required to make rapid saccades to the diamond-shaped target in order to earn a reward. The array also featured one coloured circle, termed
the *distractor*; this distractor was either orange or blue on each trial (all other shapes were grey). For each participant, one distractor colour was designated as the high-value colour, and the other colour was designated as the low-value colour (with assignment of blue and orange to these roles being counterbalanced across all participants). If participants made a rapid saccade to the diamond target while the search display contained a high-value distractor (e.g., an orange circle), then they could earn a large reward (500 points). Similarly, if they made a saccade to the diamond target while the display contained a low-value distractor (e.g., a blue circle), then they could earn a small reward (10 points); points were later converted into a cash reward, so the more points participants earned, the more money they would receive at the end of the experiment.

![Figure 1.2](image)

*Figure 1.2.* Example trials of a single-phase visual search task (e.g., Pearson et al., 2016). On each trial, participants had to make a rapid saccade to the diamond target to earn a reward; the colour of a distractor circle in the search display signalled whether a large (500 points) or small (10 points) reward was available, but participants knew that if they looked at the distractor before looking at the target, the reward would be cancelled. In the examples shown here, orange defines the high-value distractor and blue the low-value distractor; these assignments were counterbalanced across participants.

Importantly, participants were explicitly informed that if they ever looked at either of the coloured distractors, then the reward they could have earned would be cancelled for that trial. Hence, while the distractors signalled the magnitude of reward available on each trial, participants were never required to look at these coloured stimuli (their task was to look at the diamond), and in fact looking at the distractors was directly counterproductive to the participants’ goals of earning money. Consequently there was no goal-directed reason to
attend to the distractors; the best strategy for participants to recruit in this task was to ignore the coloured distractor circles entirely in order to maximise their earnings. Yet, this is not what participants did. Instead it was found that participants sometimes looked at the coloured distractors – and in particular, they were significantly more likely to look at the high-value distractor than the low-value distractor, even though this pattern of behaviour meant that they were more likely to miss out on high-value rewards (Pearson et al., 2016).

Subsequent studies using the single-phase VMAC procedure have further found that the bias towards the high-value distractor is particularly pronounced among fastest saccades (~200ms after onset of the search display: see e.g. Pearson et al., 2016). Moreover, the ‘attention-grabbing’ nature of the distractors (particularly the distractor that signals a high-value reward) has been found to exert an enduring effect on attention even when rewards are subsequently removed (Watson et al., 2019b). Interestingly, reward learning not only exerts an influence on attentional capture by reward-related stimuli, but also attentional disengagement; Watson et al. (2020) found that participants were slower to disengage attention from the high-value distractor than the low-value distractor, even though this pattern was counterproductive since slower responses meant that participants were less likely to earn reward in this task. Taken together, these findings suggest that the human attentional system rapidly prioritises stimuli that have learned associations with rewards, even when doing so is counterproductive to our goals, and when the rewards themselves are no longer delivered. These findings suggest that reward-signalling stimuli can be prioritised by attention independently of top-down goals and intentions.

1.2.2.1 Associative Relationships in VMAC Procedure

The above findings from the single-phase VMAC procedure highlight that stimuli can be prioritised by our attentional systems even when they have only merely signalled the availability of rewards (particularly high-value outcomes over low-value outcomes). The implication of these findings is that reward can exert an influence on attention through learning about what the stimulus signals. In associative learning terms, the reward-related distractors capture attention by virtue of the Pavlovian relationship between the stimulus (high- or low-value distractor) and the outcome (high- or low-value reward), rather than instrumental relationships featured in previous procedures (wherein the distractors are first trained and reinforced as targets of search).
1.2.2.2 Sign-Tracking and VMAC

Through repeated Pavlovian pairings of stimuli and rewarding outcomes, it has been suggested that reward-related stimuli can become imbued with incentive salience, wherein the cues gain salience through a change in their perceptual representation (Berridge & Robinson, 1998). With greater incentive salience, these reward cues can become ‘motivational magnets’ that beckon us toward them and drive reward-seeking behaviour (Berridge & Robinson, 1998). The ability of reward-related cues to elicit approach behaviours by virtue of their signal-value is well-established in animal models distinguishing between sign-tracking and goal-tracking behaviour (Boakes, 1977). In a typical task used to investigate these distinct behaviours, a neutral stimulus (e.g., a lever insertion into a conditioning chamber) is repeatedly paired with a rewarding outcome (e.g., delivery of a food pellet to a food well). When the stimulus is presented, an animal demonstrating goal-tracking behaviour learns to approach the location wherein the reward is delivered (i.e., the food well). By contrast, an animal displaying sign-tracking behaviour approaches the signal of reward (i.e., the lever) – licking, biting, and gnawing at the lever – even though it has no operant value in producing the reward, indicating that the signal of the reward itself has become imbued with incentive salience. Clearly there are parallels between VMAC in the single-phase procedure and sign-tracking, wherein the distractors (i.e., the signals of reward) can elicit attentional prioritisation over the target (i.e., the goal), despite never being trained and reinforced as targets of search. In this sense, VMAC can be viewed as a human analogue of the sign-tracking response.

1.2.3 Clinical Implications of VMAC

The effect of reward on attention has potentially far-reaching consequences considering that reward is central to the lifestyles that many of us live. We are in an unprecedented era of constant access to rewards, such as food, alcohol, drugs, gambling, shopping, gaming, sex, pornography, and social media. It is no wonder that the world we live in today has been termed ‘Dopamine Nation’ (Lembke, 2021). With each reward, there are inevitably a growing number of stimuli that come to be associated with these rewards through learning (fast food logos, tobacco stores on every street, bright advertisements, flashing lights on gambling machines, notification sounds from our phones). More concerning is that many of these stimuli are specifically designed and funded to captivate our attention, interest, time, and money. For example, in 2021 alone, the gambling industry in Australia spent nearly $300
million on advertising, with an average of 948 advertisements for gambling broadcast on television per day around the country (Victorian Responsible Gambling Foundation, 2021).

With the abundance of reward-related stimuli around us, it has been suggested that their incentive salience may play a role in reward-seeking behaviours implicated in compulsive disorders such as addiction (Berridge & Robinson, 1998). On the idea that there are considerable parallels between VMAC and sign-tracking (discussed above), the propensity for an animal to sign-track has been associated with neurobiological and behavioural markers of addiction (when compared to goal-tracking; Flagel et al., 2007, 2008, 2009; Robinson & Flagel, 2009). Taking this account to a human level, someone who smokes may over time associate smoking-related cues (such as a lighter, or the sign for the tobacco store down the street) with the pleasurable feelings associated with smoking. These smoking-related cues (mere signals of the actual rewarding outcome – i.e., the cigarette) may be imbued with incentive salience to increase the likelihood that this person will engage in cigarette-seeking behaviour (heading into the tobacco store to purchase cigarettes). In line with this kind of account in humans, evidence suggests that the tendency for attention to prioritise drug-related stimuli may play a role in maintaining maladaptive behaviours implicated in addiction and compulsive disorders (review by Field & Cox, 2008). Indeed, the relationship between reward-related attention and compulsive behaviour seems to be even more general, with recent work demonstrating that individual differences in attentional capture by signals of (non-drug) monetary reward predict illicit drug use (Albertella et al., 2017), risky alcohol use (Albertella, Watson, et al., 2019), trait compulsivity and impulsivity (Albertella, Le Pelley, et al., 2019). Moreover, individual differences in attentional biases to drug-related cues and signals of monetary reward have been shown to predict the likelihood of abstinence success (Albertella et al., 2021; Marissen et al., 2006; Waters et al., 2003). Together these findings suggest that VMAC may have clinical utility in assessing and tracking intervention effectiveness for addiction (Anderson, 2016b).

1.2.4 Section Summary

While traditional models of attention have typically studied top-down and bottom-up effects on attention, recent evidence suggests that our prior experiences can exert independent, automatic, and persistent influences on attentional prioritisation. Much of the existing work on selection history has investigated consequences of reward learning, and has shown that stimuli associated with large rewards are more likely to be prioritised by our
attention than stimuli associated with small rewards. This process of value-modulated attentional capture (VMAC) operates independently of top-down influences on attention, given that attending to reward-related distractors occurs even when it is directly counterproductive to an observer’s top-down goals and intentions. The ‘attention grabbing’ properties of reward-related stimuli also transcend their physical salience (i.e., bottom-up influences on attention), shaping the likelihood of capture by stimuli that are otherwise equivalent in their physical salience. Effects of reward are observed under conditions in which attending to stimuli has been directly and instrumentally reinforced in the past, but also seen for stimuli that are merely Pavlovian signals of reward – providing evidence consistent with the idea of ‘attentional sign-tracking’. Moreover, reward-related stimuli continue to be prioritised by the attentional system even once the rewards themselves are removed (Anderson et al., 2011b, 2011a; Watson, Pearson, Most, et al., 2019). It also appears that reward may not only have an influence on the extent to which a stimulus captures attention, but also the extent to which it sustains and holds attention, as suggested by recent demonstrations of reward modulating the speed of disengagement from distractors (Watson et al., 2020). Clearly, the above studies demonstrate that reward learning has a far-reaching influence on our attention, and this may have important implications for our understanding of attentional biases in the clinical realm. Existing evidence already suggests that VMAC may play a role in maintaining maladaptive reward-seeking behaviours, such as those seen in drug addiction (Field & Cox, 2008).

1.3 Attentional Prioritisation as a ‘Habit-Like’ Response

As outlined above, learning that a stimulus is associated with a rewarding outcome can imbue that stimulus with enhanced attentional priority. Importantly, this process of attentional prioritisation appears to occur independently of, and indeed contrary to, top-down goals and intentions. Viewed another way, reward-related stimuli can be said to automatically elicit a conditioned attentional prioritisation response. Conceptualising attention as a response that can be conditioned through learned associations with reward raises a question that is central to this thesis: does attentional prioritisation of reward-related stimuli exhibit features of a ‘habit’? In the next section, I will review research on habits as traditionally studied in the field of behavioural neuroscience, and outline how VMAC shows some features in common with instrumental habits.
1.3.1 Defining Habitual versus Goal-Directed Behaviours

The word ‘habit’ is commonly used to describe regular tendencies or practices – for example, a smoking habit, or an exercise habit. Crystallising behaviours as habits is argued to be critical to simplifying our interactions with the complex world (Daw, 2015). For example, driving the same route to work, or ordering coffee from the same café every day, reduces the number of cognitively demanding decisions we must make. Despite being so ubiquitous in the layman’s vocabulary and generally understood as behaviours that are performed automatically, without conscious deliberation, and repeatedly in the same context (Gardner, 2015; Robbins & Costa, 2017), it is difficult to operationalise and measure habits for empirical research based on these properties alone (Watson et al., 2022).

Researchers in the field of behavioural neuroscience have typically studied habitualbehaviours as distinct from goal-directed behaviours, with the relative balance between these two systems forming the backbone of instrumental actions. According to this framework, a response must meet the belief-desire criteria in order to be considered goal-directed; conversely, behaviour that fails to meet these criteria is considered habitual (de Wit & Dickinson, 2009; Heyes & Dickinson, 1990). According to the belief criterion, goal-directed behaviour is based upon knowledge of the causal relationship between the action and the outcome – i.e., that the response will cause the goal. The desire criterion further captures the idea that goal-directed actions are controlled by the current motivational value of the outcome, in that goal-directed actions are performed when the expected outcome is desired or valuable.

Consider the example of a rat trained to press a lever (i.e., the response) to produce a food pellet (the outcome). This behaviour would be considered goal-directed if: 1) it is based on knowledge that there is a causal instrumental relationship between lever pressing and the food pellet (i.e., the belief criterion), and 2) the rat desires the food pellet at the time of lever pressing (i.e., the desire criterion). To diagnose whether the response meets these two criteria, researchers have used procedures which degrade either the causal relationship between the response and outcome, or the desire for the outcome. Here I will focus my review on the desire criterion since this forms the basis of my subsequent studies.
1.3.1.1 The Outcome Devaluation Procedure

To diagnose whether behaviour is habitual by the desire criterion, researchers have used the *outcome devaluation procedure* (Adams & Dickinson, 1981). In a typical animal paradigm using this procedure, an instrumental response (e.g., a lever press) is first reinforced by the delivery of a reward outcome (e.g., a food pellet). Following the training phase, the outcome is devalued in order to diminish the rats’ desire for the outcome; for example rats might be fed to satiety, or the pellets may be paired with the induction of illness, reducing desire for pellets – and hence reducing the value of the pellets (Adams, 1982). The impact of this devaluation on conditioned lever-pressing is then tested in extinction, wherein food pellets are no longer delivered. Testing behaviour in extinction is essential to ensuring that any observed reduction in responding reflects sensitivity to outcome devaluation; if outcomes continued to be delivered following devaluation, the Law of Effect (Thorndike, 1911) could account for a gradual decline in responding for the devalued outcome, however this explanation would not be able to explain immediate sensitivity to outcome devaluation in the absence of further experience of response-devalued outcome pairings (de Wit & Dickinson, 2009).

Studies using this outcome devaluation procedure have shown that behaviours can be mediated by the outcome (i.e., goal-directed) after a small amount of initial training (Adams, 1982). Following brief training, instrumental behaviour is typically sensitive to the devaluation of the outcome – rats no longer press the lever when tested in extinction in line with the reduced desire for the outcome. This suggests that the lever pressing response is mediated by a representation of the outcome and its value, and hence behaviour is said to be goal-directed. By contrast, following extended training, behaviour can become insensitive to outcome devaluation – rats can continue pressing the lever even though they no longer desire the food (as demonstrated in a separate *consumption test* wherein rats who have experienced the devaluation no longer eat the food when given the choice). This kind of responding suggests that the lever pressing had become independent from a representation of the (now-devalued) outcome, and hence behaviour is driven by a conditioned S–R association (see lever ➔ press lever) in the absence of desire for the outcome – hence, behaviour in this case is said to be habitual.
1.3.2 The Habit Theory of Addiction

Habitual behaviour in particular has been a rapidly growing avenue of interest in the clinical domain, where it has been suggested that a propensity for habits underlies the compulsive behaviours implicated in a range of psychopathologies, particularly drug addiction. This ‘habit theory of addiction’ (Everitt & Robbins, 2005, 2016) posits that drugs are initially rewarding targets of goal-directed approach. Over time however, they become associated with a plethora of health, social, and financial consequences, resulting in reductions in desire for the drug (much as outcome devaluation reduces desire for food in the animal paradigm described above). The habit account proposes that despite these detriments, an abstaining smoker might reflexively seek for and light up a cigarette due to the presence of smoking cues in the environment (an S–R association), for example.

This account has been supported mostly by animal research examining the effect of chronic drug exposure on performance in the outcome devaluation task. It has been found that drug-seeking in animals is particularly insensitive to devaluation of drug outcomes (i.e., prone to habitual control), relative to food-seeking (Dickinson et al., 2002; Loughlin et al., 2017; Mangieri et al., 2012; Miles et al., 2003). Further, chronically drug-exposed rats are more likely to respond habitually after outcome devaluation than rats not pre-exposed to drugs (Corbit et al., 2012; LeBlanc et al., 2013; Nelson & Killcross, 2006; Nordquist et al., 2007; Schmitzer-Torbert et al., 2015), which has been interpreted to suggest that chronic drug use sensitises the habit system to drive drug-seeking behaviours.

1.3.3 Translating Outcome Devaluation in Humans

Given the potential importance of the distinction between goal-directed actions and habits in understanding motivated behaviour, more recent translational work has attempted to investigate this distinction in human behaviour. This research has used a variety of procedures, typically based on outcome devaluation procedures (see Watson & de Wit, 2018, for review). However, studying habits in humans has been met with some translational and theoretical problems. As discussed previously, one of the key defining features of stimulus-response habits is that they develop following extensive experience of the outcome being delivered in the presence of the stimulus; this finding has been replicated a number of times in animals (e.g., Dickinson et al., 1995; Thrailkill & Bouton, 2015). Habitual behaviour induced by extensive training has been found in some experimental studies in humans (Luque
et al., 2020; Tricomi et al., 2009), however this finding has sometimes been difficult to replicate. In five large-scale experiments with humans, de Wit et al. (2018) failed to replicate the transition from goal-directed to habitual behaviour first reported by Tricomi et al. (2009), with participants continuing to make correct responses to earn valued outcomes over now-devalued outcomes. That is, participants in these studies continued to demonstrate goal-directed sensitivity to outcome devaluation even after extensive training over multiple days. de Wit et al. (2018) concluded that tasks currently available to habit researchers may be tapping predominantly into goal-directed control (and momentary slips in this control), rather than habitual behaviour per se. Some have suggested that a possible explanation for this is that participants in the lab context remain in an active state of goal-directed control, carefully monitoring outcome values and responding accordingly (Watson & de Wit, 2018; Wood & Rünger, 2016).

Some researchers have gone further to suggest that so-called habitual behaviours in humans are actually goal-directed (Buabang et al., 2021, 2022; Kruglanski & Szumowska, 2020). Kruglanski and Szumowska (2020) argued that intrusions of overtrained behaviour are not potentiated by stimulus-response habits, but are instead motivated by alternative goals (not necessarily in the observer’s conscious awareness), known as hidden goals. For example, a classic ‘habit’ example of eating popcorn at the cinema even after satiation may reflect the attainment of goals that other than fulfilment of hunger, such as fulfilling social goals or having a complete cinema experience.

With regard to the more specific issue of the habit theory of addiction, Hogarth (2018, 2020) has argued that habits demonstrated in the animal model of outcome devaluation are unlikely to operate naturalistically in human drug users. In his review of outcome devaluation studies with drug-user populations (Hogarth, 2018), he examined the claim that habitual behaviours are more readily formed in drug-users, and found that this idea is not well supported by evidence from human studies – in fact, there was a 12:5 ratio of studies against this idea. Of the five studies that demonstrated a greater propensity toward habit formation in drug users, Hogarth (2018) argued that these findings could be attributed to general cognitive impairments or weaker motivation to engage, impairing knowledge of stimulus-outcome contingencies, rather than specifically reflecting habitual behaviours. In line with the ideas presented by Kruglanski and Szumowska (2020), Hogarth (2018) argues that drug-seeking is instead driven by excessive goal-directed behaviour. In support of this view, drug choice has
been found to be reliably increased by motivation to alleviate aversive withdrawal states (e.g., Hogarth et al., 2017; Hutcheson et al., 2001) and negative mood (e.g., Hardy et al., 2017; Hogarth et al., 2017; Hogarth & Hardy, 2018), and willingness to work to pay for drugs (e.g., Gray & MacKillop, 2014; MacKillop & Murphy, 2007, 2007; Murphy et al., 2011). Together, these findings present theoretical problems for understanding the role of habits in maintaining addiction in the real-world.

1.3.4 ‘Habit-like’ Attention

On the one hand, the idea that habits drive many of our behaviours has great intuitive appeal based on our everyday experience, and has been supported by a wealth of animal research using the outcome devaluation procedure. Notably, however, efforts to substantiate this claim in humans have been met with some translational and interpretive issues. It is here that this thesis aims to provide another perspective on habitual processes in humans.

As described above, much of the existing research on goal-directed and habitual behaviour has examined only overt behavioural choice in the context of instrumental actions: does the rat press the lever or not? Does the participant choose option A or option B? Does a smoker pick up and light a cigarette or not? Importantly, however, behavioural choice is merely the end-point of a chain of cognitive processes that eventuate in a final decision (Krajbich, 2019; Rangel et al., 2008). As discussed at the beginning of this chapter, our decisions begin with information-gathering – attending to and encoding the stimuli in our environment. Clearly if we do not notice that a stimulus is present, that stimulus will not be a target of our overt behaviour; conversely if our attentional system flags a particular stimulus as being significant, then it is more likely to be a focus of subsequent action. When shopping at the supermarket, for example, the availability of thousands of items on shelves means that our possible overt choices are practically limitless. We cannot carefully weigh up the pros and cons of every possible option. Instead we must be selective about the items that we consider, and this is the critical role of attention. Of course, much of our search will be goal-directed (if we have run out of milk we will prioritise the dairy section), but demonstrations of value-modulated attentional capture (VMAC) suggest that prioritisation of potential options can also be influenced by our previous experiences of outcome value: certain items may be more likely to ‘jump out’ at us from the display, and hence we may be more likely to buy them. Viewing attention as a ‘response’ that can be conditioned through learned associations with reward raises the idea that ‘habit-like’ processes possibly exist early on in
the decision-making cascade, in the form of an attentional prioritisation of reward-related stimuli.

It is important to note an issue of terminology here. As discussed earlier, the critical associative relationship between the distractor stimuli and the delivery of reward in the VMAC task is Pavlovian in nature. That is, distractors signal the availability of reward, but participants are not required to respond to the distractors themselves in order to earn this reward. Typically, however, habits are considered in the behavioural neuroscience literature to involve instrumental relationships. To reflect this distinction, I refer to the attentional patterns throughout this thesis as being ‘habit-like’, rather than reflecting “true” instrumental habits as they are traditionally understood. Furthermore, while we draw an analogy between the current research and the distinction between habitual and goal-directed behaviour, this should not be taken to imply that if VMAC is not found to have ‘habit-like’ properties then it is goal-directed. Clearly the pattern of attending to distractors in this task could never be considered goal-directed, as it is always counterproductive to look at these stimuli (since doing so cancels the reward). Instead, we intend the parallel here in relation to whether (1) the behaviour has become divorced from the value of (i.e., desire for) the outcome (which would be labelled a habit in the context of instrumental behaviour, and which we refer to as ‘habit-like’ in the context of VMAC), or (2) whether behaviour is mediated by retrieval of a representation of the outcome signalled by a stimulus (which would be labelled goal-directed in the context of instrumental behaviour, but not in the context of VMAC). Another way of considering this distinction between whether attentional prioritisation of reward-related stimuli is outcome-independent versus outcome-mediated is in terms of the distinction between conditioned attentional prioritisation being driven by stimulus-response associations (i.e., the distractor directly primes the attentional response) versus stimulus-stimulus associations (the distractor retrieves a representation of the outcome, and this is what results in capture of attention).

While much of the work on habits is based on instrumental relationships, there is considerably less work on ‘habit-like’ processes in Pavlovian relationships. A handful of studies in rats have examined whether Pavlovian signalling stimuli elicit responses that are insensitive to changes in the value of the outcome signalled by a stimulus. In animal studies, researchers have examined whether sign-tracking behaviour towards reward-signalling stimuli (e.g., appetitive behaviour towards a lever stimulus that signals a food outcome) is
sensitive to changes in the value of the outcomes through outcome devaluation. Some studies (e.g., Morrison et al., 2015; Nasser et al., 2015) have found that sign-tracking behaviours in rodents persist following outcome devaluation, in favour of the idea that sign-tracking becomes outcome-independent and hence ‘habit-like’. However in other studies (Derman et al., 2018; M. J. F. Robinson & Berridge, 2013), sign-tracking has been found to reduce following outcome devaluation, suggesting outcome-mediated sensitivity of sign-tracking to changes in outcome value. Potentially reconciling these mixed findings, Amaya et al. (2020) found that the sensitivity of sign-tracking could be both flexible and inflexible to changes in outcome value, depending on the specific experimental procedures in which the outcome devaluation was learned and tested. Specifically, sign-tracking could reduce in an outcome-mediated manner if the context in which the taste aversion learned in the devaluation phase was similar to the testing context; this suggests that the updated outcome values must be ‘online’ during test if behaviour is to update accordingly. By contrast, if the taste aversion to the devalued outcome failed to transfer to the testing context, then sign-tracking persisted in an outcome-independent manner (Amaya et al., 2020). Together these findings are in line with an idea raised in the study of instrumental habits—sensitivity to devaluation is not a binary measure, but may instead be a continuum ranging from a sensitive, outcome-mediated component and an insensitive, outcome-independent component, and specific learning or test conditions may shift the relative balance between these two extremes (Daw et al., 2011; Thrailkill & Bouton, 2015).

Given the parallels between sign-tracking and VMAC, my thesis aimed to examine whether ‘habit-like’ processes would operate in humans at the level of attention, and if so under what experimental conditions these processes would emerge. As noted earlier on in this chapter, the VMAC effect has shown some features in common with the instrumental habits that are typically studied in the behavioural neuroscience literature: 1) it increases with extended training (Pearson et al., 2015), 2) it persists in extinction (Anderson et al., 2011a; Watson, Pearson, Most, et al., 2019), and, 3) it seems to result from an automatic and involuntary behavioural process, as it can operate independently of (and indeed, contrary to) an individual’s goals (Le Pelley et al., 2015). Such findings have already led to the description of attentional prioritisation of reward-related stimuli as reflecting an ‘attentional habit’ (Anderson, 2016a; Anderson et al., 2016; Anderson & Yantis, 2013; Jiang & Sisk, 2019). However, it remains unclear whether attentional prioritisation of reward-related stimuli implicated by VMAC meets the key feature of habits as defined by behavioural
neuroscience – that attentional prioritisation becomes divorced from the value of (i.e., desire for) the outcome – or alternatively whether it is mediated by retrieval of a representation of the outcome signalled by a stimulus.

1.4 Research Presented in this Thesis

Building on ideas about instrumental habits and applying them to our understanding of attention, the central aim of my thesis was to investigate whether the conditioned attentional prioritisation response occurs independently of changes in the value of the outcome (that is, desire for the outcome), and hence whether attentional prioritisation could be considered a ‘habit-like’ response. Research presented in this thesis will address the following questions.

**Question 1: Under what conditions is attentional prioritisation of reward-related stimuli sensitive to changes in the value of the outcome?**

Previous research has shown that reward learning shapes the attentional priority of signals of reward, and that this process of VMAC occurs independently of, and often contrary to, the observer’s goals (Le Pelley et al., 2015). Viewed another way, reward-related stimuli can be said to automatically elicit a conditioned attentional prioritisation response, raising the idea that prioritisation of reward-related stimuli may be ‘habit-like’. A parallel line of research suggests that sign-tracking in animals may be ‘habit-like’ following outcome devaluation under some experimental conditions (Amaya et al., 2020). Given that VMAC can be considered a human analogue of sign-tracking, this idea raises the question of whether this attentional prioritisation toward high-reward-related stimuli meets a key defining feature of habits, in that it persists despite changes in the values of (i.e., desire for) rewarding outcomes. In Chapter 2, I will describe a series of four experiments assessing the impact of various forms of outcome revaluation on performance in a single-phase VMAC procedure, in order to investigate conditions under which reward-modulated attentional capture may exhibit ‘habit-like’ properties.
Question 2: Does the amount of training have an influence on the extent to which attentional prioritisation will be ‘habit-like’ in its persistence?

As discussed in Section 1.3.1, behaviour can remain under goal-directed control with a small amount of training – that is, it remains sensitive to changes in the values of outcomes. After an extended amount of training, however, behaviour can come under habitual control, indicated by an insensitivity to changes in the values of outcomes (Adams, 1982; Dickinson et al., 1995). In Chapter 3, I will present a study which examines whether attentional prioritisation demonstrates this transition from goal-directed to habit-like persistence, comparing sensitivity to outcome devaluation after moderate versus extended training.

Question 3: Does attentional prioritisation of reward-related stimuli result in ‘habit-like’ delayed disengagement?

Attentional prioritisation of reward-related stimuli influences not only the extent to which a stimulus captures attention, but also the extent to which it sustains and holds attention. That is, reward can also modulate attentional disengagement from stimuli, such that signals of high reward are more likely to take hold of attention than signals of low reward (Watson et al., 2020). The experiment presented in Chapter 4 examines whether attentional disengagement from reward-related distractors will be ‘habit-like’ in its persistence following outcome revaluation.
Chapter 2

The Effect of Outcome Revaluation on Attentional Prioritisation of Reward-Related Stimuli

2.1 Introduction

As discussed in Chapter 1, our attention often operates in a goal-directed way, filtering information from the world and prioritising stimuli that are relevant to our current goals (Connor et al., 2004). For example, if we needed to buy salad at the supermarket, our attention might scan ahead for displays of vegetables so that we can locate them faster. Notably, attention can also operate independently of our goals, with recent work demonstrating that prior experiences with rewards can exert an influence on the likelihood that stimuli will automatically capture attention (for reviews, see Failing & Theeuwes, 2018; Le Pelley et al., 2016; Rusz et al., 2020). Imagine that in the past, we have eaten a particular type of chocolate many times, and always found it delicious. On our way to the vegetable section, our attention may be captured by the wrapper of this rewarding chocolate, even though our goal is to look for salad. This fundamental effect of reward on attention is notable because the modern world is filled with reward cues: wrappers on high-calorie foods, logos of restaurant chains, billboards showing attractive models, advertisements for alcohol and cigarettes, and the bright flashing lights of gambling machines.

Many different procedures have been developed to study the influence of reward learning on attentional priority (Le Pelley et al., 2016; Rusz et al., 2020); here, we focus on a single-phase visual search procedure discussed in Chapter 1 (Section 1.2.2). In a variant of a single-phase visual search task, Pearson et al. (2016) found that people are more likely to attend to, and move their eyes toward, signals of high reward relative to signals of low reward. This attentional prioritisation of signals of high reward has reliably been shown to occur even when attending to the stimuli results in the cancellation of a larger reward (relative to looking at the distractor that signals a low-value reward; Pearson et al., 2016). The ‘attention-grabbing’ nature of the distractors (particularly the distractor that signals a high-value reward) persists even when rewards are subsequently removed (Watson, Pearson, Most, et al., 2019). This bias towards the high-value distractor was particularly pronounced among the fastest saccades that participants made (~200ms after onset of the search display:
e.g., Pearson et al., 2016). These findings suggest that the human attentional system rapidly prioritises stimuli that have learned associations with rewards, even when doing so is counterproductive to our goals, and when the rewards themselves are no longer delivered. Moreover, stimuli signalling high-value rewards compete more effectively for attentional prioritisation than do signals of low-value rewards. This modulating influence of reward learning has been termed value-modulated attentional capture (VMAC).

As discussed in Chapter 1, the VMAC effect shows some features in common with the instrumental habits that are typically studied in the behavioural neuroscience literature: it increases with extended training (Pearson et al., 2015), 2) persists in extinction (Anderson et al., 2011a; Watson, Pearson, Most, et al., 2019), and, 3) it seems to result from an automatic and involuntary behavioural process, as it can operate independently of (and indeed, contrary to) an individual’s goals (Le Pelley et al., 2015). Conceptualising attentional prioritisation as a response that can be conditioned through learned associations with reward raises the idea that attentional prioritisation implicated by VMAC may be ‘habit-like’ – that is, it may occur independently of changes in the value of the outcome (i.e., desire for the outcome). As discussed in the previous chapter, existing evidence from animal research on suggests that sign-tracking – which shares striking similarities with VMAC – can be ‘habit-like’, at least under certain conditions (Amaya et al., 2020).

The effect of outcome devaluation on attentional prioritisation of reward-related stimuli in humans has been investigated in only a few previous studies by other lab groups (De Tommaso et al., 2017; De Tommaso & Turatto, 2021; Pool et al., 2014; Watson, Pavri, et al., 2022). In the initial conditioning phases of these studies, participants learned to associate one stimulus with a pleasant chocolate odour (Pool et al., 2014) or a high probability of earning a desired beverage outcome (De Tommaso et al., 2017; De Tommaso & Turatto, 2021); stimuli of other colours were associated with no odour or lower probabilities of a drink reward, respectively. Participants who were hungry (Pool et al., 2014) or thirsty (De Tommaso et al., 2017; De Tommaso & Turatto, 2021) throughout the initial conditioning phase formed attentional biases toward the reward-signalling stimuli. Following the training phases in these studies, some participants experienced the devaluation of the rewarding outcomes through satiation of chocolate (Pool et al., 2014) or the drink reward (De Tommaso et al., 2017; De Tommaso & Turatto, 2021), with the idea being that desire for the rewarding outcomes in these participants would be diminished – analogous to animal paradigms.
wherin rodents are fed to satiation. Regarding the critical question of whether the attentional biases to reward-signalling stimuli formed in the training phases would persist following outcome devaluation, findings in these prior studies are mixed. Pool et al. (2014) found that participants who were sated on chocolate no longer exhibited an attentional bias to the chocolate-signalling stimulus (over the stimulus signalling no-reward), while participants who were not sated on chocolate (and hence still desired the chocolate outcome) continued to demonstrate the attentional bias formed in training. This finding was interpreted by the authors to suggest that the attentional prioritisation of stimuli signalling chocolate was mediated by a representation of the chocolate outcome and its current value, and hence flexible to the devaluation of the outcome. By contrast, De Tommaso et al. (2017, 2021) found that in the visual-search test phase, the attentional bias towards the high-probability drink-signalling stimulus (over the lower-probability stimuli) persisted even after participants quenched their thirst. These findings were interpreted to suggest that the high-probability stimulus maintained its attentional priority despite a reduction in desire for the beverage outcome; these findings were hence consistent with the idea that the attentional prioritisation of the high-probability stimulus became ‘habit-like’ in its independence from the (devalued) beverage outcome that the stimulus signalled.

There are some caveats that should be noted with regard to these prior studies that may account for the mixed findings. First, even after outcome devaluation in De Tommaso et al.’s (2017, 2021) procedure, the value of the high-probability cue (with a strong association to a weakly desired – but still appetitive – outcome) presumably remained higher than that of the low-probability cue (with a weak association to the same outcome). Consequently, the high-probability cue likely remained the ‘best’ stimulus; given recent evidence on the importance of relative value in determining reward-related attentional bias (Kim & Beck, 2020), it is perhaps unsurprising that the attentional bias towards this high-probability cue persisted under these conditions. Notably, whereas De Tommaso et al. report inferential statistics indicating that participants’ self-reported thirst reduced significantly as a consequence of the devaluation procedure, they do not report the mean pleasantness ratings following devaluation that would establish the extent of this devaluation – i.e., whether participants now found water unpleasant as a result. By contrast, Pool et al. (who found evidence for an influence of devaluation on attention) do report mean ratings following devaluation which suggest that devaluation may have been successful in rendering their chocolate reward no longer desirable.
Notwithstanding these issues regarding the potential (in)effectiveness of devaluation in these prior studies, the nature of the attentional process that produces the observed behaviour is also unclear. In the visual search task used in all of these studies to measure attention, the location of the search target was independent of the locations of reward-related colours, so the target sometimes appeared in the location of the reward-related stimulus. As a consequence, there was no cost of attending to the reward-related stimulus. This raises the possibility that attention to the reward-signalling stimuli in these studies may not reflect an automatic, reward-conditioned mechanism of attentional priority targeted by previous studies of VMAC, but rather may have been mediated by strategic, goal-directed attentional processes: participants may have deliberately prioritised reward-signalling stimuli because they (previously) provided useful information or predicted reward. Since the procedures in these prior studies do not allow clear distinction between reward-driven processes of selection history versus strategic allocation of attention, the level at which any observed devaluation effect (or lack thereof) was mediated remains unclear. In particular, the mixed findings in the test phases of the above studies may reflect differences in the relative balance between the automatic influences of reward on attention versus participant recruitment of strategic attentional prioritisation.

### 2.2 Experiment 1

Given the limits of previous studies on ‘habit-like’ attention (De Tommaso et al., 2017; De Tommaso & Turatto, 2021; Pool et al., 2014), we aimed to provide a clearer examination of the effect of outcome revaluation on automatic attentional capture by reward-related stimuli. In the current study, attending to reward-signalling cues was explicitly counterproductive as it resulted in omission of rewards that would otherwise have been earned, so that attention to these cues was always directly contrary to participants’ goals. Hence, our procedure allowed us to examine whether the automatic component of attentional prioritisation implicated by VMAC is mediated by retrieval of the outcome signalled by a stimulus, or whether instead prioritisation becomes divorced from the value of the events involved; that is, whether attention can become ‘habit-like’.

We designed a series of experiments that incorporated an instructed change in outcome values into a VMAC procedure that was based on the task used by Pearson et al. (2016). Hereafter, this will be referred to as the attentional revaluation task. The key
innovation of this task was to introduce ‘outcome’ elements that mediated between stimuli (colours) and rewards (high- vs. low-value points). Specifically, the colour of a distractor stimulus signalled the type of fruit that could be won on the current trial, with different fruits having different point values. By introducing the mediating fruit outcomes, we were able to keep the relationship between a stimulus and an outcome constant (e.g., a blue distractor always signalled that banana was available) while changing the value of that outcome by manipulating how much money each fruit was worth – much as in analogous studies of habits in rats, where the identity of the outcome remains constant throughout the task (e.g., lever pressing always produces food pellets) but the value of that outcome is changed (e.g., through feeding to satiety). This procedure allowed us to examine the effect of these changes in outcome value on patterns of attentional prioritisation of the reward-related distractors.

At the start of each experiment, participants were told that the aim of the task was to earn as many points as possible, since points would later be converted into a cash reward. They were also told that they could earn points by winning fruits, with one type of fruit worth a large number of points (high-value fruit), and the other fruit worth a small number of points (low-value fruit). On each trial of an initial training phase, participants had to make a rapid saccade to a diamond-shaped target among circles to earn a fruit, with the colour of a colour-singleton distractor circle signalling the type of fruit that was available: one colour (the high-value distractor) signalled that the high-value fruit was available, and the other colour (low-value distractor) signalled the low-value fruit. Following Pearson et al. (2016), if participants looked at the coloured distractor on a given trial, the fruit that could have been earned on that trial was cancelled. In line with previous findings of VMAC, we expected participants to have their attention—and gaze—captured more often by the high-value distractor than the low-value distractor, even though this meant that participants would be more likely to miss out on outcomes of higher value. This pattern of results would suggest the high-value distractor stimulus has a greater ability to elicit an attentional prioritisation response, relative to the low-value distractor. As in previous work, we also expected this pattern of a gaze-bias to the high-value distractor (relative to the low-value distractor) to be particularly pronounced among the fastest-latency saccades that participants made.

Following this training phase, participants in Experiment 1 were instructed that the values of the outcomes had changed – that both fruits were now worth only a small number of points, thus implementing a devaluation manipulation. All participants then completed the
test phase of the search task. The central question was how this change in outcome value influenced patterns of attentional prioritisation of the reward-related distractors that had previously formed during the training phase. Half of the participants in Experiment 1 – the devaluation (Dev) group – completed the test phase under a ‘nominal extinction’ procedure: participants were told that they could still earn fruits, but were not told about the identity of the fruit they earned on each trial. Nominal extinction is commonly used in tests of habits in humans, as it prevents new learning of relationships between stimuli and outcome values, while preserving participants’ motivation to perform (e.g., Luque et al., 2020; Tricomi et al., 2009). Data from participants tested under nominal extinction allowed us to assess whether value-modulated attentional capture is mediated by explicit knowledge of outcome value, in the absence of further experience of stimulus-outcome (colour-fruit) pairings under the new value regime.

The other half of participants continued to receive trial-by-trial feedback on the identity of the specific fruit earned in each trial of the test phase – this group was called the devaluation-with-feedback (DevFB) group. That is, in addition to knowledge of the revalued fruit outcomes, participants in this group had direct experience of the relationship between distractor colours and now-revalued fruits. Hence, the contrast between the Dev and the DevFB condition allowed for the comparison of the role of knowledge versus experience-driven learning in determining updating of attentional priority following a change in outcome value.

In line with previous findings on VMAC (e.g., Le Pelley et al., 2015; Pearson et al., 2015; Watson, Pearson, Chow, et al., 2019; Watson, Pearson, Most, et al., 2019), we expected both groups to demonstrate greater attentional capture by the high-value distractor than the low-value distractor during the training phase, prior to the devaluation manipulation. Of greater interest was performance during the test phase following the instructed devaluation. If conditioned prioritisation of the high-value distractor relative to the low-value distractor is insensitive to the current value of the associated outcome, then we would expect the pattern learned in training (greater capture by the high-value than low-value distractor) to persist following devaluation of the outcomes during the test phase, despite the change in outcome values. By contrast, if prioritisation of the high-value distractor was mediated by a representation of the associated outcome – and hence dependent on the current value of that outcome – we would expect the pattern of attentional bias to reflect changed values of the
fruit outcomes, with both high-value and low-value distractors now equally likely to capture attention, consistent with the equal low value of these distractors during the test phase.

As noted above, if we were to find that attention during the test phase was not influenced by outcome devaluation, this would be in line with the idea that conditioned attentional prioritisation is independent of knowledge of the current value of the associated outcome. However, an alternative explanation of this finding would be that the outcome devaluation procedure was ineffective in changing participants’ knowledge of the outcome values; e.g., perhaps participants did not encode the updated values of the fruits. We guarded against this possibility to some extent through the use of check questions: immediately after reading the instructions regarding updated outcome values, participants were required to select the new values of each fruit before they could proceed. However, while this procedure ensured that participants had encoded the new values, it remained possible that they then rapidly forgot those updated values. We therefore implemented an additional test of participants’ knowledge and retention of the updated outcome values. Specifically, in a knowledge check following the test phase, participants were asked to select the current value of each fruit outcome; this allowed us to verify that participants had explicit knowledge of the updated outcome values that persisted through the whole of the test phase. We note that previous studies of instrumental habits in humans have sometimes instead used a ‘consumption test’ to assess the efficacy of devaluation (based on the approach used in animal studies that assesses willingness to consume a now-devalued food outcome: e.g., Adams & Dickinson, 1981). In a typical consumption test in human studies, participants are shown two outcomes (one still-valuable, the other devalued) and are asked to select the outcome they would prefer to receive – on the rationale that, if devaluation has been effective, participants will choose the still-valuable outcome – with these consumption test trials being interspersed with ‘standard’ test trials (assessing conditioned responding) during the test phase (e.g., Gillan et al., 2015; Luque et al., 2017). Our test of explicit knowledge following the test phase served a similar function to consumption test trials by assessing knowledge of the current values of the outcomes, and has some advantages over this previous approach. First, the explicit knowledge check is arguably a more conservative test of knowledge of outcome values, since it requires participants to reflect on and report this knowledge for each outcome individually and directly; by contrast, the forced choice between outcomes inherent in a consumption test may be open to additional (potentially more automatic) influences that are invoked by the requirement to choose between items, that go
beyond explicit knowledge about the values of those outcomes. That is, consumption choice trials seem a less direct way of assessing participants’ knowledge than simply asking them to report this information. Moreover, by assessing knowledge after the test phase – rather than at points during that phase – our procedure allows us to verify that participants’ knowledge spanned the whole of this critical phase.

2.2.1 Methods and Materials

2.2.1.1 Participants

Previous studies have found medium to very large effects ($d_z = 0.54–2.20$) for the influence of reward on attentional capture (e.g., Le Pelley et al., 2015; Pearson et al., 2015; Watson, Pearson, Chow, et al., 2019). Hence we aimed to recruit at least 24 participants per condition; G*Power (with default settings) revealed that this would give power of .80 to detect a medium-sized effect ($d_z = 0.6$) of reward on attention in each condition, and power >.90 to detect a medium-sized effect ($\eta^2_p = .06$) for the interaction relating to differences in the reward-related attentional bias across conditions. In total, 51 UNSW Sydney students participated (31 females; age $M = 18.56$, $SEM = 0.19$ years; Dev group $n = 26$, DevFB group $n = 25$). Group assignment alternated based on order of arrival. Participants earned course credit, and received a monetary bonus depending on the number of points earned in the attentional revaluation task ($M = A$9.67, $SEM = A$0.33). Research reported here was approved by the UNSW Human Research Ethics Advisory Panel (Psychology).

2.2.1.2 Apparatus

Stimuli were presented on a 23-inch monitor (60 Hz refresh rate, 1920×1080 resolution); stimulus presentation was controlled by MATLAB with Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner, Brainard & Pelli, 2007). Monitor-mounted eye-trackers were used to record eye movements in the attentional revaluation task. All participants were tested using Tobii Pro Spectrum eye-trackers (sampling rate 600 Hz). Gaze data were down-sampled to 100 Hz for gaze-contingent calculations during stimulus presentation. Head position was stabilised using a chin-rest 60 cm from the monitor. The eye-tracker was calibrated at the start of the experiment using a 5-point procedure.
2.2.1.3 Design and Procedure

The attentional revaluation task consisted of four components: the initial value instructions, training phase, revaluation instructions, and test phase.

**Initial Value Instructions.** At the start of the experiment, participants were told that their aim should be to earn points, since these points would be converted into cash at the end of the session, and that they could win points by earning fruits — lemons and bananas. Half of the participants were initially informed that lemons were worth 500 points and bananas were worth 10 points (Figure 2.1A); for the other half of participants, this was reversed. Participants were told that they could earn these fruits by moving their eyes to the diamond target on each trial “as quickly and as directly as possible”, but that if they looked at the coloured circle that was present in the search display, the fruit they could have earned would be cancelled. However, participants were not informed of the specific colour–fruit contingencies (e.g., that a blue circle signalled availability of lemons and an orange circle signalled bananas).

![Figure 2.1](image.png)

*Figure 2.1. Examples of stimulus presentation in Experiment 1. A) Participants were informed of the values of each fruit outcome at the start of the attentional revaluation task, and reminded of these values prior to each block of trials. B) An example trial. On each trial, participants first focused on a central fixation cross. A search display then appeared, containing a diamond-shaped target among circles. One of these circles was a colour-singleton distractor, whose colour signalled the type of fruit available for making a rapid saccade to the target. The example here shows a high-value distractor trial, in which a blue distractor signals availability of a fruit (lemon in this case) that is worth 500 points; note that fruit-value and colour-fruit contingencies were counterbalanced across participants. If participants looked at the distractor before looking at the diamond, or if they did not respond quickly enough, the fruit reward would not be delivered on that trial.*
Training Phase. Each trial of the search task consisted of a fixation display, search display, and feedback display (Figure 2.1B). All stimuli were presented on a black background. The fixation display comprised a central white cross (0.5° visual angle) surrounded by a white circle (diameter 3.0°). When participants had accumulated 700 ms of gaze time inside this circle, or after 4000 ms, the cross and circle turned yellow to indicate that the search display was imminent. After 300 ms the screen blanked, and after a 150-ms delay the search display appeared: a diamond and 5 circles, each 2.3 × 2.3° visual angle, distributed evenly around the centre of the screen, with the centre of each shape at an eccentricity of 5.1°. One of the circles was coloured either orange (CIE x,y chromaticity coordinates .493/.445) or blue (CIE x,y .192/.216) with similar luminance (~24.5 cd/m²): we term this colour-singleton circle the distractor. All other shapes were grey (CIE x,y .327/.400, luminance ~8.3 cd/m²).

Participants’ task was to move their eyes to the diamond target as quickly as possible; a response was registered when they had accumulated 100 ms of gaze dwell time within a region of diameter 3.5° centred on this target. The colour of the distractor in the search display signalled the type of fruit that was available for a rapid response: for half of the participants, a blue distractor signalled that a rapid saccade to the diamond target would earn a lemon, and an orange distractor signalled a banana; for remaining participants this assignment was reversed. The distractor signalling the high-value fruit (worth 500 points) was termed the high-value distractor, and the distractor signalling the low-value fruit (worth 10 points) was the low-value distractor. If any gaze fell on or near the distractor (within a region of diameter 5.1° centred on this distractor) prior to a response being registered, no reward was given on that trial; these were recorded as distraction trials.

The search display was replaced by the feedback display immediately when a response was registered, or after 2000 ms (timeout). If response time was less than 1000 ms and it was not a distraction trial, then the feedback display stated “Fruit won!” and showed a picture of the appropriate fruit. If the trial was a distraction trial, feedback stated “No reward: You could have won”, and showed the fruit overlaid with a large red ‘X’. If response time was greater than 1000 ms, feedback stated “Too slow: You could have won”, and presented the fruit overlaid with a red ‘X’. If no response was registered before the trial timed-out, feedback read: “Too slow: Please try to look at the diamond more quickly”. Feedback appeared for 1400 ms; the next trial then began after a blank inter-trial interval of 1400 ms.
Participants completed 16 blocks of trials in the training phase, with each block containing 24 trials (384 trials in total). Half of the trials in each block featured a high-value distractor, and the other half featured a low-value distractor: trial order in each block was random. The locations of the target and distractor were randomly determined on each trial. At the end of each block participants took a short break, during which they were informed of the total number of points they had earned so far. During each break, participants also received a reminder of the fruit values (as in Figure 2.1A) which appeared on the screen for at least 10s; following this 10s period, participants opted when to continue with the task.

Revaluation and Feedback Instructions. Immediately following the training phase, participants were instructed that the fruit previously worth 500 points in training was now worth 10 points, while the fruit previously worth 10 points was still worth 10 points (Table 2.1). Participants in the Dev group were further informed that while they would still be earning fruits during the test phase, they would no longer be told whether they had earned a lemon or banana on each trial; “instead you will simply be told whether or not you won a fruit – and we will keep track of how many points you have earned”. By contrast, participants in the DevFB group were told that “as before, you will be told whether you earned a lemon or a banana on each trial, and we will keep track of how many points you have earned”. After this instruction, all participants answered check questions to ensure their knowledge of the current fruit values: participants were shown a picture of a lemon and a banana and were asked to select the current value of each fruit. Both responses were required to be correct before they could proceed.

Table 2.1. Design of each phase for one counterbalance condition in Experiment 1. Colours refer to colours of the distractor in the search display; fruits refer to outcomes that could be won for saccades to the target. Note: fruit–value and colour–fruit contingencies were counterbalanced across participants.

<table>
<thead>
<tr>
<th>Initial Value Instructions</th>
<th>Training Phase</th>
<th>Revaluation Instructions</th>
<th>Group</th>
<th>Test Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lemon = 500 pts Banana = 10 pts</td>
<td>Blue → Lemon Orange → Banana</td>
<td>Lemon = 10 pts Banana = 10 pts</td>
<td>Dev</td>
<td>Blue → ?? Orange → ??</td>
</tr>
<tr>
<td>DevFB</td>
<td>Blue → Lemon Orange → Banana</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Test Phase.** All participants then completed the test phase of the search task. In the test phase, participants in the Dev group continued to earn fruits for responding correctly, but did not receive feedback on the identity of the fruit earned on each trial. If response time was below 1000 ms and it was not a distraction trial, then feedback stated “Fruit won!”, accompanied by “??” where previously a picture of the fruit had appeared during the training phase. If the trial was a distraction trial, feedback stated “No reward: You could have won a fruit”, and displayed “??” overlaid with a red ‘X’. If response time was above 1000 ms, feedback stated “Too slow: You could have won a fruit”, and displayed “??” overlaid with a red ‘X’. Participants in the DevFB group received feedback on the identity of the fruit earned – or omitted – on each trial (lemon or banana), as in the training phase. Participants completed 8 blocks of trials in the test phase, with blocks structured as in the training phase. Participants in all groups were reminded of the current values of each fruit in the short break that followed each block.

**Knowledge Checks.** Following the test phase, participants’ knowledge of the colour–fruit contingencies was assessed. Participants were told that the type of fruit that could be won on each trial depended on the colour of the coloured circle in the search display. They were then presented with an orange and a blue circle, in random order, and for each they were asked to select which fruit (banana or lemon) they could win when that stimulus appeared in the search display. We also included a final knowledge check for the fruit-value contingencies to ensure that participants were aware of the new fruit values following revaluation. In this latter knowledge check, participants were presented with each fruit in random order and were asked to select whether it was currently worth 500 points or 10 points.

**2.2.1.4 Data Preparation**

Data analysis followed our established protocols (Le Pelley et al., 2015; Pearson et al., 2016; Watson et al., 2021). We discarded data from the first two trials after each break, trials that timed out with no response (0.57% of all trials), and trials with less than 25% valid gaze data (as a result of blinks, poor eye-tracking etc.: 0.55% of trials). For the remaining trials, valid gaze data were registered in over 99% of samples from the eye-tracker. The main dependent variable of interest was the proportion of distraction trials: the proportion of trials on which participants looked at the coloured distractor, resulting in cancellation of the fruit outcome. We analysed proportion of distraction trials as a function of whether the trial
featured a high- or low-value distractor; note that we label distractors according to the value they signalled during the training phase.

In line with previous work from our lab (Le Pelley et al., 2019; Pearson et al., 2016; Watson et al., 2021), we also analysed the direction of participants’ first saccade on each trial as a function of the latency of that saccade (i.e., the time between onset of the search display and initiation of the first saccadic eye movement). A velocity-threshold identification algorithm (Salvucci & Goldberg, 2000) identified saccades using raw data from the eye-tracker, sampled at 600 Hz. Gaps in the raw gaze data shorter than 75 ms were first interpolated using linear interpolation. The gaze data were then smoothed using a five-point moving average filter. The first saccade on each trial was then identified as the first eye movement that remained above a velocity criterion of 40° visual angle per second for at least 10 ms. This saccade was classified as moving in the direction of the distractor if the saccade vector had an angular deviation less than 30° to the left or right of the centre of the distractor.

For these latency-based analyses, in addition to the exclusions described previously, trials were excluded if the start point of the saccade was not within 100 pixels of the central fixation point, if saccade latency was below 80 ms, if there were gaps in the gaze data which were too large to be interpolated, or if there was insufficient gaze data to identify a saccade. There were 20 participants who had more than 30% of total trials excluded from either phase of the task and hence, as in previous work (Le Pelley et al., 2019; Pearson et al. 2016, 2021; Watson et al., 2021), were excluded from the saccade-latency analyses. Across the remaining 37 participants (n = 22 in the Dev group, n = 15 in the DevFB group), these additional trial-exclusion criteria resulted in removal of a further 7.7% of all trials. Included trial data for each participant were grouped by the phase (training vs. test) and distractor-type (high- vs. low-value). The Vincentizing procedure (Ratcliff, 1979) was then used to separate first saccade latencies into three time bins (defined by the tertiles of the distribution) representing the fastest, middle, and slowest groups of saccades. For each time bin in each phase of the experiment, we calculated the proportion of first saccades that went towards the distractor.
2.2.2 Results

2.2.2.1 Training Phase

**Proportion of Distraction Trials.** We first examined the proportion of distraction trials across the training and test phases using a 2 (phase: training vs. test phase) × 2 (distractor-type: high- vs. low-value distractor) × 2 (group: Dev vs. DevFB) ANOVA. Notably, preliminary analysis revealed that the three-way interaction was not significant, $F(1,49) = 0.004, p = .95, \eta^2_p < .001$. Nevertheless, planned analyses focused on the proportion of distraction trials in the training and test phases separately.

Figure 2.2A shows proportion of distraction trials during the training phase. We analysed these data using ANOVA with factors of distractor-type and group. There was a significant main effect of distractor-type, $F(1,49) = 26.72, p < .001, \eta^2_p = .35$, with more distraction trials when the display contained a high-value distractor than a low-value distractor. That is, participants were more likely to look at the distractor signalling availability of a high-value fruit than the distractor signalling a low-value fruit, even though this pattern of behaviour was counterproductive because looking at the distractor caused cancellation of the fruit. There was no significant main effect of group, $F(1,49) = 1.64, p = .21, \eta^2_p = .03$, or distractor-type × group interaction, $F(1,49) = 1.57, p = .22, \eta^2_p = .03$. These latter null findings are unsurprising, since both groups received equivalent treatment until after the training phase.

Notably, the above analyses examine performance in the training phase collapsed across all blocks of training, including the initial training blocks wherein differences in the proportion of distraction trials on high- and low-value distractor trials may have been small. Hence we repeated the above analyses, restricting data from the training phase to only the last two blocks immediately prior to the value-switch manipulation – these analyses did not affect the pattern of significant findings. There was a significant main effect of distractor-type, $F(1,49) = 40.65, p < .001, \eta^2_p = .45$, with more distraction trials when the display contained a high-value distractor than a low-value distractor. There was no significant main effect of group, $F(1,49) = 2.03, p = .16, \eta^2_p = .04$, or distractor-type × group interaction, $F(1,49) = 0.90, p = .35, \eta^2_p = .02$. 
Figure 2.2. A) Proportion of distraction trials for trials featuring high- and low-value distractors in the training phase of Experiment 1. Distraction trials occurred when participants looked at the coloured distractor in the search display, resulting in omission of the fruit outcome on that trial. Error bars show within-subjects SEM (Morey, 2008). B) Proportion of first-saccades in the direction of the high- and low-value distractors as a function of mean first saccade latency in the training phase of Experiment 1 in the Dev and DevFB groups. Mean proportion of first saccades and first saccade latency (and their respective SEM error bars) were calculated separately for each time bin of individual-participant saccade latency distributions (fastest, middle, and slowest), defined by the tertiles of the distribution.

Distraction as a Function of Saccade Latencies. We examined the proportion of first saccades that went towards distractors during the training phase of Experiment 1 using a 2 (distractor-type: high- vs. low-value distractor) × 2 (group: Dev, DevFB) × 3 (latency time bin: fastest, middle, slowest) ANOVA; data are shown in Figure 2.2B. This revealed a significant main effect of distractor-type, $F(1,35) = 17.36, p < .001, \eta^2_p = .33$, with more first saccades toward the high-value distractor than the low-value distractor. There was also a main effect of saccade latency time bin, $F(2,70) = 78.37, p < .001, \eta^2_p = .69$, with shorter-
latency saccades being more likely to go towards distractors. The interaction between these two factors was also significant, $F(2,70) = 9.09, p < .001, \eta^2_p = .21$, with the bias towards the high-value distractor over the low-value distractor being more pronounced at shorter saccade latencies. There were no significant effects involving group, $F$s < 0.90, $p$s > .34, $\eta^2$s < .03.

### 2.2.2.2 Test Phase

**Proportion of Distraction Trials.** Figure 2.3A shows the proportion of distraction trials for each group in the test phase, following the devaluation of the high-value fruit. A distractor-type × group ANOVA revealed a significant main effect of distractor-type, $F(1,49) = 30.86, p < .001, \eta^2_p = .39$, with participants more likely to look at high-value than low-value distractors, even though during this test phase the outcome signalled by each type of distractor had the same (low) value. The main effect of group was not significant, $F(1,49) = 1.95, p = .17, \eta^2_p = .04$. Importantly, the interaction between distractor-type and group was not significant, $F(1,49) = 1.19, p = .28, \eta^2_p = .024$.

To further analyse this non-significant interaction, we calculated distractor difference scores for each participant by taking the difference in the proportion of distraction trials between high- and low-value distractor trials; comparing these difference scores between Dev and DevFB groups via a Bayesian independent samples $t$-test (using the default prior in JASP: JASP Team, 2020) yielded a Bayes factor of $BF_{01} = 2.19$ in favour of the null hypothesis (no difference in attentional bias towards the high-value distractor between the two groups) over the alternative hypothesis. Notably, Figure 2.2A shows that the mean attentional bias was numerically (but not significantly: see analyses above) smaller in the Dev group than the DevFB group during the training phase, a pattern that carried through to the test phase. The implication is that any small between-group difference in attentional bias during the test phase was not a consequence of the difference in their treatment. Consistent with this idea, an analysis comparing the distractor difference scores in each group from the test phase, while controlling for the size of each group’s attentional bias during the training phase (by using distractor difference scores from the training phase as a covariate in a Bayesian one-way ANOVA), yielded $BF_{01} = 3.56$, suggesting moderate evidence in favour of the null hypothesis of no difference between the groups during the test phase (Lee & Wagenmakers, 2014).
We conducted planned analyses of simple effects to test the effect of distractor type in each group during the test phase. These analyses revealed a significant effect of distractor type in both the Dev group, $t(25) = 3.14, p = .004, d_z = .62$, and the DevFB group, $t(24) = 4.74, p < .001, d_z = .95$. In both groups participants were significantly more likely to look at the distractor that previously signalled a high-value fruit in the training phase relative to the distractor that previously signalled a low-value fruit.
Distraction as a Function of Saccade Latencies. Figure 2.3B shows proportion of first saccades that went towards distractors during the test phase of Experiment 1, following the devaluation of the high-value fruit, for the Dev and DevFB groups respectively. As for the training phase, a distractor-type × group × latency bin ANOVA revealed significant main effects of distractor-type, $F(1,35) = 17.20, p < .001, \eta^2_p = .33$, and latency bin, $F(2,70) = 66.16, p < .001, \eta^2_p = .65$, and a significant distractor-type × latency bin interaction, $F(2,70) = 5.01, p = .009, \eta^2_p = .13$, with a bias towards the high-value distractor over the low-value distractor that was more pronounced at shorter saccade latencies. As in the training phase, there were no significant effects involving group, $F$s < 0.82, $p$s > .44, $\eta^2$s < .03.

As noted earlier, prior studies have shown that the influence of reward on gaze is most apparent among the fastest saccades that participants make (e.g., Failing et al., 2015; Pearson et al., 2016). Consequently, and following that previous research, follow-up analyses focused on data from the fastest saccade latency time bin. Paired samples $t$-tests revealed that the bias towards the high-value distractor (versus the low-value distractor) among these rapid saccades was significant in the Dev group, $t(21) = 2.76, p = .012, \delta = .59$, and approached significance in the DevFB group, $t(14) = 2.04, p = .06, \delta = .53$.

A 2 × 2 ANOVA comparing the Dev and DevFB groups in the proportion of saccades made to the high- and low-value distractors in the fastest latency bin revealed a significant main effect of distractor-type, $F(1,35) = 11.15, p = .002, \eta^2_p = .24$. The main effect of group was not significant, $F(1,35) = 0.12, p = .74, \eta^2_p = .003$, nor was the interaction between group and distractor-type, $F(1,35) = 0.01, p = .93, \eta^2_p < .001$. Thus the pattern of performance for the fastest saccades did not differ significantly between the two groups.

2.2.2.3 Knowledge Checks

All participants correctly identified the current values of each fruit (the fruit–value contingencies) in the knowledge check following the test phase. In each group, only 2 participants failed to correctly identify which fruit was signalled by each coloured distractor (the colour–fruit contingencies). Re-analysis of proportion of distraction trials while excluding these participants who failed to correctly identify the colour–fruit contingencies revealed that the patterns of significant and non-significant findings were unchanged.
2.2.3 Discussion

Experiment 1 examined the effect of changing the values of outcomes on counterproductive attentional prioritisation of reward-related distractor stimuli. In the training phase, prior to the devaluation manipulation, participants in both groups were more likely to have their attention captured by a distractor signalling a high-value outcome versus a low-value outcome, even though looking at the distractor resulted in omission of the outcome. Moreover, this pattern of greater capture by high-value distractors (relative to low-value distractors) was particularly pronounced among participants’ fastest saccades. These findings are consistent with previous demonstrations of value-modulated attentional capture (e.g., Failing et al., 2015; Le Pelley et al., 2015; Pearson et al., 2016; Watson et al., 2021), wherein the size of the reward associated with a stimulus modulates the extent to which that stimulus rapidly captures attention. Viewed another way, findings from the training phase suggest that the high-value distractor stimulus was more likely to elicit an attentional prioritisation response than the low-value distractor.

We were particularly interested in whether the pattern of attentional prioritisation changed as a result of the change in outcome values prior to the test phase of the attentional revaluation task. In the test phase, participants in both groups continued to show greater attentional capture by the high-value distractor (i.e., the distractor that had previously signalled a high-value reward during the training phase) relative to the low-value distractor (the distractor that had previously signalled a low-value reward during the training phase). This persistent orienting toward — and prioritisation of — the high-value distractor occurred even though it no longer signalled a high-value outcome in the test phase, and in fact signalled a fruit outcome that was now of equal (low) value to the fruit associated with the low-value distractor. Results of latency-based analyses showed a similar pattern: rapid first saccades in both groups were more likely to be initiated toward the high-value distractor than the low-value distractor across both the training and test phases. Strikingly, this pattern of attentional prioritisation in the Dev group occurred despite the fact that participants were aware of the new values of the fruits in the test phase; they were explicitly informed of the new values, required to pass check questions on the values of the fruits following instructed revaluation, reminded of these values repeatedly at the end of each block, and able to correctly report the equal values of the two fruits in a knowledge check at the end of the experiment. The last point in particular demonstrates that participants had explicitly encoded
the information about the updated values of the outcomes, and retained that knowledge throughout the test phase. Consequently, the insensitivity of the attentional bias to outcome devaluation cannot simply be ascribed to a failure of the devaluation manipulation to shift explicit knowledge regarding the values of the fruits (analogous to data from consumption tests in prior studies of instrumental habits in animals and humans: e.g., Adams & Dickinson, 1981; Gillan et al., 2015; Luque et al., 2017). Overall, then, devaluation of the previously-high-value outcome did not seem to have an effect on attention to a signal of that outcome. Despite having explicit knowledge of the changed outcome values, participants did not appear to update the attentional priority of distractor colours to reflect that change.

Interestingly, providing trial-by-trial feedback on the identity of the fruit earned on each trial during the test phase (in the DevFB group) also did not result in a change in the attentional prioritisation of the high-value distractor in line with its current value. We postpone further examination of these findings in the feedback condition to the Discussion of Experiment 2.

Findings from the Dev group mirror those of a recent study by Watson, Pavri, Le, Pearson, & Le Pelley (2022). Using a similar approach to that employed here, but with food rather than monetary rewards, Watson et al. (2022) showed that devaluation of a specific food outcome (through feeding participants on that food to satiety) did not reduce the attentional priority of a distractor colour that signalled the availability of the now-devalued food in a test conducted under nominal extinction. These findings were also in line with previously discussed studies by De Tommaso et al. (2017, 2021) on the persistence of attentional biases to reward-signalling stimuli following devaluation by satiation on the food outcomes (conflicting findings are reported by Pool et al., 2014 potentially due to differences in methodology as discussed in introduction of this chapter). The findings of the current experiment extend these prior results. Data from the DevFB group demonstrate that attentional bias following outcome devaluation is also resistant to change in the face of direct experience of stimulus–outcome pairings under the new value regime, further underlining the resistance to updating attentional priorities.
2.3 Experiment 2

Considering the findings from Experiment 1, it is possible attentional prioritisation of the high-value distractor following outcome devaluation occurred because no substantial rewards were available during the test phase (as all outcomes now had low value), and attentional capture by distractors had consequently become of little consequence. That is, since attentional capture by distractors led to the omission of only a small reward (10 points) in the test phase, perhaps participants were no longer incentivised to recruit cognitive resources to change their pattern of responding from the training phase.

Hence, in Experiment 2 we examined whether prioritisation of the high-value distractor would persist if capture by distractors was always meaningful for participants’ earnings. By contrast to the devaluation manipulation of Experiment 1, in Experiment 2 we used a super-valuation manipulation to examine whether patterns of attentional prioritisation would change if both outcomes were associated with large rewards in the test phase – that is, if incentive to earn both outcomes was presumably high.

The training phase of Experiment 2 was as in Experiment 1: participants learned that one distractor signalled a high-value fruit (worth 500 points) and the other distractor signalled a low-value fruit (10 points). We again expected that participants’ attention would be captured more often by the high-value distractor than the low-value distractor, consistent with the idea that the high-value distractor was imbued with a greater ability to elicit an attentional prioritisation response than the low-value distractor. Following the training phase, participants were allocated to either the super-valued (Super) group or the super-valued-with-feedback (SuperFB) group. Participants in both groups were informed that the low-value fruit had increased in value, such that both fruits were thereafter worth 500 points. Both groups then completed the test phase, with trial-by-trial feedback on the fruit-outcome identity provided only to the SuperFB group.

As in Experiment 1, we were interested in whether or not conditioned attentional prioritisation of the high-value distractor that developed during the training phase would be sensitive to changes in the values of outcomes. If conditioned attentional responses were indeed mediated by representations of the fruit outcomes – and their new, equal (high) values – then the attentional bias formed in the training phase should disappear in the test phase. If instead conditioned attentional responses had become ‘habit-like’, operating independently of
representations of the outcomes, then we would expect the attentional bias toward the high-
value distractor to persist in the test phase. Performance of participants in the Super group
allowed us to examine whether knowledge of the new outcome values alone was sufficient to
induce a change in attentional bias. Performance of participants in the SuperFB groups
allowed us to further assess whether direct experience of pairings of distractor stimuli with
newly devalued outcomes during the test phase was necessary for a change in attentional
bias.

2.3.1 Methods and Materials

2.3.1.1 Participants and Apparatus

Fifty-five UNSW Sydney students participated in Experiment 2 (39 females; age $M = 19.18, SEM = 0.39$ years; Super group $n = 28$, SuperFB group $n = 27$). Group assignment
alternated based on order of arrival. All participants earned course credit and received a
monetary bonus dependent on their performance in the attentional revaluation task ($M = \text{A$9.95, SEM = A$0.18}$). Apparatus was as for Experiment 1.

2.3.1.2 Design and Procedure

The training phase was as for Experiment 1. Following the training phase, all
participants were told that the values of the fruits had changed (see Table 2.2). The fruit
previously worth 10 points in training was now worth 500 points, while the fruit previously
worth 500 points was still worth 500 points. Check questions were used to ensure that
participants had understood these instructions: participants were required to select the correct
current value of each fruit before they could proceed.

Participants in the Super group were further informed that while they would still be
earning fruits during the subsequent test phase, they would no longer be told the identity of
the fruit earned on each trial. Participants in the SuperFB group were informed that they
would continue to be told whether they had earned a lemon or a banana on each trial. In other
respects, the test phase was as in Experiment 1. As in Experiment 1, knowledge checks
following the test phase were used to assess participants’ explicit knowledge of the colour–
fruit contingencies, and of the current value of the fruit outcomes.
Table 2.2. Design of each phase for one counterbalance condition in Experiment 2. Colours refer to colours of the distractor in the search display; fruits refer to outcomes that could be won for saccades to the target. Note: fruit–value and colour–fruit contingencies were counterbalanced across participants.

<table>
<thead>
<tr>
<th>Initial Value Instructions</th>
<th>Training Phase</th>
<th>Revaluation Instructions</th>
<th>Group</th>
<th>Test Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lemon = 500 pts Banana = 10 pts</td>
<td>Blue → Lemon Orange → Banana</td>
<td>Lemon = 500 pts Banana = 500 pts</td>
<td>Super</td>
<td>Blue → ?? Orange → ??</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SuperFB</td>
<td>Blue → Lemon Orange → Banana</td>
</tr>
</tbody>
</table>

2.3.1.3 Data Preparation

Data preparation was as in Experiment 1. We discarded the first two trials after each break, time-out trials (1.07% of all trials), and trials with less than 25% valid gaze data (0.03% of all trials). Valid gaze data were registered on over 99% of the gaze-location samples in remaining trials.

2.3.2 Results

2.3.2.1 Training Phase

Proportion of Distraction Trials. Analysis of the proportion of distraction trials across Experiment 2 using a 2 (phase) × 2 (distractor-type) × 2 (group: Super vs. SuperFB) ANOVA revealed that the three-way interaction was not significant, $F(1,53) = 0.006, p = .94$, $\eta_p^2 < .001$. Planned analyses focused on the proportion of distraction trials in the training and test phases separately.

Figure 2.4A shows the proportion of distraction trials during the training phase, which was analysed using a distractor-type × group ANOVA. There was a significant main effect of distractor-type, $F(1,53) = 15.18, p < .001, \eta_p^2 = .22$, with more distraction trials when the display contained a high-value distractor than a low-value distractor. There was no significant main effect of group, $F(1,53) = 0.04, p = .84, \eta_p^2 < .001$, or interaction, $F(1,53) = 0.93, p = .34, \eta_p^2 = .02$. 
As in Experiment 1, we repeated the above analyses, restricting data from the training phase to only the last two blocks immediately prior to the value-switch manipulation – these analyses did not affect the pattern of significant findings. There was a significant main effect of distractor-type, $F(1,53) = 16.84, p < .001, \eta_p^2 = .24$, with more distraction trials when the display contained a high-value distractor than a low-value distractor. There was no significant main effect of group, $F(1,53) = 0.23, p = .63, \eta_p^2 = .004$, and critically no significant distractor-type × group interaction, $F(1,53) = 1.48, p = .23, \eta_p^2 = .03$. 

Figure 2.4. A) Proportion of distraction trials for trials featuring high- and low-value distractors in the training phase of Experiment 2; error bars show within-subjects SEM (Morey, 2008). Panel B) shows proportion of first-saccades in the direction of the high- and low-value distractors as a function of mean first saccade latency in the training phase of Experiment 2. Mean proportion of first saccades and first saccade latency (and their respective SEM error bars) were calculated separately for each time bin of individual-participant saccade latency distributions (fastest, middle, and slowest), defined by the tertiles of the distribution.
Distraction as a Function of Saccade Latencies. We examined the proportion of first saccades going towards distractors during the training phase of Experiment 2 using a 2 (distractor-type) × 2 (group) × 3 (latency time bin) ANOVA (see Figure 2.4B). As in Experiment 1, there were significant main effects of distractor type, $F(1,34) = 14.52, p < .001, \eta_p^2 = .30$, and latency bin, $F(2,68) = 109.65, p < .001, \eta_p^2 = .76$, and a significant distractor-type × latency bin interaction, $F(2,68) = 19.02, p < .001, \eta_p^2 = .36$. The latency bin × group interaction was also significant, $F(2,68) = 3.46, p = .037, \eta_p^2 = .09$, with the pattern of increased saccades toward the distractors at shorter latencies being more pronounced in the SuperFB than the Super group. Since both groups received equivalent treatment during the training phase, this difference must have been a quirk of random assignment and we do not interpret it further. No other effects involving group were significant, $F$s < 0.88, $p$s > .35, $\eta_p^2$s < .03.

2.3.2.2 Test Phase

Proportion of Distraction Trials. Figure 2.5A shows the proportion of distraction trials for each group in the test phase, following the super-valuation of the low-value fruit. A distractor-type × group ANOVA found a main effect of distractor-type, $F(1,53) = 11.33, p = .001, \eta_p^2 = .18$, with participants more likely to look at high-value than low-value distractors, even though during this test phase the outcome signalled by each type of distractor had the same (high) value. The main effect of group was not significant, $F(1,53) = 0.12, p = .73, \eta_p^2 = .002$. Importantly, the group × distractor-type interaction was not significant, $F(1,53) = 0.74, p = .39, \eta_p^2 = .014$.

Following up this non-significant interaction, Bayesian analysis of distractor difference scores from the test phase yielded a Bayes factor of $BF_{01} = 2.71$ in favour of the null hypothesis (no difference in attentional bias towards the high-value distractor between the two groups) over the alternative hypothesis. When the size of each group’s attentional bias during the training phase was further controlled for (by using distractor difference scores from the training phase as a covariate in a Bayesian one-way ANOVA), this analysis yielded $BF_{01} = 3.67$, suggesting moderate evidence in favour of the null hypothesis of no difference between the groups during the test phase (Lee & Wagenmakers, 2014).
Figure 2.5. A) Proportion of distraction trials for trials featuring high- and low-value distractors in the test phase of Experiment 2; error bars show within-subjects SEM (Morey, 2008). Panel B) shows proportion of first-saccades in the direction of the high- and low-value distractors as a function of mean first saccade latency in the test phase of Experiment 2 in the Super and SuperFB groups. Mean proportion of first saccades and first saccade latency (and their respective SEM error bars) were calculated separately for each time bin of individual-participant saccade latency distributions (fastest, middle, and slowest), defined by the tertiles of the distribution. Note that distractors are defined by the value of the outcome they signalled during the training phase.

Planned analyses of the simple effect of distractor type in each group during the test phase revealed a significant effect in both the Super group, $t(27) = 2.52, p = .018, d_z = .45$, and the SuperFB group, $t(26) = 2.37, p = .025, d_z = .46$. In both groups participants were significantly more likely to look at the distractor that previously signalled a high-value fruit in the training phase relative to the distractor that previously signalled a low-value fruit.
Distraction as a Function of Saccade Latencies. Figure 2.5B show proportion of first saccades that went towards distractors during the test phase, following the supervaluation of the low-value fruit, for the Super and SuperFB groups respectively. As in the training phase, ANOVA revealed main effects of distractor-type, $F(1,34) = 7.79, p = .009$, $\eta^2_p = .19$, and latency bin, $F(2,68) = 90.72, p < .001$, $\eta^2_p = .73$, and a significant distractor-type $\times$ latency bin interaction. There were no significant effects involving group, $Fs < 1.75$, $ps > .18$, $\eta^2_ps < .05$. Follow-up analyses restricted to the fastest saccade latency time bin revealed that the effect of distractor type was significant in the Super group, $t(19) = 3.24, p = .004$, $d_z = .73$, but did not reach significance in the SuperFB group, $t(15) = 1.76, p = .10$, $d_z = .44$.

A $2 \times 2$ ANOVA comparing the Super and SuperFB groups in the proportion of saccades made to the high- and low-value distractors in the fastest latency bin revealed a significant main effect of distractor-type, $F(1,34) = 12.22, p = .001$, $\eta^2_p = .26$. The main effect of group was not significant, $F(1,34) = 1.84, p = .18$, $\eta^2_p = .05$, nor was the interaction between group and distractor-type, $F(1,34) = 1.27, p = .27$, $\eta^2_p = .04$. Thus the pattern of performance for the fastest saccades did not differ significantly between the two groups.

2.3.2.3 Knowledge Checks

All participants correctly identified the current values of each fruit (the fruit–value contingencies) in the knowledge check following the test phase. Five participants in the Super group and three in the SuperFB group failed to correctly identify which fruit was signalled by each coloured distractor (the colour–fruit contingencies). Re-analysis of proportion of distraction trials while excluding these participants who failed to correctly identify the colour–fruit contingencies revealed that the patterns of significant and non-significant findings were unchanged.

2.3.3 Discussion

Findings from the training phase of Experiment 2 replicated findings from Experiment 1: participants were more likely to look at the high-value distractor than the low-value distractor, even though this pattern meant that high-value rewards were more often cancelled. Moreover, this pattern of greater attentional capture by high-value distractors (relative to low-value distractors) was particularly pronounced among participants’ fastest
saccades. Once again, this finding suggest that the high-value distractor stimulus was more likely to elicit an attentional prioritisation response than the low-value distractor.

As in Experiment 1, data from the test phase of Experiment 2 suggested that the pattern of attentional bias formed during training was not sensitive to a subsequent change in the value of the outcome implemented via an increase in the value of the low-value outcome in Experiment 2 (as opposed to a devaluation of the high-value outcome in Experiment 1). That is, participants in both groups continued to show a significant attentional bias towards the high-value distractor relative to the low-value distractor, even though both distractors now signalled a fruit outcome of equal (high) value. Results of latency-based analyses showed a similar pattern: rapid first saccades were more likely to be initiated toward the high-value distractor than the low-value distractor across both the training and test phases (though contrasts did not always reach significance, presumably as a consequence of additional participant exclusions reducing sample size coupled with analysis being restricted to only the fastest third of responses). Moreover – and as in the DevFB group of Experiment 1 – additional trial-by-trial feedback on the identity of the fruit earned on each trial in the SuperFB group did not result in a change in the attentional prioritisation of the high-value distractor in line with its current value. Overall, then, the super-valuation of a previously-low-value outcome did not seem to have an effect on attention to a signal of that outcome.

Findings of persistence of reward-related attentional bias in the face of outcome feedback following revaluation (in the DevFB and SuperFB groups of Experiments 1 and 2, respectively) are notable, as they are seemingly at odds with existing evidence from animal studies on habitual behaviour and sign-tracking. As noted earlier, animal studies of instrumental and sign-tracking behaviour have reported persistence of trained behaviour following outcome devaluation in tests conducted in extinction. By contrast, if (revalued) outcomes are provided in the test phase the typical finding is that conditioned behaviour rapidly alters in accordance with the new outcome values (e.g., lever-pressing rapidly ceases if doing so produces a now-devalued food pellet: see Amaya et al., 2020; de Wit et al., 2007; Dickinson et al., 1995) – presumably because direct experience of this outcome allows for new, goal-directed learning based on the outcome’s current reinforcement value. Indeed, such reacquisition tests—with outcome feedback provided—are often used precisely because they allow the researchers to verify that the devaluation procedure has been effective in changing the value of the outcome.
Could the failure to find an influence of outcome feedback on conditioned attention in the DevFB and SuperFB groups of Experiments 1 and 2 therefore be taken to suggest that our outcome revaluation procedure was ineffective? This seems unlikely. In Experiments 1 and 2, participants were explicitly told about the revaluation manipulation, reminded of it repeatedly, and required to pass check questions on the values of the fruits following revaluation. Moreover, data from the final tests of explicit knowledge showed that all participants could correctly report the current (updated) outcome values, verifying that participants’ knowledge spanned the whole of this critical phase.

We posit that there may be possible explanations for the differences between existing research in animals showing that behaviour is sensitive to changes in outcome value in reacquisition versus our findings on insensitivity of attentional prioritisation to changes in outcome value in our DevFB and SuperFB groups. First, in animal studies, receiving a now-devalued outcome (which has been paired with nausea-inducing compounds or consumed to satiety) during a reacquisition test comes with direct experience of the incentive value of that outcome – that is, its aversiveness – which would drive a subsequent reduction in responding. By contrast, in our experiments participants in the feedback groups continue to receive fruit outcomes that are not inherently aversive (in the case of the devalued outcome in Experiment 1) nor appetitive (in the case of the super-valued outcome in Experiment 2) but are merely symbolic outcomes worth high- or low-value. Hence, relative to animal paradigms, it is possible that participants in Experiments 1 and 2 had little drive to update their attentional control settings in the test phase – even with further feedback and explicit knowledge of their new values – since the fruit outcomes received were not experienced as ‘directly’ aversive nor appetitive.

2.4 Experiment 3

Related to the idea that participants in Experiments 1 and 2 had little motivation to change their patterns of attention (even with further outcome feedback following revaluation), a notable feature of both is that, following revaluation, both outcomes had the same value in the test phase: either both had low value (Experiment 1) or both had high value (Experiment 2). This equivalence of outcome values may have meant that participants did not strive to earn one particular fruit over the other. Hence there may have been insufficient incentive to drive effortful updating of attentional control settings to reflect the new fruit
values – regardless of whether or not outcome feedback was provided – such that the previous pattern of reward-related attentional capture persisted across groups. This possibility provided for the rationale for Experiment 3, in which the values of the two outcomes were reversed following the training phase: the (previously) high-value fruit was thereafter worth a small number of points, and the low-value fruit was worth a large number of points. So not only did the outcome values change following the training phase, but there was also a difference in their relative value during the test phase, which we hoped would provide a greater incentive for participants to exert the cognitive resources needed to update attentional control settings to reflect the changed outcome values.

Experiment 3 featured three groups of participants. Participants in the revalued (Rev) group experienced the value-switch outcome revaluation following the training phase: these participants were instructed regarding the changed values of the outcomes and were subsequently tested under nominal extinction. Participants in the revalued-with-feedback (RevFB) group experienced the same instructed value-switch, and continued to receive feedback in the test phase regarding the identity of the fruit earned on each trial. Finally, Experiment 3 also included a group of participants who did not undergo revaluation of the outcomes following the training phase (NoRev). Instead, following training these participants were instructed that the values of the fruits would remain the same during the subsequent test phase, and then completed the test under nominal extinction. Data from this NoRev group were used to provide a baseline against which to assess the effect of revaluation in the other groups (as opposed to comparing performance before and after revaluation as in Experiment 1 and 2).

2.4.1 Method

2.4.1.1 Participants and Apparatus

In total, 87 UNSW Sydney students (65 females; age \( M = 19.52, \text{SEM} = 0.31 \) years) participated in Experiment 3 for course credit, and received a monetary bonus depending on the points they earned in the attentional revaluation task (\( M = \text{A$9.83}, \text{SEM} = \text{A$0.23} \)). Participants were allocated to one of three groups (NoRev, Rev, RevFB: \( n = 29 \) per group) in three-way alternation based on their order of arrival.
Apparatus was similar to that for Experiment 1. For the first 34 participants, eye-tracking was via Tobii TX300 eye-trackers (sampling rate 300 Hz); subsequent participants were tested using Tobii Pro Spectrum eye-trackers (sampling rate 600 Hz). As in previous experiments, for all participants gaze data were down-sampled to 100 Hz for gaze-contingent calculations during stimulus presentation. Preliminary analyses indicated that the model of eye tracker had no significant effect on the pattern of results, and hence we collapsed across this factor in all analyses reported here.

2.4.1.2 Design and Procedure

The training phase was as for Experiment 1 and 2. Following the training phase, participants received further instructions depending on their group assignment (see Table 2.3). Participants in the NoRev group were reminded of the values of the fruits, which were unchanged from training (e.g., if lemons had previously been worth 500 points and bananas 10 points, participants in the NoRev group were simply reminded of these relationships). By contrast, participants in the Rev and RevFB groups were told that the values of the fruits had changed, such that the fruit previously worth 500 points in training was now worth 10 points, and the fruit previously worth 10 points was now worth 500 points. Check questions were used to ensure that participants had understood these instructions: participants were required to select the correct current value of each fruit before they could proceed.

Table 2.3. Design of each phase for one counterbalance condition in Experiment 3 (in which blue defined the high-value distractor and orange defined the low-value distractor). Colours refer to colours of the distractor in the search display; fruits refer to outcomes that could be won for a saccade to the target. Note: fruit–value and colour–fruit contingencies were counterbalanced across participants.

<table>
<thead>
<tr>
<th>Initial Value Instructions</th>
<th>Training Phase</th>
<th>Group</th>
<th>Revaluation Instructions</th>
<th>Test Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lemon = 500 pts, Banana = 10 pts</td>
<td>Blue → Lemon, Orange → Banana</td>
<td>NoRev</td>
<td>Lemon = 500 pts, Banana = 10 pts</td>
<td>Blue → ??, Orange → ??</td>
</tr>
<tr>
<td>Rev</td>
<td>Lemon = 10 pts, Banana = 500 pts</td>
<td>Rev</td>
<td>Blue → ??, Orange → ??</td>
<td></td>
</tr>
<tr>
<td>RevFB</td>
<td>Lemon = 10 pts, Banana = 500 pts</td>
<td>RevFB</td>
<td>Blue → Lemon, Orange → Banana</td>
<td></td>
</tr>
</tbody>
</table>
Participants in the NoRev and Rev groups were further informed that while they would still be earning fruits during the subsequent test phase, they would no longer be told the identity of the fruit earned on each trial. Participants in the RevFB group were informed that they would continue to be told whether they had earned a lemon or a banana on each trial. In other respects, the test phase was as in Experiments 1 and 2.

Following the test phase, participants completed a knowledge check assessing their understanding of which fruit was signalled by each colour of distractor (i.e., the colour–fruit contingencies). Due to a technical error, the knowledge check of fruit-point contingencies (how many points each fruit was currently worth) was omitted from Experiment 3. However, given that data from Experiments 1 and 2 showed that every participant could correctly identify the fruit values, it seems safe to assume that they would also have been able to do so here.

### 2.4.1.3 Data Preparation

Data preparation was as in previous experiments. We discarded the first two trials after each break, time-out trials (1.46% of all trials), and trials with less than 25% valid gaze data (0.42% of trials). Valid gaze data were registered in 99% of the gaze-location samples in remaining trials. As in previous experiments, in addition to analyses of the proportion of distraction trials (i.e., trials on which participants looked at the distractor, thereby cancelling reward), we also analysed the proportion of first saccades made towards the distractor as a function of the latency of those first saccades.

### 2.4.2 Results

#### 2.4.2.1 Training Phase

**Proportion of Distraction Trials.** We first examined the proportion of distraction trials across the training and test phases using a 2 (phase: training vs. test phase) × 2 (distractor-type: high- vs. low-value distractor) × 2 (group: NoRev, Rev, RevFB) ANOVA. Notably, the three-way interaction was significant, $F(2,84) = 10.94$, $p < .001$, $\eta^2_p = .21$, suggesting a significant difference in the distractor-type × group interaction between the training and test phases. Planned analyses focused on the proportion of distraction trials in the training and test phases separately. We first analysed proportion of distraction trials during the training phase, before the revaluation manipulation, using a 2 (distractor-type) × 3
(revaluation group) ANOVA (see Figure 2.6A). There was a significant main effect of distractor-type, $F(1,84) = 54.19$, $p < .001$, $\eta^2_p = .39$, with more distraction trials when the display contained a high-value distractor than a low-value distractor. There was no significant main effect of revaluation group, $F(2,84) = 0.12$, $p = .89$, $\eta^2_p = .003$, nor revaluation group $\times$ distractor-type interaction, $F(2,84) = 0.30$, $p = .75$, $\eta^2_p = .01$.

As in Experiments 1 and 2, we repeated the above analyses restricting data from the training phase to only the last two blocks immediately prior to the value-switch manipulation. These analyses did not affect the pattern of significant findings. There was a significant main effect of distractor-type, $F(1,84) = 77.38$, $p < .001$, $\eta^2_p = .48$, with a greater proportion of distraction trials when the display featured a high-value than a low-value distractor. There was no significant main effect of revaluation group, $F(2,84) = 0.12$, $p = .89$, $\eta^2_p = .003$, nor was there a revaluation group $\times$ distractor-type interaction, $F(2,84) = 0.15$, $p = .86$, $\eta^2_p = .004$. These null findings are unsurprising, since all groups received equivalent treatment until after the training phase.
Distraction as a Function of Saccade Latencies. We examined the proportion of first saccades going towards distractors during the training phase as a function of saccade latency, using a 2 (distractor-type: high- vs. low-value distractor) × 3 (revaluation group: NoRev, Rev, RevFB) × 3 (latency time bin: fastest, middle, slowest) ANOVA (Figure 2.6B). There was a significant main effect of distractor-type, $F(1,128) = 38.60, p < .001$, $\eta_p^2 = .38$, with more first saccades toward the high-value distractor than the low-value distractor. There was also a significant main effect of saccade latency time bin, $F(2,128) = 227.11, p < .001$, $\eta_p^2 = .78$, with shorter-latency saccades being more likely to go towards distractors. There was a significant interaction between these two factors, $F(2,128) = 26.49, p < .001$, $\eta_p^2 = .29$, with the bias towards the high-value distractor over the low-value distractor being more
pronounced at shorter saccade latencies. There were no significant effects involving revaluation group, \( F < 0.87, ps > .49, \eta^2_s < .03 \).

### 2.4.2.2 Test Phase

**Proportion of Distraction Trials.** Figure 2.7A shows the proportion of distraction trials for each revaluation group in the test phase, following the instructions regarding revaluation of the fruit outcomes. A 2 (distractor-type) × 3 (revaluation group) ANOVA revealed a significant main effect of distractor-type, \( F(1,84) = 4.77, p = .032, \eta^2_s = .05 \), while the main effect of revaluation group was not significant, \( F(2,84) = 0.26, p = .78, \eta^2_s = .006 \). Critically, we found a significant interaction between distractor-type and revaluation-group, \( F(2,84) = 8.38, p < .001, \eta^2_s = .17 \), indicating that the pattern of attentional bias differed across the three groups during the test phase.

To decompose this interaction, we first examined the simple effect of distractor-type in each group using paired \( t \)-tests. In the NoRev group, there was a significant effect of distractor-type, \( t(28) = 3.89, p < .001, d_z = 0.72 \), with participants more likely to look at the high-value distractor than the low-value distractor (difference scores are shown in Figure 2.7B). By contrast, in the Rev group, there was no significant effect of distractor type, \( t(28) = 1.14, p = .26, d_z = 0.21 \). A follow-up Bayesian \( t \)-test was used to assess evidence for the one-tailed alternative hypothesis of greater capture by the low-value distractor than the high-value distractor in this group (the pattern expected if attention were purely mediated by current outcome value). The resulting \( BF_{0+} = 9.97 \) represents moderate-to-strong evidence in favour of the null hypothesis over this alternative (Lee & Wagenmakers, 2014). Finally, in the RevFB group, participants were more likely to look at the low-value distractor than the high-value distractor, though this difference only approached significance, \( t(28) = 2.04, p = .051, d_z = 0.38 \).
Figure 2.7. A) Proportion of distraction trials for trials featuring high- and low-value distractors in the test phase of Experiment 3; error bars show within-subjects SEM (Morey, 2008). B) Distraction difference scores (given by the difference in the mean proportion of distraction trials between high- versus low-value distractor trials) for each group during the test phase. C) Proportion of first saccades in the direction of the high- and low-value distractors as a function of mean first saccade latency in the test phase of Experiment 3 in the NoRev, Rev, and RevFB groups. Mean proportion of first saccades and first saccade latency (and their respective SEM error bars) were calculated separately for each time bin of individual-participant saccade latency distributions (fastest, middle, and slowest), defined by the tertiles of the distribution. Note that distractors are defined by the value of the outcome they signalled during the training phase.

We next compared between-group differences in attentional capture by examining the interaction between revaluation group and distractor-type. First, the Rev group and NoRev group differed only in whether or not the values of the fruits changed from the training phase to the revaluation phase (neither group received feedback on the specific fruit won on each trial during the test phase). Hence, differences in patterns of attentional capture between these two groups during the test phase provided an index of the effect of outcome revaluation on attentional capture. A $2 \times 2$ ANOVA revealed that the revaluation group $\times$ distractor-type approached significance, $F(1,56) = 3.56, p = .064, \eta^2_p = .06$, with a trend towards a greater
difference in capture by high-value versus low-value distractors in the NoRev group than the Rev group (Figure 2.7B).

Our second comparison of interest was between the Rev and RevFB groups. These groups differed only in whether or not trial-by-trial feedback on specific fruit outcomes was provided during the test phase (both groups experienced the revaluation of fruit outcomes). Hence, any difference in patterns of attentional capture between these two groups provided an index of the effect of feedback following outcome revaluation. Here, a 2 × 2 ANOVA found a significant revaluation group × distractor-type interaction, \(F(1,56) = 4.52, p = .038, \eta_p^2 = .075\), with the RevFB group exhibiting a greater difference in attentional capture by high-versus low-value distractors than the Rev group (Figure 2.7B).

Finally, ANOVA comparing the NoRev and RevFB groups revealed a highly significant interaction, \(F(1,56) = 18.78, p < .001, \eta_p^2 = .251\), in line with the very different pattern of attentional bias observed during the test phase in these groups (Figure 2.7B).

**Distraction as a Function of Saccade Latencies.** Figure 2.7C shows the proportion of first saccades to the distractors in the test phase as a function of saccade latency. A 2 (distractor-type) × 3 (revaluation group) × 3 (latency time bin) ANOVA revealed a significant main effect of latency bin, \(F(2,128) = 143.95, p < .001, \eta_p^2 = .69\), with shorter-latency saccades more likely to go towards distractors. Notably a significant three-way interaction, \(F(4,128) = 9.09, p < .001, \eta_p^2 = .22\), indicated that the latency-modulated pattern of attentional bias differed across the three groups. There was also a significant distractor type × revaluation group interaction, \(F(2,128) = 9.63, p < .001, \eta_p^2 = .23\). No other main effects or interactions were significant, \(F_s < 2.8, ps > .09, \eta_p^2s < .05\).

As for Experiments 1 and 2, follow-up analyses focused on data from the fastest saccade latency time bin. Paired \(t\)-tests revealed a significant effect of distractor-type in the NoRev group, \(t(20) = 4.20, p < .001, d_z = 0.92\), with fastest saccades more often initiated toward the high-value distractor than the low-value distractor. In the Rev group, however, there was no significant effect of distractor type, \(t(22) = 0.60, p = .56, d_z = 0.12\); a Bayesian \(t\)-test assessing evidence for the one-tailed alternative hypothesis of greater rapid capture by the low-value distractor than the high-value distractor yielded \(BF_{0+} = 6.77\), representing moderate evidence in favour of the null hypothesis over this alternative (Lee & Wagenmakers, 2014). Finally, in the RevFB group the fastest saccades were significantly
more likely to move toward the low-value distractor than the high-value distractor, \( t(22) = 2.27, p = .02, d_z = 0.54 \).

To test for between-group differences in attentional bias among the fastest saccades, we ran a series of 2 (revaluation group) × 2 (distractor-type) ANOVAs using the data from the fastest saccade latency time bin. Analysis contrasting the NoRev and Rev groups revealed a significant group × distractor-type interaction, \( F(1,42) = 8.64, p = .005, \eta_p^2 = .17 \), with the NoRev group showing a greater attentional bias in rapid saccades to high-value versus low-value distractors than the Rev group. Analysis contrasting the Rev and RevFB groups again yielded a significant interaction, \( F(1,44) = 4.80, p = .03, \eta_p^2 = .10 \), indicating that provision of outcome feedback in the test phase led to a significant difference in the resulting pattern of attentional bias following outcome revaluation. Finally, contrast of the NoRev and RevFB groups also revealed a significant group × distractor-type interaction, \( F(1,42) = 24.63, p < .001, \eta_p^2 = .37 \), in line with the reversal in the pattern of gaze bias in the RevFB group relative to the NoRev group.

2.4.2.3 Knowledge Check

In each of the three groups, 26 out of 29 participants reported the correct colour–fruit associations in the knowledge check at the end of the experiment. We repeated the previously described analyses of proportion of distraction trials while excluding the nine participants who failed to correctly identify the colour–fruit contingencies. These analyses did not affect the pattern of significant findings, with the exception that the pattern of attentional bias in the proportion of distraction trials during the test phase now differed significantly between Rev and NoRev groups (\( F[1,50] = 6.22, p = .016, \eta_p^2 = .111 \)), where previously this difference had only approached significance (\( p = .064 \)).

2.4.3 Discussion

Experiment 3 examined the effect of switching the values of outcomes on attentional prioritisation of distractor stimuli. As in Experiments 1 and 2, during the training phase participants in all three groups were more likely to have their attention captured by a distractor signalling a high-value fruit outcome versus a low-value fruit, with this pattern particularly pronounced among the fastest saccades. For participants in the NoRev group, there was no change in the value of the fruit outcomes between training and test phases: these
participants were instructed that the fruit that had high value in the training phase continued to have high value in the test phase, and vice versa. Participants in this group continued to show greater attentional capture by the high-value distractor than the low-value distractor during the test phase, even though no feedback was provided as to the specific fruit earned on each trial of the test phase. Thus for the NoRev group, the attentional bias established during the training phase persisted during the test phase, even in the absence of further feedback (cf. Watson et al., 2019).

Crucially, for participants in the Rev group, the values of the fruit outcomes switched between the training and test phases; the fruit that had high value in the training phase thereafter had low value, and the fruit that had low value in the training phase had high value. During the test phase, these participants no longer exhibited significantly greater attentional capture by the high-value distractor than the low-value distractor, either in the data relating to proportion of distraction trials or the direction of their fastest saccades. Notably, the difference in pattern of attentional bias between NoRev and Rev groups was significant in focused analyses of the fastest saccades, and also in broader analyses of the proportion of distractor trials when we excluded participants who failed the explicit knowledge check regarding stimulus-outcome relationships (the difference in proportion of distraction trials only approached significance in the full sample).

However, the pattern of attention in the Rev group did not show a full reversal as might be expected based on the change in outcome values. Participants in this group were aware of the new values of the fruits in the test phase, since they were explicitly told this information, reminded of it repeatedly, and required to pass check questions on the values of the fruits following revaluation. Moreover, almost all participants in the Rev group correctly reported which fruit outcome was associated with each colour of distractor in the final knowledge test. Yet, in the test phase, attentional prioritisation and rapid orienting toward the high-value distractor was not completely overcome to follow the changed values of the fruits. This finding suggests that in the Rev condition, there is an outcome-mediated component of attention that attempted to update attentional control, but was counteracted by an inflexible ‘habit-like’ process that continued to prioritise the (previously) high-value distractor in an outcome-independent manner.

Like the Rev group, participants in the RevFB group experienced a switch in the values of the fruit outcomes between the training and test phases. Notably, however,
participants in the RevFB group also received feedback on the identity of the fruit earned on each trial of the test phase. During the test phase, these participants showed a reversal of the pattern formed during training; they were significantly more likely to make rapid saccades towards the low-value distractor than the high-value distractor, in line with the current values of the fruit outcomes (the corresponding pattern in overall proportion of distraction trials also closely approached significance, \( p = .051 \)).

There was a significant difference in the pattern of attentional bias between the Rev and RevFB groups during the test phase. Both of these groups underwent outcome revaluation; the only difference was that participants in the RevFB group continued to receive trial-by-trial feedback during the test phase, whereas the Rev group did not. Although participants in the Rev group had knowledge of the new values of the fruit outcomes, and knowledge of which fruit outcome was associated with each distractor, this knowledge alone was not sufficient to overturn the attentional bias toward the high-value distractor. Only with the additional experience of the pairings of distractor colours and revalued outcomes was a full reversal of the attentional prioritisation response toward high- and low-value distractors observed (in the RevFB group). Together, the findings suggest that in the Rev group, the attentional prioritisation response had become partially divorced from the value of the outcome, failing to show a full switch despite the reversal of outcome values, and updating of this ‘habit-like’ component of conditioned attention for full reversal of the pattern established in training requires further training through outcome feedback. It appears that such an attentional shift must be triggered by learning and experience of stimuli and outcomes (and their associated values).

2.5 Experiment 4

To recap, data from Experiments 1 to 3 suggested that the degree of sensitivity to changes in outcome value depended on the specific revaluation regime. Under conditions of devaluation in Experiment 1 or ‘super-valuation’ in Experiment 2, attentional prioritisation of the high-value distractor persisted, even though it now signalled an outcome that was of equal low- or high-value to the outcome signalled by the low-value distractor. Strikingly, this persistence was seen even in groups who received feedback on the identity of the outcome earned on each trial following revaluation. By contrast, findings from Experiment 3 revealed that when outcome values were switched between training and test, the distractor that had
signalled a high-value outcome during initial training was no longer more likely to elicit attentional prioritisation than the low-value distractor. That is, there was an outcome-mediated component of attention which updated attentional prioritisation of the high-value versus low-value distractor. However, this updating was opposed by an inflexible outcome-independent component which continued to prioritise the distractor that previously signalled the high-value outcome; in sum there was no longer a significant bias towards the high-value distractor, although the pattern of attentional prioritisation did not show a ‘full’ switch in line with the new value regime. The habit-like component of conditioned attention was overcome only with the provision of further outcome feedback – that is, experience of the distractor-outcome pairings – in the test phase.

One shared feature of Experiments 1–3 is that participants are (presumably) tracking the relationships between coloured distractor stimuli, fruit outcomes, and point values. For Experiment 3 in particular, one way of viewing the value-reversal manipulation is that it rendered the high-value outcome devalued, and the low-value outcome ‘super-valued’, as opposed to Experiments 1 and 2, where participants experienced only one of these changes in outcome value. Considering the example of one counterbalance condition in the Rev group of Experiment 3, participants would have learned in training that the blue distractor signals a banana worth 500 points, while the orange distractor signals a lemon worth only 10 points. Then in the instructed value-reversal, participants in this condition would have been told that the values of the banana and lemon outcomes have reversed, such that bananas are worth only 10 points and lemons are worth 500 points. For this instructional manipulation to result in a goal-directed change in behaviour during the test phase, participants would have to work through this associative structure (from values to outcomes to stimuli signalling those outcomes) to understand that the blue distractor signals an outcome of now-low-value, and the orange distractor signals an outcome of now-high-value, in order to update their attentional control settings in line with the new fruit values. It has been argued in the literature on instrumental habits that performance in devaluation tasks may be influenced by the ease of integrating updated outcome-value information into the existing associative structure (Buabang et al., 2020). Hence, one possible explanation for why participants in the Rev group showed only partial updating in Experiment 3 could be related to their capacity (e.g., working memory capacity, or motivation) to update their attentional control settings under this relatively complex associative structure, rather than the influence of an outcome-mediated process counteracted by a ‘habit-like’ outcome-independent process.
To examine this possible account, we investigated whether the findings of partial value-sensitivity observed in Experiment 3 would generalise to a situation involving a simpler associative structure. In Experiment 4, we simplified the task by omitting the mediating fruit outcomes—participants were directly rewarded with points for correctly responding to the diamond target, with the colour of distractors signalling the magnitude of point rewards (i.e., high- versus low-points) available on each trial. It is important to note that by removing the mediating fruit outcomes and having the values themselves as outcomes, it was not possible to change outcome value and identity independently. Hence this simplified associative structure provided a less pure test of the effect of ‘habit-like’ processes drawn from the study of traditional instrumental habits, relative to Experiments 1 to 3. Nonetheless, Experiment 4 was conducted with the aim of examining the alternative possibility that findings in Experiment 3 were a consequence of the specific (more complex) associative structure used there, wherein participants had to integrate instructed knowledge about outcome value with learned knowledge of the relationships between stimuli and outcomes.

As in Experiment 3, the current experiment featured three groups—NoRev, Rev, and RevFB groups. Following training under this simplified colour-point structure, participants in the Rev and RevFB groups were instructed that the colour-reward relationships had reversed, such that the high-value distractor thereafter signalled the availability of a small number of points (and vice versa for low-value distractor). The Rev group were tested under nominal extinction, while the RevFB group continued to be provided with trial-by-trial feedback on the points earned on each trial. By contrast, the NoRev group did not undergo the reversal of colour-point relationships, and then completed the test under nominal extinction. Data from this NoRev group were used to provide a baseline against which to assess the effect of revaluation in the other groups (as in Experiment 3).

Notably, due to restrictions resulting from the COVID-19 pandemic, we were unable to test participants face-to-face in Experiment 4, and hence could not use the eye-tracking procedure that we had previously employed in Experiments 1 to 3. Consequently in Experiment 4, we instead used a variant of the attentional revaluation task that could instead be completed by participants online. This procedure was based on a version of the attentional revaluation task in which the critical measure was response time (RT) to make a manual keypress based on the identity of the target. Previous research using this task, including online studies (e.g., Le Pelley et al., 2015, 2022; Watson, Pearson, Most, et al., 2019) has
found that participants are slower (but not more accurate) to respond to the target when the display contains a distractor signalling availability of high-value reward than a distractor signalling low-value reward, even though the task is arranged so that slower responses result in lower earnings (somewhat analogous to the omission of reward caused by looking at the distractor in Experiments 1 to 3). The implication is that the high-value distractor was more likely to capture participants’ attention than the low-value distractor, hence slowing responses to the target: a VMAC effect.

In Experiment 4, we incorporated a value-reversal manipulation into the RT variant of the VMAC task. Participants were initially informed – and learned through experience in the training phase – the colour of the distractor in the search display on each trial indicated whether the current trial was a low- or high-reward trial. In line with previous work on the VMAC effect, we expected participants to show slower (but not more accurate) responses to the target when the display contained the high-value distractor relative to the low-value distractor in the training phase. Following the training phase, participants in the Rev and RevFB groups were then informed that the distractor-reward contingencies had switched, such that the high-value distractor now signalled the availability of a low-value reward, and vice versa. Participants in the Rev group were subsequently tested under nominal extinction, while participants in the RevFB group continued to receive feedback in the test phase regarding the reward on each trial. As in Experiment 3, participants in the NoRev group did not undergo the reward reversal following the training phase; they were instead instructed that the distractor-reward contingencies would remain the same during the subsequent test phase, and then completed the test under nominal extinction. Data from this NoRev condition were used to provide a baseline against which to assess the effect of revaluation in the other groups.

If participants in the Rev group showed a reversal of the attentional bias formed in training (i.e., slower but not more accurate responses when the display contained a low-value distractor relative to a high-value distractor) in line with the reversed reward contingencies, then this would suggest that attentional control settings could be updated under the simplified associative structure of Experiment 4 where the identity of the outcome specified the value of that outcome. This pattern of results would be in line with the account that previous findings of only partial sensitivity to changes in outcome value were a consequence of the complexity of the associative structure in Experiment 3 (where outcome identity and value were distinct
properties that could be varied independently). By contrast, if participants in the Rev group failed to switch their patterns of attention in line with the reward-switch, even under the simplified task structure, then this would strengthen the idea that outcome-independent ‘habit-like’ processes may indeed be operating to counteract updating of attentional control settings.

2.5.1 Method

2.5.1.1 Participants and Apparatus

In total, 148 students from UNSW Sydney (108 females; age $M = 19.97$ years, $SEM = 0.28$ years) completed Experiment 4 for course credit. Participants who performed in the top 50% also earned a $10$ AUD bonus dependent on their performance (see Design and Procedure). Eighteen participants were excluded from all analyses (see Data Preparation); the remaining participants were randomly allocated to either the NoRev ($n = 37$), Rev ($n = 42$), or RevFB group ($n = 51$).

The task was programmed using jsPsych (de Leeuw, 2015), a JavaScript library that allows behavioural experiments to be conducted over the internet. These scripts are able to access high-performance native system components on the local machine, allowing for the collection of millisecond-precision RTs. Participants completed the task on their own desktop devices (the task was not compatible with tablets or phones).

2.5.1.2 Design and Procedure

As in Experiments 1 to 3, the task consisted of three main components: the training phase, revaluation instructions, and test phase.

Training Phase. At the start of the experiment, participants were told that their aim should be to earn as many points as possible, as the top 50% of participants who received the most points would be given a bonus of $10$ AUD. In order to earn points, participants were informed that they had to respond quickly and accurately in a visual search task.

Each trial of the training phase of the search task consisted of a fixation display, search display, and feedback display; all stimuli were presented on a black background. All trials began with a central fixation cross, on which participants were informed to fixate (see Figure 2.8). After a random interval of 300–500 ms, the fixation cross disappeared and was
followed by a blank screen for 150ms. The search display then appeared, consisting of the diamond target and 5 circles distributed evenly around the centre of the screen, each containing a line segment. The line inside the diamond was always oriented either horizontally or vertically, while the circles contained a line tilted 45° randomly to the left or right; orientations of lines and locations of the shapes were selected randomly on each trial. Importantly, one of the circles was either an orange or blue distractor on each trial, while all other circles were grey.

![Diagram of trial events](image)

*Figure 2.8. Example of trial events in the RT version of the attentional revaluation task. Each trial began with a central fixation cross. Following a blank interval, a search display appeared, containing a diamond-shaped target among circles. The faster that participants responded (correctly) according to the orientation of the line segment inside the diamond, the more points they could earn on that trial. Notably, the colour of a colour-singleton distractor circle in the display signalled the size of the available reward: if the display contained a distractor rendered in the high-value colour, the trial was a bonus trial on which points were multiplied by 10; if the distractor was rendered in the low-value colour, the trial was a standard (non-bonus) trial. The example here shows a high-value distractor trial, in which an orange distractor signals a 10x bonus trial. Note: the colour-reward contingencies were counterbalanced across participants.*

Participants’ task was to respond to the orientation of the line segment inside the diamond target by pressing the ‘C’ key (if the line was horizontal) or the ‘M’ key (if the line was vertical). The faster participants made a correct response below 1000 ms, the more points they could earn on that trial. Notably, the colour of the distractor in the search display on each trial indicated whether the current trial was a low- or high-reward trial, with the colour-reward relationships counterbalanced across all participants. On trials featuring the low-value
distractor, participants could earn 1 point for every 10 ms they responded below 1000 ms; for example, an RT of 550 ms would earn 45 points. Trials featuring the high-value distractor were marked as ‘bonus trials’ on which points were multiplied by 10; so here an RT of 550 ms would earn 450 points. Participants were explicitly informed of these colour-reward contingencies at the outset (e.g., that if the display contained an orange circle it would be a 10x bonus trial, whereas if the display contained a blue distractor it would not, and they could earn only a small number of points). Participants were also told that if their RT was over 1000 ms, or if they made an error, they would not earn any points.

The feedback display replaced the search display only once a response was registered. If the response was correct, feedback showed “Correct” and the number of points earned on that trial; for high-reward trials, feedback additionally included the text “10x BONUS TRIAL”. If RT was above 1000ms, feedback showed “Too Slow: 0 points”, and for incorrect responses, feedback showed “Error: 0 points”. Feedback was displayed for 700ms, followed by a blank inter-trial interval of 1000ms.

In total, participants completed 20 blocks of trials in the training phase, with each block containing 24 trials (480 trials in total). Half of the trials in each block featured a high-reward distractor, and the other half featured a low-reward distractor: trial order in each block was random. At the end of each block participants took a short break, during which they were informed of the total number of points they had earned so far. During each break, they were also reminded that the faster they responded on each trial, the more points they would earn.

Revaluation and Feedback Instructions. Immediately following the training phase, participants received further instructions depending on their group assignment. Participants in the NoRev group were reminded of the colour-reward relationships, which were unchanged from training (e.g., if the blue distractor had previously been the high-reward colour and orange had been the low-reward colour, participants in the NoRev group were simply reminded of these relationships). As shown in Table 2.4, participants in the Rev and RevFB groups were informed that the allocation of high- and low-reward colours had switched, such that the previously high-reward colour was now associated with low-reward, and the low-reward colour was now associated with high-reward (that is, it signalled 10x bonus trials).

In addition, participants in the NoRev and Rev groups were instructed that while they would still be earning points during the subsequent test phase, they would no longer be told
whether they had earned a small number of points or a 10x bonus number of points on each trial; “we will only tell you whether you made a correct or incorrect response. You will still be earning these points, but we will keep track of your total”. Participants in the RevFB group were told that “as before, we will tell you how many points you earned on each trial”. After this instruction, all participants answered check questions to ensure their knowledge of the two current colour-reward associations and feedback structure. All three responses were required to be correct before they could proceed.

Table 2.4. Design of each phase in Experiment 4 for one counterbalance condition (in which blue defined the high-value distractor and orange defined the low-value distractor). Colours refer to colours of the distractor in the search display. Note: colour–reward contingencies were counterbalanced across participants.

<table>
<thead>
<tr>
<th>Initial Value Instructions</th>
<th>Training Phase</th>
<th>Group</th>
<th>Revaluation Instructions</th>
<th>Test Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue → 10x bonus Orange → No bonus</td>
<td>Blue → 10x bonus Orange → No bonus</td>
<td>NoRev</td>
<td>Blue → 10x BONUS Orange → No bonus</td>
<td>Blue → ?? Orange → ??</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev</td>
<td>Blue → No bonus Orange → 10x BONUS</td>
<td>Blue → ?? Orange → ??</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RevFB</td>
<td>Blue → No bonus Orange → 10x BONUS</td>
<td>Blue → No bonus Orange → 10x BONUS</td>
</tr>
</tbody>
</table>

Test Phase. All participants then completed the test phase of the search task. In the test phase, participants in the Rev and NoRev groups continued to earn points for responding correctly, but did not receive feedback on whether they earned a high-reward or low-reward. If RT was below 1000 ms and it was not a distraction trial, then feedback stated “Correct”, accompanied by “??” where previously the number of points earned on that trial appeared during the training phase. If the response was incorrect, feedback stated “Error”. If RT was above 1000 ms, feedback stated “Too slow”. Participants in the RevFB group received feedback on the number of points earned – or omitted – on each trial (high- or low-reward), as in the training phase. Participants completed 8 blocks of trials in the test phase, with blocks structured as in the training phase (192 trials in total). Participants in all groups were informed of the total number of points they had earned so far during breaks at the end of each block.
Knowledge Check. Following the test phase, participants’ knowledge of the colour–point contingencies was assessed. Participants were told that the number of points that could be won on each trial depended on the colour of the coloured circle in the search display. For each distractor colour, participants were asked whether they earned a small number of points or a 10x bonus number of points for making a correct response.

2.5.1.3 Data Preparation

Data analysis followed established protocols for analysing RTs from this version of the task (e.g., Le Pelley et al., 2022; Watson, Pearson, Most, et al., 2019). We discarded data from the first two trials after each break, trials with responses that were faster than 150ms (indicating anticipatory responding), and trials with responses slower than 1500ms. Eighteen participants who had more than 25% of trials excluded based on these criteria, or whose mean accuracy was low (<65% correct) were excluded from all analyses.

Following previous work, the main dependent variable of interest was participants’ RT (for correct responses only), on the grounds that attentional capture by distractors would lead to slower responses to the target. We analysed RT as a function of whether the trial featured a high- or low-value distractor; note that (as in previous experiments) we label distractors according to the value they signalled during the training phase. To examine potential trade-offs between speed and accuracy of responding, our secondary analyses focused on proportion of errors as a function of distractor-type.

2.5.2 Results

2.5.2.1 Training Phase

Response Times. We first examined RTs across the training and test phases using a 2 (phase: training vs. test phase) × 2 (distractor-type: high- vs. low-value distractor) × 2 (group: NoRev, Rev, RevFB) ANOVA. The three-way interaction was not significant, $F(2,127) = 10.94, p = .17, \eta^2_p = .03$; nevertheless, following previous experiments, planned analyses focused on RTs in the training and test phases separately. We analysed mean RTs during the training phase using a 2 (distractor-type) × 3 (revaluation group) ANOVA (see Figure 2.9A). There was a significant main effect of distractor-type, $F(1,127) = 71.58, p < .001, \eta^2_p = .36$, with slower RT when the display contained a high-value distractor than a low-value distractor. There was no significant main effect of revaluation group, $F(2,127) = 0.78, p =$
.46, $\eta_p^2 = .01$, nor revaluation group $\times$ distractor-type interaction, $F(2,127) = 0.70$, $p = .50$, $\eta_p^2 = .01$.

We repeated the above analyses restricting data from the training phase to only the last two blocks immediately prior to the value-switch manipulation. These analyses did not affect the pattern of significant findings. There was a significant main effect of distractor-type, $F(1,127) = 23.08$, $p < .001$, $\eta_p^2 = .15$. There was no significant main effect of revaluation group, $F(2,127) = 0.70$, $p = .50$, $\eta_p^2 = .01$, nor was there a revaluation group $\times$ distractor-type interaction, $F(2,127) = 2.08$, $p = .13$, $\eta_p^2 = .03$. These null findings are unsurprising, since all groups received equivalent treatment until after the training phase.

**Figure 2.9.** A) Mean response times (RT) and B) proportion of errors on trials with high- and low-value distractors in the training phase of Experiment 4. Error bars show within-subjects SEM (Morey, 2008)
Proportion of Errors. We examined proportion of errors across the training and test phases using a 2 (phase: training vs. test phase) × 2 (distractor-type: high- vs. low-value distractor) × 2 (group: NoRev, Rev, RevFB) ANOVA. The three-way interaction was marginally significant, \( F(2,127) = 2.57, p = .08, \eta_p^2 = .04 \). Planned analyses focused on mean proportion of errors in the training and test phases separately. As shown in Figure 2.9B, we analysed proportion of errors during the training phase using a 2 (distractor-type) × 3 (revaluation group) ANOVA. There was a significant main effect of distractor-type, \( F(1,127) = 10.16, p = .002, \eta_p^2 = .07 \), with a greater proportion of errors when the display contained a high-value distractor than a low-value distractor. There was no significant main effect of revaluation group, \( F(2,127) = 0.18, p = .84, \eta_p^2 < .01 \), nor revaluation group × distractor-type interaction, \( F(2,127) = 0.97, p = .38, \eta_p^2 = .02 \).

2.5.2.2 Test Phase

Response Times. Figure 2.10A shows the mean RTs for each revaluation group in the test phase. A 2 (distractor-type) × 3 (revaluation group) ANOVA revealed a significant main effect of distractor-type, \( F(1,127) = 5.24, p = .024, \eta_p^2 = .04 \), while the main effect of revaluation group was not significant, \( F(2,127) = 0.67, p = .51, \eta_p^2 = .01 \). Importantly, we found a significant interaction between distractor-type and revaluation group, \( F(2,127) = 4.57, p = .012, \eta_p^2 = .07 \), indicating that differences in RTs on high- versus low-value distractor trials differed across the three groups during the test phase.

To decompose this interaction, we examined the simple effect of distractor-type in each group using paired t-tests. In the NoRev group, there remained a significant effect of distractor-type, \( t(36) = 3.60, p < .001, d_z = 0.60 \), with slower responses with the display contained a high-value distractor than a low-value distractor (difference scores are shown in Figure 2.10B). In the Rev group, there was no significant effect of distractor type, \( t(41) = 1.18, p = .24, d_z = 0.18 \). A follow-up Bayesian t-test was used to assess evidence for the one-tailed alternative hypothesis of slower RTs on low-value distractor trials than the high-value distractor trials in this group (the pattern expected if attention were purely mediated by current reward value). The resulting \( BF_{01} = 12.16 \) represents strong evidence in favour of the null hypothesis over this alternative (Lee & Wagenmakers, 2014). Finally, in the RevFB group there was a numeric trend towards slower responses for the low-value distractor than the high-value distractor, but this trend did not approach significance, \( t(50) = 0.69, p = .49, d_z = 0.10 \).
We next compared between-group differences in performance during the test phase by examining the interaction between revaluation group and distractor-type. As in Experiment 3, comparing performance between the Rev group and NoRev group provided an index of the effect of outcome revaluation on attentional capture. A $2 \times 2$ ANOVA revealed that the revaluation group $\times$ distractor-type interaction approached significance, $F(1,77) = 3.24, p = .076, \eta^2_p = .04$, with a trend towards a greater difference in RTs when the display contained a high-value versus low-value distractors in the NoRev group than the Rev group (Figure 2.10B). Next, comparing performance between the Rev and RevFB groups provided an index of the effect of feedback following outcome revaluation. Here, a $2 \times 2$ ANOVA revealed the revaluation group $\times$ distractor-type interaction was not significant, $F(1,91) = 1.67, p = .20, \eta^2_p = .02$. Finally, ANOVA comparing the NoRev and RevFB groups revealed a significant interaction, $F(1,86) = 8.58, p = .004, \eta^2_p = .09$, in line with the very different pattern of attentional bias observed during the test phase in these groups (Figure 2.10B).
Figure 2.10. A) Mean response times (RT) on trials with high- and low-value distractors in the test phase of Experiment 4, and B) shows RT difference scores (given by the difference in the mean RTs between high- versus low-value distractor trials) for each group during the test phase. C) Proportion of errors on trials with high- and low-value distractors in the test phase, and D) shows error difference scores (given by the difference in proportion of errors between high- versus low-value distractor trials) for each group during the test phase. Error bars in Panels A) and C) show within-subjects SEM (Morey, 2008) and in Panels B) and D) show between-subjects SEM. Note that distractors are defined by the value of the outcome they signalled during the training phase.

We next compared between-group differences in performance during the test phase by examining the interaction between revaluation group and distractor-type. As in Experiment 3, comparing performance between the Rev group and NoRev group provided an index of the effect of outcome revaluation on attentional capture. A 2 × 2 ANOVA revealed that the reevaluation group × distractor-type interaction approached significance, $F(1,77) = 3.24$, $p = .076$, $\eta^2_p = .04$, with a trend towards a greater difference in RTs when the display contained a high-value versus low-value distractors in the NoRev group than the Rev group (Figure 2.10B). Next, comparing performance between the Rev and RevFB groups provided an index
of the effect of feedback following outcome revaluation. Here, a $2 \times 2$ ANOVA revealed the revaluation group × distractor-type interaction was not significant, $F(1,91) = 1.67, p = .20, \eta^2_p = .02$. Finally, ANOVA comparing the NoRev and RevFB groups revealed a significant interaction, $F(1,86) = 8.58, p = .004, \eta^2_p = .09$, in line with the very different pattern of attentional bias observed during the test phase in these groups (Figure 2.10B).

**Proportion of Errors.** Figure 2.10C shows the mean proportion of errors for each revaluation group in the test phase. A $2$ (distractor-type) × $3$ (revaluation group) ANOVA found non-significant main effects of distractor-type, $F(1,127) = 0.56, p = .46, \eta^2_p = .004$, and revaluation group, $F(2,127) = 1.27, p = .29, \eta^2_p = .02$, and a non-significant interaction, $F(2,127) = 2.23, p = .11, \eta^2_p = .03$. Nevertheless, we conducted planned analyses of the simple effect of distractor-type on proportion of errors in each group using paired $t$-tests. There was no significant effect of distractor-type in the NoRev group $t(36) = .62, p = .54, d = 0.10$, nor the Rev group, $t(41) = .16, p = .88, d = 0.02$. Notably, however, the simple effect of distractor type was significant for the RevFB group, $t(50) = 2.31, p = .025, d = 0.32$, indicating a greater proportion of errors on low-value trials than high-value trials (Figure 2.10D).

**2.5.2.3 Knowledge Check**

One participant in the Rev group and two participants in the RevFB group reported incorrect colour-point associations in the knowledge check at the end of the experiment. We repeated the previously described analyses of proportion of distraction trials while excluding these participants. These analyses did not affect the pattern of significant findings.

**2.5.3 Discussion**

Experiment 4 used an online, RT-based version of the attentional revaluation task to examine the generalisability of sensitivity to changes in outcome value under a simplified associative structure. Specifically, we were interested in the effect of switching values of points on attentional prioritisation of distractor stimuli, while omitting the mediating fruit outcomes that were featured in Experiments 1 to 3.

Consistent with data from previous experiments using this version of the VMAC task (Le Pelley et al., 2015, 2022; Watson, Pearson, Most, et al., 2019), during the training phase participants were slower (and less accurate) to respond to the target when the display featured
a high-value distractor relative to a low-value distractor. These data provided a conceptual replication of findings from the training phases of Experiments 1 to 3, suggesting that the distractor that signalled a high-value reward was more likely to capture attention – slowing responding and eliciting more errors – than the distractor that signalled a low-value reward: that is, a VMAC effect. As in previous experiments, this pattern of performance was counterproductive for participants’ earnings, since both speed and accuracy influenced the reward that was earned on each trial.

Interestingly, even under the simplified associative structure of Experiment 4 wherein mediating outcomes were removed, findings from the test phase appeared to mirror patterns of attention found in Experiment 3. The NoRev group experienced no change in the colour-fruit relationship between the training and test phases, and these participants continued to show slower RTs on high-value distractor trials than low-value distractor trials in the test phase, even though feedback on the magnitude of reward earned on each trial was now omitted. That is, the attentional bias toward the high-value distractor established in the training phase persisted through to the test phase, even under conditions of nominal extinction.

By contrast, participants in the Rev group experienced a switch in the colour-reward relationships; the distractor that signalled availability of a large number of points in the training phase thereafter signalled a small number of points in the test phase (and vice versa). Critically, the Rev group no longer exhibited slower RTs on high-value distractor trials than low-value distractor trials. Moreover, the difference in the VMAC effect between the NoRev and Rev groups approached significance ($p = .076$) with the NoRev group showing a numerically greater difference in RTs on high- versus low-value distractor trials than the Rev group. Together these findings suggest that with the instructed reversal of the reward relationships, the Rev group no longer showed a bias in attention towards the high-value distractor over the low-value distractor in the test phase.

Critically, however — and again consistent with Experiment 3 — the Rev group did not show a full reversal of their patterns of attention in the test phase in line with the new reward scheme. Participants in this group were aware of the new values of the fruits in the test phase, since they were explicitly told this information and required to pass check questions on new colour–point pairings following the instructed reversal. Moreover, all but one participant in the Rev group correctly reported which colour signalled bonus trials and which did not in the
final knowledge test (and exclusion of this erroneous participant did not change the pattern of findings). Yet, in the test phase, attentional prioritisation of the high-value distractor was not completely overcome to follow the new reward scheme. This finding suggests that the Rev group did not fully update attentional priority to reflect the new colour-point relationships, even under the simpler associative structure relative to Experiment 3.

Participants in the RevFB group also experienced an instructed switch in the colour-reward pairings following the training phase. In addition to receiving this information explicitly and passing check questions on the new colour–point pairings, participants in the RevFB group also received experience of feedback on the magnitude of reward (i.e., a small number of points versus a 10x bonus number of points) earned on each trial of the test phase. In general the findings from this RevFB group were in line with those of the corresponding group in Experiment 3, though somewhat less pronounced. Mean RT for low-distractor trials was slower (but not significantly so) for low-value than high-value distractor trials. Notably, accuracy was also lower for low- than high-value distractor trials, a reversal of the pattern observed during training – and this difference was significant. Overall, then, data from the RevFB group suggest poorer performance for low-value than high-value distractor trials during the test phase, providing some evidence that participants in this group had updated attentional priority of distractor colours in line with their updated status following revaluation. Taken together, these findings suggest that, even under simplified learning conditions, both knowledge and experience of stimulus–outcome pairings are needed to update attentional biases in line with new reward relationships.

Findings in the RevFB group are in line with a study by Albertella, Watson, et al., (2019) on the persistence of VMAC following a switch in the distractor-reward contingencies. This study also trained on the RT variant of the VMAC task, which was followed by an informed value-switch manipulation – participants were informed that the relationships between the distractors and bonus trials had reversed. Albertella et al. (2019) found that, when sensitivity to the value-switch manipulation was tested with trial-by-trial feedback (analogous to the RevFB group in the current experiment), participants showed a rapid change in their pattern of attentional prioritisation over a short (48 trial) test phase; the VMAC effect observed in training had disappeared, with a numerical trend toward slower RTs on trials featuring the low-value distractor trials (which now signalled a bonus trial) than high-value distractor trials. The findings of the current experiment extend these prior results
by demonstrating that in the Rev group, some sensitivity to a value-switch manipulation is observed even without the provision of additional trial-by-trial feedback.

2.6 Chapter Discussion

In the experiments presented here, we investigated whether conditioned attentional prioritisation of reward-related stimuli was sensitive to acute changes in the values of outcomes. In the training phase of all four experiments, participants were more often distracted by a stimulus that signalled a high-value outcome versus a low-value outcome. In Experiments 1 to 3, the high-value distractor stimulus became more likely to elicit a rapid oculomotor capture (than the low-value distractor stimulus) by virtue of its association with a high-value outcome. In Experiment 4, participants’ manual responses to the target were slower (and less accurate) when the display contained a distractor that signalled a high-reward relative to a low-reward. Notably, these patterns of results occurred even though looking at the high-value distractor was counterproductive to participants’ goals of maximising their payoff – in Experiments 1–3 it resulted in cancellation of a large reward, and in Experiment 4 it resulted in earning fewer points.

Interestingly, the degree to which this attentional prioritisation of high- versus low-value distractors persisted following a change in outcome values varied across experiments. In Experiment 1, instructed devaluation following training rendered both fruits of equally low value; in Experiment 2, instructed ‘super-valuation’ rendered both fruits of equally high value. Participants knew and understood the updated values of the fruits, and yet in both experiments attentional prioritisation of the high-value distractor persisted – with no evidence of a significant change in the magnitude of this effect relative to the training phase. This pattern of attentional prioritisation endured even when participants received feedback on the identity of the outcome earned on each trial of the test phase, thus providing further experience of distractor–outcome pairings under the updated value structure.

By contrast, in Experiment 3, the values of outcomes were switched following the training phase: the high-value fruit became low-value, and vice versa. This instructed value-switch influenced participants’ attentional prioritisation of distractors when tested in (nominal) extinction, such that the rapid bias towards the high-value distractor was weaker than in a control group of participants for whom outcome values remained unchanged from
training. However, knowledge of the instructed revaluation alone was insufficient for participants to completely reverse their pattern of attentional bias so that it reflected the updated outcome values. Instead, a ‘full’ reversal of attentional bias was observed only in a group of participants who received additional experience of receiving the outcome on each trial of the test phase in the form of feedback. This experience of explicit pairings of distractor colours and revalued outcomes in the test phase seemed to be sufficient for patterns of attentional prioritisation to reflect the changed values of the outcomes.

In Experiment 4, the fruit outcomes that mediated distractor colours and point values were omitted, simplifying the associative structure which participants (presumably) worked through to update patterns of attention following a change in outcome value. Crucially, during the subsequent test phase participants who experienced a switch in the colour–point relationships were no longer slower nor more error-prone to respond to the target in the presence of the high-value distractor (relative to the low-value distractor), suggesting that the bias toward the high-value distractor seen in training did not persist in the test phase. Moreover, for participants who were also provided with additional trial-by-trial feedback on the number of points earned on each trial following the value-switch, there was some evidence of a switch in the attentional bias towards the low-value distractor (which now signalled a high-reward), as indexed by a greater proportion of errors on these trials. The similarity of these findings to those of Experiment 3 suggest that the observed failures to demonstrate a ‘full’ reversal was not a consequence of the particular complexity of the associative structure (including mediating fruit outcomes) used in Experiment 3. Rather, the similarity of findings in these latter experiments strengthen the idea that what we are observing in the Rev conditions of these experiments is the operation of two competing influences on attention – an outcome-mediated component that attempts to update attentional priority, counteracted by an outcome-independent ‘habit-like’ component that has crystallised the value regimes learned in training.

Taken together, our findings provide important insights into the influence of reward on attentional capture. The results of Experiments 3 and 4 demonstrate that a change in outcome value can result in updating of attentional priority – both under conditions where the identity of the outcome paired with a stimulus changes (Experiment 4), and conditions where the stimulus–outcome (colour–fruit) relationships remain constant throughout, but the value of those outcomes changes (Experiment 3). These findings in turn suggest that the persistence
observed in Experiments 1 and 2 was a consequence of the particular revaluation procedure used there, in which both outcomes had the same value during the test phase. In this regard, our findings can be understood within the ‘adaptive view’ of attentional control proposed by Anderson (2021), wherein updating of attentional priority is based on cost-benefit accounting derived from reinforcement learning. In Experiments 1 and 2, neither outcome was worth more than the other during the test phase, and hence there may have been little incentive for participants to strive to earn a particular fruit. As such, there may have been little drive to exert (effortful) cognitive control in order to update attentional priority, with existing settings instead allowed to run on during the test phase. Existing work has shown that a difference in relative value—as opposed to absolute value—is important for the formation of value-modulated attentional biases (Kim & Beck, 2020); the current findings suggest that a difference in relative value may also be critical for the updating of existing biases. A practical implication of these findings is that revaluation procedures that result in similar outcome values (e.g., De Tommaso et al., 2017; De Tommaso & Turatto, 2021; Watson et al., 2022) may be suboptimal for detecting value-sensitivity of reward-modulated attention; instead procedures involving an ordinal change in outcome values may be better targeted in this regard.

Turning to Experiments 3 and 4, as noted above we saw evidence consistent with the idea that explicit knowledge of a reversal of outcome values can result in updating of attentional priority to reflect the new values given incentive to earn a relatively more valuable outcome over another. However the effect of this outcome-mediated updating will be opposed by an inflexible ‘habit-like’ process that continues to prioritise the (previously) high-value distractor in an outcome-independent manner. On this account, direct experience (i.e., provision of specific outcome feedback in the test phase) seems to play an important role in updating of this ‘habit-like’ component of conditioned attention, and hence full reversal of the pattern established in training. The implication is that reward-related attentional biases, once formed, can become (at least to some degree) divorced from explicit knowledge of the current value of the outcome, such that subsequent retraining is required for effective updating.
2.7 Chapter Conclusions

Overall, our findings from Experiments 1 to 4 revealed that value-modulated attentional prioritisation can show elements of both persistence and flexible updating following changes in outcome value. Clearly, the reward structure involved in the outcome revaluation plays a critical role in determining whether participants react in a ‘habit-like’ manner (allowing well-learned attentional patterns to persist in the face of devaluation or super-valuation) or whether cognitive resources will be recruited to update attentional settings (when there is a relative difference in values of outcomes). Even in the latter case, it appears that a ‘habit-like’ component will continue to exert control over attention, and that further retraining is required for effective updating in line with the new value regime.
Chapter 3

The Effect of Moderate Versus Extended Training on Attentional Prioritisation of Reward-Related Stimuli

3.1 Introduction

As discussed in previous chapters, the relative balance between goal-directed and habitual systems can influence reward-related behaviour (Dickinson, 1985; Wood & Rünger, 2016). Goal-directed responses are performed when the agent desires a valued outcome, and when the agent believes that the response will obtain the outcome (de Wit & Dickinson, 2009). Consider the example of perusing cafes after moving to a new city in order to find the best coffee; this behaviour is goal-directed by virtue of the desire to obtain delicious coffee, and the belief that going to cafes will obtain the coffee. In new learning scenarios such as this one, it is adaptive for behaviour to be under goal-directed control because this allows us to flexibly obtain valued outcomes (over less valued outcomes). Critically, after repeated and stable experience of a response producing the valued outcome, behaviour can be crystallised under habitual control to free up the agent’s cognitive resources (Wood & Rünger, 2016). After obtaining delicious tasting coffee from one particular cafe several times, it would make sense to settle on ordering our coffee from this cafe each day, as this would save us the cognitive energy of having to decide where to get our coffee each day. This example illustrates a central tenet of theorising on motivated behaviour: that a key factor in determining whether behaviour is under goal-directed or habitual control is the amount of experience of receiving of the outcome (i.e., the delicious coffee) as a result of the performing the response in the presence of particular stimuli (i.e., heading to the same cafe). That is, while behaviour may initially be goal-directed, it can transition to be under habitual control after prolonged and stable training.

This idea that instrumental behaviour can transition from goal-directed to habitual after extensive repetition has been demonstrated numerous times in animals (e.g., Adams, 1982; Dickinson et al., 1995; Thrailkill & Bouton, 2015). Adams (1982) was the first to demonstrate this effect using the outcome devaluation procedure. Rats were initially trained to press a lever for food pellets, before these pellets were devalued (through pairing with
illness induced by administration of lithium chloride). Following devaluation, a group of moderately trained rats still showed behaviour sensitive to changes in the values of outcomes, i.e., these rats reduced their lever-pressing as a consequence of outcome devaluation. By contrast, a group of extensively trained rats continued to frequently lever-press for a devalued outcome. This finding has been interpreted to suggest that following extended training, behaviour came to be under habitual control – persisting despite changes in rats’ desire for the now-devalued outcome (de Wit & Dickinson, 2009; Dickinson, 1985).

Following up on animal work showing that instrumental habits can be strengthened with repetition, translational research in recent years has shifted to investigate whether the same pattern can be observed with experimentally induced habits in human subjects. Anecdotally, we have all had the experience of forming habits through extensive repetition; for example, we might unthinkingly enter an old password even after we have changed it. Despite the importance of this issue, there are relatively few experimental demonstrations of this experience-dependent transition from goal-directed to habitual behaviour in humans in lab-based tasks. A study by Tricomi et al. (2009) measuring participants’ overt response selections found evidence consistent with greater insensitivity to devaluation (implying habit-like behaviour) following extensive training than following moderate training. However, subsequent attempts to replicate this finding have been unsuccessful. In five large-scale experiments with humans, de Wit et al. (2018) failed to replicate the transition from goal-directed to habitual behaviour, with participants continuing to make correct responses to earn valued outcomes over devalued outcomes. That is, participants in these studies continued to demonstrate goal-directed sensitivity to outcome devaluation even after extensive training over multiple days.

Taking a novel approach to this issue, Luque et al. (2020) posited that while overtraining in human laboratory studies may not be sufficient to produce errors in response selections towards devalued outcomes, it may lead to a detectable difference in the speed that correct responses are made. Hence, Luque et al. (2020) measured the strength of habit formation by the ability of previously learned instrumental responses to interfere with newly implemented, rapid goal-directed responding. In the training phase of their study – taking place over one day (moderate training) versus three days (extended training) – participants completed a task in which they could earn ‘diamonds’ of different value by carrying out one of two keypresses. In the subsequent test phase of the procedure, some diamonds were
devalued, such that alternative responses needed to be made to earn a different, more valuable diamond. The aim of this phase was to put the goal-directed system (favouring the now-optimal response) in conflict with the habit system (favouring the previously-optimal response), as this conflict was posited to slow down responding. Indeed, while participants made very few overt selection errors in the test phase (keypresses for devalued diamond outcomes), they were slower to respond on trials wherein the now-optimal response differed from the previously-optimal response than when the now-optimal response was the same as the previously-optimal response. Most importantly, the magnitude of this effect of prior training on response speed was larger in participants who had received three days of training than one day of training, consistent with stronger habit formation as a consequence of overtraining. The critical implications of Luque et al. (2020)’s findings are that while instrumental habits may not always manifest as overt behaviours, their presence may be felt through their interference with goal-directed overt responding, and this interference can indeed strengthen with overtraining.

Findings from Chapter 2 were consistent with the idea that habits may not always manifest as overt behaviours. When we examined habits beyond the traditional instrumental sense, we found that ‘habit-like’ processes may operate at the level of attention. To summarise, during the training phase of all experiments featured in Chapter 2, participants acquired attentional biases towards the distractor that signalled a high-value outcome, relative to the distractor signalling a low-value outcome (i.e., the VMAC effect; Le Pelley et al., 2015). Critically, once formed, the attentional bias towards the high-value distractor appeared to become insensitive to changes in outcome value under some conditions. Namely, despite the devaluation of a previously high-value outcome, or super-valuation of a previously low-value outcome, attentional prioritisation of the high-value distractor persisted (Experiments 1 and 2, respectively). By contrast, following the reversal of outcome values (wherein the high-value distractor now signalled a low-value outcome, and vice versa), the attentional bias toward the high-value distractor was weakened, but not completely overcome in line with the new value regime (Experiments 3 and 4). We interpreted these findings to suggest that the revaluation regime influenced the extent to which participants were motivated to flexibly change their attentional control settings, or instead let ‘habit-like’ attentional patterns learned in training play out. In other words, we take the findings of Chapter 2 to imply that the training procedure (and duration) used in those experiments resulted in a pattern of conditioned attentional bias that was influenced by an outcome-mediated component and an
outcome-independent component; that is, partly goal-directed and partly habit-like. In the devaluation and super-valuation procedures, participants have little motivation to update either aspect of attentional control, resulting in persistence of the trained pattern. By contrast, in the reversal procedure of Experiments 3 and 4, reversal instruction motivates updating of the goal-directed component (producing some change in the pattern of attentional bias), but updating of the habit-like component (to give a full reversal of attention) requires further training.

To reiterate, this interpretation of the findings of experiments in Chapter 2 assumes that the training procedure used there produced an attentional bias that had become (in sum) partly habit-like. As noted above, a foundational principle in the study of instrumental habits is that repetition is critical for habit formation: that habits are strengthened as a consequence of extended training. We therefore wished to investigate whether the habit-like attentional process identified in Chapter 2 also exhibited this cardinal feature of dependence on repetition: does the value-modulated attentional bias formed during training start out as being mediated by a representation of current value, and become increasingly outcome-independent as training proceeds?

### 3.2 Experiment 5

In Experiment 5, we investigated whether the habit-like properties of the attentional prioritisation response would vary as a function of the amount of initial training. We examined the effect of moderate versus extended training on performance on the attentional revaluation task featured in Chapter 2. Key methodological variations to the task are outlined below.

Due to restrictions to face-to-face testing resulting from the COVID-19 pandemic, we were unable to test participants face-to-face. Hence, we used a variant of the attentional revaluation task that could be completed by participants online. This procedure was based on a version of the attentional revaluation task in which the critical measure was response time (RT) to make a manual keypress based on the identity of the target. In previous research using this task (e.g., Le Pelley et al., 2015, 2022; Watson et al., 2019) and in Experiment 4, slower responses resulted in lower earnings. Despite this, participants have been found to be slower (but not more accurate) to respond to the target when the display contains a distractor
signalling availability of high-value reward than a distractor signalling low-value reward. This difference in RT is taken to suggest that the high-value distractor was more likely to capture participants’ attention.

Importantly, we re-introduced the fruit outcomes which mediated between distractor colour stimuli and point values, making the associative structure in this study identical to those featured in Experiments 1 to 3. Data from Experiments 3 and 4 suggest that similar findings obtain regardless of whether mediating outcomes are used (Experiment 3), or stimuli are directly paired with values in a simpler associative structure (Experiment 4). However, the use of mediating outcomes is preferable from a theoretical and design perspective because it allows us to keep the relationship between a stimulus and an outcome constant (e.g., a blue distractor always signalled that banana was available) while changing the value of that outcome by manipulating how much money each fruit was worth (e.g., by devaluing the banana from being worth 500 points to only 10 points). This chain of associations is closer to traditional procedures used to study instrumental habits in rodents, where the identity of the outcome remains constant throughout the task (e.g., lever pressing always produces food pellets) but the value of that outcome is changed (e.g., through feeding to satiety). This procedure allowed us a test of whether patterns of attention would be ‘habit-like’ – i.e., insensitive to these changes in outcome value – following moderate versus extensive training.

At the beginning of the experiment, participants were told that the aim of the task was to earn as many points as possible, as they could earn a cash bonus for above-average performance. They were told that they could earn points by winning fruits, with one type of fruit worth a large number of points (the high-value fruit), and the other fruit worth a small number of points (the low-value fruit). On each trial of the initial training phase, participants earned fruits for correct and rapid responses to the diamond-shaped target among circles. We motivated participants to respond quickly by using dynamic response time (RT) thresholds that adapted to each individual participant’s rate of responding: if participants responded below their RT threshold on a given trial, then they could earn a fruit, otherwise the fruit was cancelled. This procedure was intended to parallel the reward-omission contingency in our previous eye-tracking studies, providing top-down motivation for participants to try to ignore distractors – since if they attended to the distractor this would slow response to the target, increasing the likelihood of exceeding the RT threshold and hence missing out on reward.
The colour of a colour-singleton distractor circle on each trial signalled the type of fruit that could be won: one colour (the high-value distractor) signalled that the high-value fruit was available, and the other colour (low-value distractor) signalled the low-value fruit. Importantly, participants were assigned to one of two training-length conditions: they received either a moderate or extended amount of training. Notably, the amount of training received by the moderate condition was equal to the amount of training provided in Experiments 1 to 3. The key measure was response time (RT) to make a manual keypress based on the identity of the target; we examined RTs as a function of whether the trial featured a high- or low-value distractor.

Following the training phase, participants were assigned to one of three revaluation conditions – the reversal (Rev) group, devalued (Dev) group, and not-revalued (NoRev group) group. The manipulation for each condition was similar to that used in Chapter 2: participants in the Rev group were instructed that the values of the two outcomes were reversed following the training phase (the previously high-value fruit was thereafter worth a small number of points, and the previously low-value fruit was worth a large number of points); participants in the devalued (Dev) were instructed that the previously high-value fruit was thereafter worth a small number of points, such that it was now equal in value to the previously low-value fruit; and participants in the NoRev group were instructed that the values of the fruits remained unchanged during the test phase. All conditions completed the test phase of the attentional revaluation task under a nominal extinction procedure; they still earned fruit outcomes, but were not told about the identity of the fruit they earned on each trial. As discussed previously, nominal extinction is commonly used in tests of habits in humans, as it prevents new learning of relationships between stimuli and outcome values, while preserving participants’ motivation to perform (e.g., Luque et al., 2017, 2020; Tricomi et al., 2009). Omitting trial-by-trial feedback on the identity of the outcome earned allowed us to assess whether patterns of attention are mediated by explicit knowledge of outcome values, in the absence of further experience of stimulus-outcome (colour-fruit) pairings under the new value regime.

---

1 We did not include conditions in which outcome feedback was provided during the test phase in this study. Our focus in Experiment 5 was on the influence of amount of training on the degree to which the resulting pattern of attentional bias is sensitive to a subsequent change in explicit knowledge of outcome value. This question can be effectively assessed using the ‘instruction-only’ conditions, and the ‘with-feedback’ conditions are unnecessary for this purpose.
Notably, we selected these groups based on findings in Chapter 2. In the Rev condition, we previously found evidence consistent with partial sensitivity to a chance in outcome value, and partial persistence. Our hypothesis was that increasing the amount of training in this condition would increase the persistence of the conditioned attentional bias in the face of revaluation instructions. As in Experiment 3, the degree of persistence was assessed by comparing test-phase data with those of the NoRev condition. The Dev condition was included as a control. In Experiment 1, we previously found no evidence of a change in attentional bias following devaluation instructions, even on the basis of the moderate amount of training provided there. Consequently in Experiment 5 we expected to observe persistence during the test phase in the Dev condition regardless of the amount of initial training (again, assessed relative to the NoRev condition).

3.2 Method and Materials

3.2.1.1 Participants and Apparatus

A total of 193 UNSW Sydney students (123 females; age $M = 19.59$ years, $SEM = 0.24$ years) participated in the study for course credit. Participants who performed in the top 50% also earned a $10 AUD bonus (see Design and Procedure). Participants were allocated to one of three revaluation conditions (NoRev, Rev, and Dev), and for each of these revaluation conditions, participants received either a moderate length of training (NoRev $n = 35$; Rev $n = 35$; Dev $n = 34$) or extended training (NoRev $n = 30$; Rev $n = 33$; Dev $n = 26$). The task was programmed using jsPsych (de Leeuw, 2015), and participants completed the task on their own computer (it was not possible to complete the experiment on a mobile device). All research was approved by the UNSW Human Research Ethics Advisory Panel (Psychology).

3.2.1.2 Design and Procedure

Initial Value Instructions. At the start of the experiment, participants were told that their aim should be to earn as many points as possible, as the top 50% of participants who received the most points would be given a $10 AUD voucher. Participants were informed they could win points by earning fruits – lemons and bananas. Participants were randomly assigned to one of two fruit balance conditions – they were initially informed that lemons were worth 500 points and bananas were worth 10 points, or vice versa (Figure 3.1A). In
order to earn these fruits, participants were informed that they would have to respond correctly on trials of a visual search task.

Figure 3.1. Examples of stimulus presentation in Experiment 5. A) Participants were informed of the values of each fruit outcome at the start of the attentional revaluation task, and reminded of these values prior to each block of trials. B) An example trial. Each trial began with a central fixation cross. Following a blank interval, a search display appeared, containing a diamond. If participants responded correctly and rapidly to the orientation of the line segment inside the diamond, then they could earn a fruit on that trial. The colour of a colour-singleton distractor circle in the search display signalled whether a high- or low-value fruit could be earned on that trial: if the display contained a distractor rendered in the high-value colour, the participant could earn a high-value fruit worth 500 points; if the distractor was rendered in the low-value colour, the participant could earn a low-value fruit worth 10 points. The example here shows a high-value distractor trial, in which an orange distractor signals the availability of a 500-point fruit. Note: the colour-fruit and fruit-value contingencies were counterbalanced across participants.

All participants were required to answer check questions to ensure they had registered instructions regarding: the need to respond to the diamond target, which keys to press when the orientation of the line inside the diamond was horizontal or vertical, and the value of each of the fruit outcomes (10 points or 500 points). Participants needed to answer all check questions correctly before proceeding.

Training Phase. The training phase of this task was based on that used in Experiment 4 (see Figure 3.1B). Each trial began with a 400-ms central fixation cross, followed by a 150-ms blank screen. The search display then appeared: a diamond target, and 5 circles distributed evenly around the centre of the screen. The line inside the diamond was oriented either
horizontally or vertically (determined at random), while each circle contained a line tilted 45° randomly to the left or right. Importantly, there was an orange or blue distractor circle on each trial. Target and distractor locations were selected randomly on each trial.

Participants were informed that in order to earn fruits, they had to respond correctly to the orientation of the line inside the diamond target on each trial by pressing the ‘C’ key (if the line was horizontal) or the ‘M’ key (if the line was vertical). Not only were participants required to respond correctly to the orientation of the line inside the diamond target, they were also required to respond quickly, in order to earn a fruit. For the first 12 trials of the training phase, all participants had to make a correct response with an RT below 3000 ms to earn a fruit reward. Thereafter, a new RT threshold was set every 12 trials: the new threshold was calculated by adding 100 ms to the upper quartile of RTs for correct responses in the previous 12 trials (the threshold was updated only if at least six of the 12 previous trials had correct responses; if not then the previous threshold was carried over). Pilot testing revealed that this procedure for setting the RT threshold encouraged rapid responding (with slow responses being penalised) while ensuring that the task still felt ‘achivable’ such that penalties did not feel arbitrary. The use of a dynamic threshold tailored to each participant allowed for differences between participants in their baseline response rate: some participants naturally respond faster than others, and if we had applied a constant threshold for RTs across all participants (say 500 ms), this may have meant that some of the slower participants would have registered a majority of too-slow responses.

If participants responded correctly below the RT threshold, they could earn a fruit reward. Notably, the colour of the distractor on each trial indicated which type of fruit could be earned on the current trial. The relationship between colour and fruit identity was randomised for each participant: for some participants, a blue distractor signalled that a rapid response to the diamond target would earn a lemon, and an orange distractor signalled a banana; for other participants this assignment was reversed. Participants were not explicitly informed of these colour-outcome relationships. The distractor signalling the high-value fruit (worth 500 points) during training was termed the high-value distractor, and the distractor signalling the low-value fruit (worth 10 points) was the low-value distractor.

The feedback display replaced the search display only once a response was registered. If the response was made below the time threshold and correct, feedback showed “FRUIT WON!” and showed a picture of the appropriate fruit. If the response was made below the
time threshold but incorrect, feedback stated “ERROR: NO REWARD // You could have won.”, and showed the fruit overlaid with a large red ‘X’. If response time was greater than the time threshold, feedback stated “TOO SLOW: NO REWARD // You could have won.”, and presented the fruit overlaid with a red ‘X’. If the response time was registered as below 150 ms, indicating an anticipatory response, feedback stated: “Do not anticipate your responses: NO REWARD”. Feedback appeared for 1200 ms, except in the case the response was registered as an anticipatory response, wherein feedback appeared for 4000 ms (this was done to discourage anticipatory responding, by ensuring that the experiment would take much longer to complete if participants made anticipatory responses). The next trial began after a blank inter-trial interval of 700 ms.

In the training phase, participants in the extended training condition completed 30 blocks of 24 trials (720 trials in total), while participants in the moderate training condition completed 16 blocks of 24 trials (384 trials in total). The length of the training in the extended condition was determined by the maximum number of blocks that could be completed within a 1-hour testing session. The length of training in the moderate condition was chosen based on a pilot test; a group of participants received extended training (30 blocks of 24 trials) and we observed that it took approximately 16 blocks for the VMAC effect to clearly emerge. We took this as the point wherein participants had formed an attentional bias towards the high-value distractor over the low-value distractor, but could still be sensitive to changes in the values of outcomes. Incidentally, 16 blocks was also the same number of training trials provided to participants in Experiments 1 to 3 – though given the differences between studies (with Experiment 5 run online using an RT-based procedure, versus Experiments 1 to 3 which were lab-based and used eye tracking) there may not be a close functional equivalence between these training durations.

Half of the trials in each block featured a high-value distractor, and the other half featured a low-value distractor: trial order in each block was random. At the end of each block participants took a short break, during which they were informed of the total number of points they had earned so far. During each break, they were also reminded of the fruit values (presented with Figure 3.1A).

Revaluation Instructions. Immediately following the training phase, participants received further instructions depending on their group assignment, with different groups receiving different instructions about the values of the two fruits. As shown in Table 3.1,
participants in the NoRev group were simply reminded of the values of the fruits, which were unchanged from training (e.g., if lemons had previously been worth 500 points and bananas 10 points, participants in the NoRev group were reminded of these relationships). By contrast, participants in the Dev group were told that the values of the fruits had changed, such that the fruit previously worth 500 points in training was now worth 10 points, while the fruit previously worth 10 points was still worth 10 points. Participants in the Rev group were told that the values of both fruits had changed, such that the fruit previously worth 500 points in training was now worth 10 points, and the fruit previously worth 10 points was now worth 500 points.

In addition, all participants were instructed that they would no longer be told whether they had earned a lemon or a banana on each trial; “instead you will simply be told whether or not you won a fruit. You will still be earning these fruits, but we will keep track of your total. You should still try to earn as many points as possible.” After this instruction, all participants were required to answer check questions to ensure their knowledge of the current fruit values. Participants were asked to select how much a banana and lemon was then worth (500 points or 10 points), and asked whether or not they would be provided with feedback and points (they would not be provided with feedback but would still be earning points). All three responses were required to be correct before they could proceed.

Table 3.1. Design of each phase for one counterbalance condition in Experiment 5. Colours refer to colours of the distractor in the search display; fruits refer to outcomes that could be won for rapid responses. Note: fruit–value and colour–fruit contingencies were counterbalanced across participants.

<table>
<thead>
<tr>
<th>Initial value instructions</th>
<th>Training phase</th>
<th>Group</th>
<th>Revaluation Instructions</th>
<th>Test Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lemon = 500 pts&lt;br&gt;Banana = 10 pts</td>
<td>Blue → Lemon&lt;br&gt;Orange → Banana</td>
<td>NoRev</td>
<td>Lemon = 500 pts&lt;br&gt;Banana = 10 pts</td>
<td>Blue → ??&lt;br&gt;Orange → ??</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dev</td>
<td>Lemon = 10 pts&lt;br&gt;Banana = 10 pts</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev</td>
<td>Lemon = 10 pts&lt;br&gt;Banana = 500 pts</td>
<td></td>
</tr>
</tbody>
</table>

**Test Phase.** All participants then completed the test phase of the search task. RT thresholds continued to be updated every 12 trials as in the training phase. Participants
continued to earn fruits for responding correctly below the RT threshold, but did not receive feedback on the identity of the fruit earned on each trial: feedback stated “FRUIT WON!”, accompanied by “???” where previously a picture of the fruit had appeared during the training phase. If the response was registered below the RT threshold but was incorrect, feedback stated “ERROR: NO REWARD // You could have won a fruit”, and displayed “???” overlaid with a red ‘X’. If response time was above the RT threshold, feedback stated “TOO SLOW: NO REWARD // You could have won a fruit”, and displayed “???” overlaid with a red ‘X’.

Participants completed 4 blocks of trials in the test phase, with blocks structured as in the training phase (96 trials in total). The length of this test phase was shorter than the test phases featured in experiments in Chapter 2 (192 trials in total). We shortened the length here as we expected any sensitivity to the revaluation manipulation based on explicit knowledge to emerge immediately following the instructed revaluation and because this experiment did not feature any feedback conditions (compared to experiments featured in Chapter 2 wherein a longer test phase may have allowed updating of attention based on trial-by-trial outcome feedback). Participants in all groups were reminded of the current values of each fruit in the short break that followed each block.

Knowledge Checks. Following the test phase, we assessed participants’ knowledge of the colour–fruit contingencies (i.e., whether participants had learned that a particular distractor colour signalled the availability of a particular type of fruit), and the current value of each fruit. As noted in Chapter 2, the knowledge checks allowed us to verify that participants had explicit knowledge of the updated outcome values that persisted through the whole of the test phase. In order to motivate participants to try their hardest in these tests, they were informed that they could earn an additional 5000 points if they answered all questions correctly. First, to assess knowledge of the stimuli-outcome relationships, participants were told that the type of fruit that could be won on each trial depended on the colour of the coloured circle in the search display. They were then asked which fruit (banana or lemon) could be earned when an orange and blue coloured circle was in the search display. Moreover, while all participants were required to pass check questions on the current values of each fruit to progress to the test phase, we added a final knowledge check for the outcome-value contingency to further ensure that participants had encoded knowledge of the new fruit-values. In this latter knowledge check, participants were asked how many points (500 points or 10 points) the banana and lemon were currently worth (i.e., in the test phase).
3.2.1.3 Data Preparation

Data analysis followed our established protocols for analysing RTs from this version of the task (e.g., Le Pelley et al., 2022; Watson et al., 2019; and Experiment 4 of this thesis). We discarded data from the first two trials after each break, trials with responses that were faster than 150ms (indicating anticipatory responding), and trials with responses slower than 1500ms. Thirty-two participants who had more than 25% of trials excluded based on these criteria, or whose mean accuracy was low (<65% correct) were excluded from all analyses. Of the remaining participants, 94 were in the moderate condition (NoRev \(n = 32\); Rev \(n = 33\); Dev \(n = 29\)) and 67 were in the extended condition (NoRev \(n = 22\); Rev \(n = 24\); Dev \(n = 21\)). It is worth noting here that the exclusion rate was somewhat higher in the extended condition than in the moderate condition, and we return to this point in the Discussion.

Following previous work, the main dependent variable of interest was participants’ RT to the target (for correct responses only), on the grounds that attentional capture by distractors would lead to slower responses to the target. To examine potential trade-offs between speed and accuracy of responding, our secondary analyses focused on proportion of errors.

Notably, given the multiple factors involved in the following analyses (namely, phase, training length condition, revaluation condition, and distractor-type) we simplified our analyses by collapsing the distractor-type factor into a single dependent variable. Specifically, we calculated RT difference scores – the difference in mean RTs between high- and low-value distractor trials; note that (as in previous experiments) we label distractors according to the value they signalled during the training phase. Hence, RT difference scores above zero indicate a VMAC effect consistent with the reward regime in the training phase (i.e., slower responses on high-value than low-value distractor trials) with larger scores indicating a larger VMAC effect. While this approach has benefits for simplifying our analyses, for completeness we also include plots showing RTs separately for high- versus low-value distractor types in each phase, revaluation condition, and length condition. For our secondary analyses, we calculated error difference scores – the difference in the mean proportion of errors between high- and low-distractor trials.
3.2.2 Results

3.2.2.1 Training Phase

**RT Difference Scores.** We analysed RT difference scores during the training phase using a 2 (length condition: moderate vs. extended) × 3 (revaluation condition: NoRev, Dev, Rev) ANOVA (see Figure 3.2B). The main effect of length condition was not significant, $F(1,155) = 0.72$, $p = .40$, $\eta^2_p = .005$, nor was there a significant main effect of revaluation condition, $F(2,155) = 0.37$, $p = .70$, $\eta^2_p = .005$. There was a significant interaction between length and revaluation conditions, $F(2,155) = 3.89$, $p = .023$, $\eta^2_p = .048$, suggesting that the magnitude of the VMAC effect in each length condition differed across revaluation conditions. This finding was unexpected given that the training phase was completed before the revaluation manipulation, and hence all revaluation conditions had received equivalent treatment in this phase.

Setting aside for a moment the unexpected interaction between revaluation and length conditions, we examined whether participants in each length condition (collapsed across the three revaluation conditions) showed the VMAC effect. For each length condition, we conducted one-sample $t$-tests of RT difference scores against zero, where a score of zero would indicate no difference in RTs on high- versus low-value distractor trials – i.e., no significant VMAC effect. This analysis revealed a significant VMAC effect in both the moderate condition ($M = 10.05$ ms, $SEM = 2.15$ ms), $t(93) = 4.67$, $p < .001$, $d_z = .48$, and the extended condition ($M = 7.70$ ms, $SEM = 1.87$ ms), $t(66) = 4.11$, $p < .001$, $d_z = .50$. Unexpectedly, the mean VMAC effect in the extended condition was not greater than that in moderate condition – there was no significant difference in magnitude between the two conditions, $t(159) = 0.79$, $p = .43$, $d_z = .13$. 
Figure 3.2. A) Reaction time (RT) on high- and low-value distractor trials, B) RT difference scores (given by the difference in mean RTs on high- minus low-value distractor trials), and C) error difference scores (given by the difference in the proportion of errors on high- minus low-value distractor trials), for each length and revaluation condition in the training phase of Experiment 5. Error bars show within-subjects SEMs (Morey, 2008) in Panel A, and between-subjects SEM in Panels B and C.
As shown in Table 3.2, we repeated the above analyses of one-sample *t*-tests of RT difference scores against zero, this time separating the length conditions into revaluation conditions (given the revaluation × length interaction that was observed). For participants in the moderate training condition, all revaluation conditions (NoRev, Rev, Dev) showed significant VMAC effects, i.e., slower responses on high-value distractor trials than low-value distractor trials, *p* < .047. For participants in the extended training condition, the Rev and Dev condition showed significant VMAC effects, *p* < .030, but the NoRev condition did not, *p* = .674.

Table 3.2. Results of one-sample *t*-tests of RT difference scores and error difference scores against zero (which would indicate no significant VMAC effect) in each condition during the training phase of Experiment 5.

<table>
<thead>
<tr>
<th>Length Condition</th>
<th>Revaluation Condition</th>
<th>N</th>
<th>RT Difference</th>
<th>Error Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>t</em></td>
<td><em>p</em></td>
</tr>
<tr>
<td>Moderate</td>
<td>NoRev</td>
<td>32</td>
<td>3.85</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Dev</td>
<td>29</td>
<td>2.13</td>
<td>.042</td>
</tr>
<tr>
<td></td>
<td>Rev</td>
<td>33</td>
<td>2.08</td>
<td>.046</td>
</tr>
<tr>
<td>Extended</td>
<td>NoRev</td>
<td>22</td>
<td>0.43</td>
<td>.674</td>
</tr>
<tr>
<td></td>
<td>Dev</td>
<td>21</td>
<td>2.35</td>
<td>.029</td>
</tr>
<tr>
<td></td>
<td>Rev</td>
<td>24</td>
<td>4.15</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Error Difference Scores. For our secondary analyses, we analysed error difference scores during the training phase (shown in Figure 3.2C) using a 2 (length condition: moderate vs. extended) × 3 (revaluation condition: NoRev, Dev, Rev) ANOVA. The main effect of length condition was not significant *F*(1,155) = 0.22, *p* = .64, *η*² = .001, nor was the main effect of revaluation condition *F*(2,155) = 0.20, *p* = .82, *η*² = .003. The interaction between length and revaluation condition was not significant, *F*(2,155) = 1.16, *p* = .32, *η*² = .015.

As above for data on RT difference scores, we examined error difference scores for each of the length conditions (collapsed across the three revaluation conditions). For each length condition, we conducted one-sample *t*-tests of error difference scores against zero, which would indicate no difference in the proportion of errors on high- versus low-value distractor trials. Mirroring findings on RT difference scores, a significantly greater proportion of errors were made on high- versus low-value distractor trials in both the moderate condition
(\(M = 0.008, \text{SEM} = 0.004\)), \(t(93) = 2.15, p = .034, d_z = .22\), and the extended condition \((M = 0.010, \text{SEM} = 0.003), t(66) = 2.91, p = .005, d_z = .36\). The difference in the magnitude of error difference scores between these two length conditions was not significant, \(t(159) = 0.64, p = .48, d_z = .08\).

For completeness we also conducted one-sample \(t\)-tests of error difference scores against zero separating the length conditions into revaluation conditions. As shown in Table 3.2, for participants in the moderate training condition, only those in the NoRev condition showed a significantly greater proportion of errors on high- versus low-value distractor trials, \(p = .013\). For participants in the extended training condition, only those later assigned to the Rev condition showed a significantly greater proportion of errors on high- versus low-value distractor trials, \(p = .033\).

### 3.2.2.2 Test Phase

**RT Difference Scores.** Figure 3.3A shows the mean RT difference scores for each length condition and revaluation condition in the test phase. A 2 (length condition: moderate vs. extended) \(\times\) 3 (revaluation condition: NoRev, Dev, Rev) ANOVA revealed that the main effect of length condition was not significant, \(F(1,155) = 0.66, p = .42, \eta^2_{p} = .004\), nor was the main effect of revaluation condition, \(F(2,155) = 0.59, p = .56, \eta^2_{p} = .008\). Importantly, the interaction between distractor-type and revaluation-group was not significant, \(F(2,155) = 1.97, p = .14, \eta^2_{p} = .025\).
Figure 3.3. A) Reaction time (RT) on high- and low-value distractor trials, B) RT difference scores (given by the difference in mean RTs on high- minus low-value distractor trials), and C) error difference scores (given by the difference in the proportion of errors on high- minus low-value distractor trials), for each length and revaluation condition in the test phase of Experiment 5. Note that high- and low-value distractors are referred to by the value of the outcome they signalled during the training phase, with scores above zero indicating a VMAC effect consistent with the reward regime in the training phase (i.e., slower responses on high-value than low-value distractor trials). Error bars show within-subjects SEMs (Morey, 2008) in Panel A, and between-subjects SEM in Panels B and C.
Notwithstanding the non-significant findings of the omnibus ANOVA, we conducted planned analyses to investigate whether VMAC effects were present in each revaluation condition, and to investigate between-group contrasts that were of a priori interest. Table 3.3 shows the results of one-sample t-tests of RT difference scores against zero in each condition – i.e., tests of whether there was a significant VMAC effect in each condition. Following moderate training, the VMAC effect was not significant in any of the revaluation conditions (NoRev, Dev, Rev), \( ps > .08 \). Following extended training, the VMAC effect was non-significant in the NoRev and Rev conditions, \( ps > .36 \), and marginally significant in the Dev condition, \( p = .057 \). A notable aspect here was that there was no appreciable VMAC effect in the NoRev condition following extended training (i.e., mean RT difference score was very close to zero); this finding was unexpected as this condition was intended to be a ‘baseline’ in which the VMAC effect was strongest, which would have allowed us to observe changes in this effect under other revaluation conditions.

Table 3.3: Results of one-sample t-tests of RT difference scores and error difference scores against zero (which would indicate no significant VMAC effect) in each condition during the test phase of Experiment 5.

<table>
<thead>
<tr>
<th>Length Condition</th>
<th>Revaluation Condition</th>
<th>N</th>
<th>RT Difference</th>
<th></th>
<th>Error Difference</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>t</td>
<td>p</td>
<td>dz</td>
</tr>
<tr>
<td>Moderate</td>
<td>NoRev</td>
<td>32</td>
<td>1.78</td>
<td>.085</td>
<td>.31</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td>Dev</td>
<td>29</td>
<td>0.39</td>
<td>.697</td>
<td>.07</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>Rev</td>
<td>33</td>
<td>0.30</td>
<td>.766</td>
<td>.05</td>
<td>0.20</td>
</tr>
<tr>
<td>Extended</td>
<td>NoRev</td>
<td>22</td>
<td>0.13</td>
<td>.896</td>
<td>.03</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>Dev</td>
<td>21</td>
<td>2.02</td>
<td>.057</td>
<td>.44</td>
<td>2.45</td>
</tr>
<tr>
<td></td>
<td>Rev</td>
<td>24</td>
<td>0.93</td>
<td>.361</td>
<td>.19</td>
<td>0.06</td>
</tr>
</tbody>
</table>

The lack of significant VMAC effects in the baseline (NoRev) conditions limits the utility of between-group contrasts with the revaluation (Dev, Rev) conditions. Nevertheless we report these contrasts for completeness. First, comparing the NoRev and Dev conditions provided an index of the effect of the devaluation manipulation on performance during the test phase, since the only difference in treatment between these conditions was whether participants were informed that the high-value fruit had been devalued. Independent \( t \)-tests of RT difference scores in the NoRev and Dev conditions were conducted for both moderately and extensively trained participants; this approach is equivalent to examining the revaluation
group × distractor-type interaction with 2 x 2 ANOVAs used in Chapter 2. These analyses revealed that RT difference scores did not significantly differ between NoRev and Dev conditions after moderate training, \( t(59) = 0.94, p = .35, d_z = 0.24 \), nor after extended training, \( t(41) = 1.68, p = .10, d_z = 0.51 \). Bayesian analyses (conducted using the default prior in JASP: JASP Team, 2020) yielded a Bayes factor of \( BF_{01} = 2.65 \) for moderately trained participants, again suggesting only anecdotal evidence in favour of the null hypothesis. For extensively trained participants, the Bayes factor was equivocal, \( BF_{01} = 1.10 \) (Lee and Wagenmakers, 2014).

Next, comparing the NoRev and Rev conditions was intended to provide an index of the effect of the value-reversal manipulation on performance during the test phase. Independent \( t \)-tests revealed that RT difference scores did not significantly differ between NoRev and Rev conditions after moderate training, \( t(63) = 1.08, p = .28, d_z = 0.27 \), or after extended training, \( t(44) = 0.63, p = .53, d_z = .19 \). As above, we compared the RT difference scores between NoRev and Dev conditions using Bayesian independent \( t \)-tests. These analyses indicated only anecdotal evidence in favour of the null hypothesis (no difference in performance between the two groups; Lee & Wagenmakers, 2014), with a Bayes factor of \( BF_{01} = 2.40 \) for moderately trained participants, and \( BF_{01} = 2.91 \) for extensively trained participants.

**Error Difference Scores.** As shown in Figure 3.3B, a 2 (length condition: moderate vs. extended) × 3 (revaluation condition: NoRev, Dev, Rev) ANOVA of error difference scores revealed that the main effect of length condition was not significant, \( F(1,155) = 0.66, p = .42, \eta^2_p = .004 \), nor was the main effect of revaluation condition, \( F(2,155) = 0.59, p = .56, \eta^2_p = .008 \). Importantly, the interaction between distractor-type and revaluation-group was not significant, \( F(2,155) = 1.97, p = .14, \eta^2_p = .025 \).

Notwithstanding the non-significant findings of the omnibus ANOVA, we again examined error difference scores for each of the revaluation conditions using one-sample \( t \)-tests of RT difference scores against zero. Of all six conditions, only those in the extended Dev condition showed a significantly greater proportion of errors on high- versus low-value distractor trials, \( p = .033 \).
3.2.2.3 Knowledge Checks

Two participants in the moderate training group and one participant in the extended training group reported incorrect fruit-point associations in the knowledge check at the end of the experiment. Four participants in the moderate training group and three participants in the extended group reported incorrect colour-fruit associations in the knowledge check at the end of the experiment. We repeated the previously described analyses while excluding these participants. These analyses did not affect the pattern of significant findings.

3.2.3 Discussion

A foundational principle in the study of instrumental habits is that repetition is critical for habit formation: that habits are strengthened as a consequence of extended training. That is, while behaviour may initially be goal-directed, prolonged and stable experience of receiving the outcome for making a particular response in the presence of particular stimuli can strengthen habits (de Wit & Dickinson, 2009). Experiment 5 used an online, RT-based version of the attentional revaluation task to investigate the effect of training length on sensitivity to outcome revaluation in the context of conditioned attentional prioritisation. Specifically, we were interested in whether the habit-like attentional process identified in Chapter 2 also exhibited this cardinal feature of dependence on repetition: that is, whether the attentional bias formed in training starts out as being mediated by a representation of current value, and becomes increasingly outcome-independent with extended training.

In the training phase, prior to the instructed revaluation, participants who had received moderate training were slower (and less accurate) to respond to the target when the display featured a high-value distractor relative to a low-value distractor – i.e., the VMAC effect. This finding is consistent and broadly similar in magnitude to the corresponding effect observed in Experiment 4, and also provided a conceptual replication of Experiments 1 to 3 of this thesis, suggesting that the distractor that signalled a high-value fruit outcome was more likely to capture attention than the distractor that signalled a low-value fruit outcome. As in previous experiments, this pattern of performance was counterproductive for participants’ earnings, in that slower responses were more likely to result in the cancellation of high-value fruit outcomes (here due to failure to meet the RT threshold).
On the other hand, the VMAC effect was more variable for participants who had received an extended amount of training. A priori, we expected that the VMAC effect would be stable or (if anything) increase with longer training (cf. Le Pelley et al., 2015, Experiment 2). While there was still a significant VMAC effect overall in the extended conditions (collapsed across revaluation conditions), the magnitude of the VMAC effect in the training phase varied between revaluation conditions. Only the Dev and Rev conditions demonstrated a significant VMAC effect during the training phase, indicated by slower responding on high-value trials than low-value trials, while there was no significant VMAC effect in the NoRev condition (that is, the group of participants who did not experience a change in outcome values following extensive training). This was an unanticipated finding, since all revaluation conditions had received equivalent treatment in this initial training phase. This unusual finding may reflect noise in the data, an idea which I will return to shortly.

Patterns of attention in the test phase, following outcome revaluation, also did not support a clear story. For participants who received moderate training, the pattern of RT difference scores was broadly in line with evidence of sensitivity to changes in outcome value: relative to the NoRev condition, participants who experienced devaluation and value-reversal of outcomes showed numerically smaller VMAC effects. Notably however, none of these group differences reached significance, and in fact the NoRev condition no longer exhibited a significant VMAC effect during the test phase ($p = .08$), which may once again be a product of noisy data (perhaps exacerbated by the relatively limited number of trials involved in the test phase). For participants who received extended training, there was also no significant VMAC effect in the NoRev condition during the test phase – and here the effect did not even approach significance ($p = .896$). The lack of test-phase VMAC effects in both the moderate and extended training NoRev conditions was problematic for our intended analyses, because (as noted earlier) these NoRev conditions were intended as the primary baseline conditions against which we would compare changes in attention as a consequence of revaluation following moderate or extended training. Hence, we are unable to draw meaningful conclusions about sensitivity of VMAC to outcome devaluation or value-reversal, since we did not observe a VMAC in the control groups that did not experience an outcome revaluation manipulation. In fact, the mean VMAC effects for the extended Dev and Rev groups were actually numerically larger than for the corresponding NoRev group, although again these differences were not significant, and in the absence of further evidence it seems most likely that these patterns are another reflection of noisy data. Bayesian analyses of the
difference in performance between the NoRev and Rev conditions, and between the NoRev and Dev conditions, yielded inconclusive Bayes factors, again consistent with the idea that the amount of noise in the data was too great for the relatively subtle revaluation manipulations we were investigating.

In light of the equivocal findings, the critical question for understanding findings from Experiment 5 becomes why there seems to have been such noise in the data. It seems reasonable to conclude that the amount of noise largely reflects poor participant engagement with the task. As mentioned early on in this chapter, due to restrictions on face-to-face testing during the COVID-19 pandemic, participants completed the study online on their own devices, without experimenter supervision. For the extended training group in particular, it seems plausible that these participants responded more inaccurately because they became fatigued or bored from the sheer number of training trials they had to complete. While we aimed to fit in as many training trials in the extended condition timeslot as possible, this may have had the unintended (but understandable) effect of fatiguing this group of participants. After completing the task, participants had the option of providing written feedback on their experience of the experiment, and we found that many in the extended condition remarked that they found the study to be too long and tiresome. In line with this idea, random allocation to moderate versus extended training meant that initially, there were relatively similar numbers of participants in the moderate and extended training conditions. Notably, however, after restrictions were applied during Data Preparation, we saw a greater rate of exclusion in the extended training condition (22 participants excluded) than the moderate training condition (only 10 participants excluded). It seems plausible that participants in the extended training condition responded more inaccurately (and hence had to be excluded from analyses) because they became fatigued from the sheer number of training trials they had to complete. Hence, the revaluation instructions and test phase may have come at a time in the experiment where some participants were already disengaged from the task, and this would have meant that they were more likely to respond inaccurately during the test phase.

It is worth noting that we previously used an online, RT-based version of the attentional revaluation task in Experiment 4 where participants had also received a moderate amount of training: in that experiment we found clear evidence of a significant VMAC effect during the training phase, and found evidence of at least partial sensitivity to a reversal manipulation. The key methodological difference between Experiment 4 and the moderate
training condition of the current study was the re-introduction of the mediating fruit outcomes. As noted earlier, there was a benefit to re-introducing the mediating fruit outcomes, as this allowed us to have a more pure test of ‘habit-like’ persistence following changes in outcome value, analogous to what is often used in the instrumental habit literature. However, the potential trade-off of this procedure (particularly when used in an online self-completed version of the VMAC task) may have been that participants had low motivation to respond accurately as they were incentivised only by earning symbolic fruit outcomes with the possibility of a bonus for better-than-average performance. It is understandable that this reward structure would have been less motivating than the one used in Experiment 4, wherein participants directly earned a small or large number of points for each correct response. Of course Experiments 1–3 successfully used the more complex reward structure with mediating fruit outcomes, but these experiments were run in person, in the laboratory, featuring interactions with an experimenter and few external distractions—and this may have encouraged greater motivation and engagement at a level that was sufficient to produce satisfactory performance with the more complex task. In essence, our data may be taken to suggest that the “mediating outcomes” procedure is suitable for laboratory use, but it may be better to restrict online experiments to using simpler associative structures.

In consideration of these issues, a future study may be conducted for a more sound examination into the effect of training length on sensitivity to outcome revaluation. On the point above regarding the effect of reintroducing the mediating fruit outcomes into the RT version of the VMAC task, one option for a future study would be to remove these mediating fruit outcomes and instead use the simpler associative structure of Experiment 4—with the addition of manipulating the length of (moderate versus extended) training received with distractors that signal a 10x bonus trial versus a non-bonus trial. Of course, this approach has the downside of being a less pure test of whether attentional prioritisation is outcome-mediated or outcome-independent (i.e., ‘habit-like’) since the outcomes themselves are removed, and less directly maps onto the outcome revaluation procedure used in the study of instrumental habits.

As noted earlier, we did run a pilot version ($N = 55$) of the online task using the mediating-outcomes design prior to conducting Experiment 5, and the data from this pilot version were encouraging in the sense of showing the development of a VMAC effect over the course of training that stabilised as training continued. Unfortunately the data from the full sample of Experiment 5 appear to have been much more variable.
An alternative approach that I believe will produce the greatest improvements on the current study – by giving a strong and reliable VMAC effect in the NoRev condition – will come from conducting the experiment in-person (while retaining the mediating-outcome design). First, conducting the experiment in-person would give the experimenter control over potential distractions, and potentially participants would more consciously enter a “social contract” to complete the task to the best of their abilities. Second, participants could be incentivised for good performance more directly; much like Experiments 1 to 3, participants could earn fruit outcomes worth points that directly correspond to the bonus amount earned (rather than the possibility of a bonus for above-average performance). Third, in-person testing would also allow for the use of eye-tracking as a more direct and pure measure of attention than the RT-based measure, since saccades to a given location are closely and necessarily preceded by shifts in attention to that location (Deubel & Schneider, 1996).

In addition to running the study in person, minimising fatigue experienced by the extended training condition appears to be a priority given participants’ qualitative reports and the higher rate of exclusions of participants in the extended training conditions of the current experiment. It may be more effective to split training into multiple testing sessions – an approach used in previous studies in humans (e.g., Luque et al., 2020). Experiment 2 in Le Pelley et al. (2015) had three sessions of training on consecutive days, and the VMAC effect remained reliable over the three sessions (if anything, the relative magnitude of the VMAC effect became more pronounced each session). Having said that, taken together these changes would have significant resource implications, given the experimenter time required for running a long, eye-tracking study over multiple testing sessions with the number of participants needed for six experimental conditions; such a study was outside the time and resource confines of my PhD. One approach that would (somewhat) alleviate this load would be to run initial training sessions online (using an RT-based version of the task as in Experiment 5) before bringing participants into the lab for test sessions using eye-tracking. But this approach raises its own questions about the extent to which critical patterns of attention would transfer between different versions of the task, which may create further issues. Regardless, while this study served as a first step into the question of whether ‘habit-like’ attention can strengthen with extended training, future research with more resources may be able to improve the rigor of this study.
3.3 Chapter Conclusions

A key principle in the study of instrumental habits is that the length of training influences the relative balance between goal-directed and habitual control: that repetition through extended training can strengthen habits (de Wit & Dickinson, 2009). Building on this foundation from traditional instrumental habits, Experiment 5 investigated whether the ‘habit-like’ properties of the attentional prioritisation response (which we found in Chapter 2) would vary as a function of the amount of training. We examined the effect of moderate versus extended training on patterns of attentional prioritisation following outcome revaluation (specifically, value-reversal and devaluation). Examining findings from the initial training phase, we replicate findings from Chapter 2 on the development of an attentional bias towards a distractor signalling a high-value outcome over a distractor signalling a low-value outcome. However, results from the test phase were unclear, marred by noise in the data that likely reflected participant fatigue and boredom from a relatively long, online, self-paced version of the task. Despite the somewhat disappointing results, the current study does provide directions for a future study on the important question of whether attentional prioritisation can transition to come under habit-like control as a consequence of extended training. At present, however, this issue remains an important area for future work.
Chapter 4

The Effect of Outcome Revaluation on Attentional Disengagement from Reward-Related Stimuli

4.1 Introduction

Previous chapters in this thesis have focused on the process of attentional capture – that is, the ability of stimuli to automatically ‘pull’ attention to themselves. Importantly, however, attention can be fractionated into multiple components, and attentional capture is only one facet of a whole set of cognitive processes involved in attention. For example, once attention is captured by a stimulus, it can be ‘held’ by the stimulus before it is able to shift elsewhere; this process of attention shifting away from a stimulus is known as attentional disengagement. Clearly, attentional disengagement is another component of attention that will influence performance in a visual search task with salient distractors: once attention has been captured by a distractor, it must disengage from the distractor in order to move to the target – so any factor that influences the ease of disengagement can potentially influence the efficiency of search (Born et al., 2011; Posner & Petersen, 1990; Wang et al., 2019).

Previous research reviewed in previous chapters, and experiments in this thesis, suggest that reward learning can modulate attentional capture, such that stimuli associated with high-value outcomes are more likely to capture our attention than stimuli associated with low-value outcomes (i.e., the VMAC effect; Le Pelley et al., 2015; Pearson et al., 2016). In parallel to research on attentional capture, recent work has examined whether reward can also modulate attentional disengagement. One approach to examine the effect of reward on attentional disengagement has been via measurements of response time in tasks based on a dot-probe procedure (Müller et al., 2016), but concerns have been raised over the ability of these procedures to disentangle influences of reward on attentional disengagement from influences on attentional capture (Watson et al., 2020). An alternative approach to investigate the effect of reward on attentional disengagement is to use eye-tracking: since eye-movements are accompanied by shifts of attention (Deubel & Schneider, 1996), gaze can be used as a proxy measure of attention. To the extent that eye-tracking can provide a continuous ‘online’ measure of the location of attention, previous work using this approach has used the duration for which gaze lingered on a distractor (dwell time) as an index of the
ease of disengagement from that distractor (Le Pelley et al., 2015; Theeuwes & Belopolsky, 2012), with the idea being that the harder it is to disengage gaze from a distractor, the longer gaze will tend to linger on that distractor. For example, Le Pelley et al. (2015) used this approach in a VMAC procedure that was similar to the task used in Experiments 1 to 3 of this thesis: participants in this task earned a reward for making a saccade to the target, in the presence of colour-singleton distractors that signalled whether a high- or low-reward was available on that trial. To investigate the effect of reward on attentional disengagement, Le Pelley et al. (2015) considered the subset of trials on which gaze was captured by the distractor, and across this subset of trials calculated the mean gaze dwell time on the distractor as a function of whether it signalled a high versus low reward. Le Pelley et al. (2015) found that the mean duration of dwell times on high- versus low-value distractors was not significantly different, which could be interpreted as evidence that reward did not have an effect on attentional disengagement (see also Theeuwes & Belopolsky, 2012, for similar findings).

However, there are two problems with using this approach to measure attentional disengagement from reward-related distractors. First, dwell duration on distractors could be analysed only on the subset of trials on which the distractor had first captured gaze: we cannot measure time to disengage from a stimulus if attention never falls on that stimulus in the first place. This is an issue because the typical finding in these sorts of VMAC tasks is that the distractor captures attention on only a minority of trials (for example, across all participants in the study by Le Pelley et al., 2015, the low-reward distractor captured attention on less than 10% of trials), limiting the amount of data available for analysis of dwell times. Hypothetically considering a case wherein a participant had their gaze captured only once by the high-reward distractor and once by the low-reward distractor over the course of the experiment, the effect of reward on disengagement for this participant would have been totally determined by the difference in mean dwell time on these two trials. What this (albeit extreme) example demonstrates is that, because attentional capture by distractors occurred somewhat infrequently in this procedure, examining dwell time on distractors was likely a noisy measure that lacked power to detect an influence of reward on disengagement. Second, saccades to and from distractors are not necessarily made serially; it is possible, in principle, to program more than one saccade at a time (Born et al., 2011; Godijn & Theeuwes, 2002). The implication of this idea is that participants may have prepared corrective saccades to the target in parallel with their initial saccade to the distractor, and hence the measure of dwell
time on distractors in this procedure was not necessarily a pure measure of attentional disengagement from distractors, distinct from parallel programming of saccades towards the target. These ideas suggest that using the VMAC procedure as it stands—much like Experiments 1 to 5 in this thesis—may not provide a sensitive and pure test of the influence of reward on attentional disengagement from distractors. By contrast, measures of latency of first saccades in Experiments 1 to 3 featured a valid data point per trial, since this measure was not limited to trials on which participants had their gaze captured by the distractor.

To address these issues and more clearly examine the influence reward on attentional disengagement, Watson et al. (2020) incorporated a critical methodological change to the VMAC task. Namely, on some of the trials of this study, the coloured reward-signalling distractor circle could appear at the centre of the search display, at the location that participants were already fixating prior to the onset of the search array (i.e., on the fixation cross; Figure 4.1). Placing the distractors at the central location on these trials ensured that participants had their gaze fixed on the distractor at the outset, giving experimenters control over the number of trials that would ultimately be included in analyses of dwell time on distractors. As a result, the Watson et al. (2020) procedure had higher statistical power to detect an effect of reward on attentional disengagement, relative to the above approach that could analyse only the small subset of trials wherein distractors had captured attention (Le Pelley et al., 2015; Theeuwes & Belopolsky, 2012). Second, since overt attention was already at the location of the distractor on these central-distractor trials, participants on these trials needed to initiate only one saccade to the target (which was always presented in one of the outer locations) to earn a reward. Hence this procedure minimised the conflation of attention to the target with disengagement, instead providing a relatively pure measure of the ease with which attention could disengage from the distractor.
As in previous VMAC procedures, the colour of the distractor signalled whether a high-reward (500 points) or low-reward (10 points) was available on that trial for a rapid saccade to the target. The task was set up to motivate participants to disengage from central distractors as quickly as possible; they would earn points on each trial only if they made a saccade to the target below a time-limit, where this time-limit was adjusted dynamically and adaptively based on participants’ previous responses so that they remained under time pressure throughout (similar to the approach taken in the current Experiment 5). The primary dependent variable in this study was the latency for participants to begin moving their eyes away from the central distractor item to the target: to the extent that participants found it difficult to disengage attention from this central stimulus, they would take longer to move their eyes away from it. It was found that participants were significantly slower to move their eyes away from the high-reward distractor than the low-value distractor, even though this pattern was counterproductive since slower responses meant that participants were less likely to earn reward in this task. The implication is that participants found it harder to disengage attention from the high-reward stimulus than the low-reward stimulus, even though they were motivated to shift their attention as quickly as possible. Interestingly, this pattern of slower disengagement from the high-reward distractor endured even when rewards were removed in a subsequent unrewarded test phase. This persistence in the absence of further reward mirrors

---

*Figure 4.1. Illustration of the procedure used by Watson et al. (2020). Participants began each trial by gazing at a central fixation cross. The search display then appeared, containing a central coloured distractor at the location where participants had already been focusing their gaze. Rapid saccades away from the central distractor to the target were rewarded with points; the colour of the central distractor signalled the magnitude of reward (i.e., high- versus low-reward points) available for a rapid saccade to the target, which was always presented in an outer location.*
the persistence of the parallel process of value-modulated attentional capture in extinction (Watson, Pearson, Most, et al., 2019), as well as findings of the NoRev group during nominal extinction in Experiments 3 and 4 of this thesis. Together, these findings suggest that – in addition to its effect on attentional capture – reward learning can also modulate the ease of attentional disengagement. That is, signals of high-reward are not only more likely than signals of low-reward to ‘pull’ attention towards themselves, but they are also more likely to ‘hold’ attention at their location.

Clearly, our attentional systems prioritise stimuli associated with reward, and this prioritisation can be fractionated into multiple forms; reward-related stimuli can be imbued with the ability to capture attention (as in the VMAC effect), and also the ability to hold our attention to delay disengagement from them (Watson et al., 2020). Experiments in this thesis have gone further to demonstrate – most clearly in Experiments 1 to 4 – that reward-related stimuli can continue to capture attention even after outcomes have been revalued, suggesting that attentional prioritisation implicated by VMAC may (at least partially) ‘habit-like’ in its independence from the value of the outcomes involved.

Watson et al.’s (2020) findings suggest that – like attentional capture – delayed disengagement from signals of reward shows features in common with instrumental habits: the influence of reward on disengagement occurs even when it is counterproductive to participants’ goals, and once learned, will persist even when rewarding outcomes are removed. However, what remains unclear is whether delayed disengagement from reward-related distractors is mediated by retrieval of a representation of the outcome signalled by a stimulus, or whether delayed disengagement (like attentional capture) can also become ‘habit-like’ and divorced from the value of the events involved. Hence, the current study investigated whether delayed disengagement from reward-related stimuli would persist following outcome revaluation; that is, would this form of attentional prioritisation would be ‘habit-like’?

Rather than using the attentional revaluation procedure as it stands in previous experiments – which would ultimately be limited by the same issues as previous approaches to examining the effect of reward on attentional disengagement (Le Pelley et al., 2015; Theeuwes & Belopolsky, 2012) – the current experiment built upon the procedure used by Watson et al. (2020) which featured central reward-signalling distractors. Our key innovation was the introduction of fruit ‘outcome’ elements that mediated between the distractor stimuli
(colours) and rewards (high- vs. low-value points), as in previous experiments (1, 2, 3, and 5) of this thesis. By introducing the mediating fruit outcomes, we were able to keep the relationship between a stimulus and an outcome constant (e.g., a blue distractor always signalled that banana was available) while changing the value of that outcome by manipulating how much money each fruit was worth. This procedure allowed us to examine the effect of these changes in outcome value on patterns of attentional disengagement from the reward-related distractors. Investigating whether attentional disengagement is mediated by retrieval of the current value of the outcome signalled by a stimulus, or whether instead prioritisation becomes divorced from the value of the events involved, is crucial if we are to gain a richer understanding of the different components of attention on which ‘habit-like’ processes may exert an influence.

4.2 Experiment 6

In Experiment 6, we investigated whether delayed disengagement would be ‘habit-like’ in its persistence following outcome revaluation. As in previous experiments in this thesis, we used the attentional revaluation task with key methodological variations to the task outlined below.

During this time of my candidature, COVID-19 restrictions had come to a gradual end, meaning that once again we were able to test participants face-to-face. Hence (and as in Experiments 1 to 3) Experiment 6 was run as a lab-based study using eye-tracking as both the means by which participants made their responses during the task, and as the measure of attention. Importantly, unlike previous experiments in this thesis, the coloured distractor on each trial was positioned at the centre of the search display, where participants were already fixating prior to the appearance of the search display (i.e., at the location of the fixation cross). The task for participants was to make a saccade to the target, which required them to first disengage their attention from the central distractor, as the target was always positioned in an outer location; hence, the critical measure in this task was the latency of saccades made to the target.

As in Experiments 1 to 3, participants were told at the start of the experiment that the aim of the task was to earn as many points as possible, as they could earn a cash bonus based on their performance. They were told that they could earn points by winning fruits, with one
type of fruit worth a large number of points (500 points; the high-value fruit), and the other fruit worth a small number of points (10 points; the low-value fruit). On each trial of the initial training phase, participants earned fruits for rapid saccades to the diamond target. To motivate participants to quickly disengage from the central stimulus, participants were rewarded with fruit outcomes only if they made a saccade to the target that was faster than a dynamic response time (RT) threshold, much like the RT threshold used in Experiment 5. This RT threshold updated based on the individual participant’s rate of responding, and was used to parallel with the reward-omission contingency in previous eye-tracking tasks – the idea being that if participants disengaged slowly from the distractor, this would slow response to the target, increasing the likelihood of exceeding the RT threshold and hence missing out on the fruit outcome. This reward structure was used as an alternative to the omission schedule featured in the previous eye-tracking experiments in this thesis; it did not make sense to cancel the fruit reward each time participants looked at the distractors, since participants necessarily looked at central distractors on every trial.

As in previous experiments, the colour-singleton distractor circle on each trial signalled the type of fruit outcome that could be won: during the initial training phase, one colour (the high-value distractor) signalled that the high-value fruit was available, and the other colour (low-value distractor) signalled the low-value fruit. We examined latency of saccades to the target as a function of distractor type. In line with previous findings on the effect of reward on attentional disengagement (Watson et al. 2020), we expected that it would take participants longer to disengage from the high-value distractor than the low-value distractor in this initial training phase. This finding would suggest that the high-value distractor was more likely to elicit attentional prioritisation, and hence slow down disengagement, than the low-value distractor.

Following the training phase, participants were assigned to one of three revaluation groups – the not-revalued (NoRev group) group, devalued (Dev) group, and value-reversal (Rev) group. The manipulation for each group was the same as in Experiment 5: participants in the NoRev group were instructed that the values of the fruits remained unchanged during the test phase; participants in the Dev were instructed that the previously high-value fruit was thereafter worth a small number of points, such that it was now equal in value to the previously low-value fruit; and participants in the Rev group were instructed that the values of the two outcomes were reversed following the training phase (the previously high-value
fruit was thereafter worth a small number of points, and the previously low-value fruit was worth a large number of points). As in Experiments 3 to 5, the degree of persistence was assessed by comparing test-phase data in the NoRev group to the Dev and Rev groups. As discussed in Experiment 5, we selected the Dev and Rev groups because they were cases in the ‘capture’ version of the task where we saw evidence for persistence (Dev; see Experiment 1) and – to some extent at least – updating (Rev; see Experiment 3). And further, as in Experiment 5, we did not include groups in which outcome feedback was provided during the test phase in this study, as our focus in this study was on whether attentional disengagement is insensitive to a subsequent change in explicit knowledge of outcome value; this question can be effectively assessed using the ‘instruction-only’ groups, and the ‘with-feedback’ groups are unnecessary for this purpose.

Following these revaluation instructions, all groups then completed the test phase of the task under a nominal extinction procedure (similar to the approach used in Experiments 1, 2, 3, and 5); they still earned fruit outcomes, but were not told about the identity of the fruit earned on each trial. As discussed previously, omitting trial-by-trial feedback on the identity of the outcome allowed us to assess whether patterns of attention are mediated by explicit knowledge of outcome values, in the absence of further experience of stimulus-outcome (colour-fruit) pairings under the new value regime.

If the conditioned influence of reward on attentional disengagement is insensitive to explicit knowledge of the current value of the associated outcome – i.e., if it reflects the operation of a habit-like process – then we would expect the pattern of slower disengagement from the high-value distractor than the low-value distractor to persist during the test phase even following the change in outcome values resulting from the value-reversal and devaluation manipulations. By contrast, if the conditioned influence of reward on disengagement is mediated by a representation of the associated outcome – and hence dependent on the current value of that outcome – we would expect the speed of disengagement to be in line with the current value regime. For the Dev group, sensitivity would be indicated by no significant differences in saccade latencies to the target on high-value and low-value distractor trials (in line with the equally low values of the outcomes); for the Rev group, this sensitivity would be indicated by slower saccade latencies to the target on low-value trials than high-value trials during the test phase (in line with the reversed values of outcomes). In each case we assessed whether devaluation/value-switch instructions led to a
change in the pattern of reward-related attentional disengagement by contrasting test-phase performance of participants in the Dev/Rev groups with those in the NoRev group, which provided our baseline.

4.2.1 Method and Materials

4.2.1.1 Participants

Previous studies have found medium to very large effects ($d_z = 0.54–2.20$) for the influence of reward on attentional capture (e.g., Le Pelley et al., 2015; Pearson et al., 2015; Watson, Pearson, Chow, et al., 2019). Hence we aimed to recruit at least 24 participants per group; G*Power (with default settings) revealed that this would give power of .80 to detect a medium-sized effect ($d_z = 0.6$) of reward on attention in each group, and power >.90 to detect a medium-sized effect ($\eta^2_p = .06$) for the interaction relating to differences in the reward-related attentional bias across groups. In total, 96 UNSW Sydney students participated (59 females; age $M = 20.08$, $SEM = 0.46$ years). Group assignment alternated based on order of arrival; NoRev $n = 33$, Dev $n = 31$, Rev $n = 32$. Participants earned course credit, and received a monetary bonus depending on the number of points earned in the attentional revaluation task ($M = A$8.67, $SEM = A$0.19). Research reported here was approved by the UNSW Human Research Ethics Advisory Panel (Psychology).

4.2.1.2 Apparatus

Stimuli were presented on a 23-inch monitor (60 Hz refresh rate, 1920×1080 resolution); stimulus presentation was controlled by MATLAB with Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner, Brainard & Pelli, 2007). Participants were tested using Tobii Pro Spectrum eye-trackers (sampling rate 600 Hz). Gaze data were downsampled to 100 Hz for gaze-contingent calculations during stimulus presentation. Head position was stabilised using a chin-rest 60 cm from the monitor. The eye-tracker was calibrated at the start of the experiment using a 5-point procedure.

4.2.1.3 Design and Procedure

The attentional revaluation task was broadly based on the eye-tracking tasks featured in Experiments 1 to 3. It consisted of four components: the initial value instructions, training phase, revaluation instructions, and test phase.
**Initial Value Instructions.** At the start of the experiment, participants were told that their aim should be to earn points, since these points would be converted into cash at the end of the session, and that they could win points by earning fruits – lemons and bananas. Half of the participants were initially informed that lemons were worth 500 points and bananas were worth 10 points (Figure 4.2A); for the other half of participants, this was reversed. Participants were told that they could earn these fruits by moving their eyes to the diamond target on each trial “as quickly and as directly as possible”, but that if they were too slow to look at the diamond, the fruit they could have earned would be cancelled. However, participants were not informed of the specific colour–fruit contingencies (e.g., that a blue circle signalled availability of lemons and an orange circle signalled bananas).

![Figure 4.2](image.png)

*Figure 4.2. Examples of stimulus presentation in Experiment 6. A) Participants were informed of the values of each fruit outcome at the start of the attentional revaluation task, and reminded of these values prior to each block of trials. B) An example trial. On each trial, participants first focused on a central fixation cross. A search display then appeared, containing a diamond-shaped target among circles. The central circle was a colour-singleton distractor, whose colour signalled the type of fruit available for making a rapid saccade to the target. The example here shows a high-value distractor trial, in which a blue distractor signals availability of a fruit (lemon in this case) that is worth 500 points; note that fruit-value and colour-fruit contingencies were counterbalanced across participants. If participants did not make a saccade to the target faster than the RT threshold, the fruit reward would not be delivered on that trial.*

**Training Phase.** Each trial of the search task consisted of a fixation display, search display, and feedback display (Figure 4.2B). When participants had accumulated 700 ms of gaze time inside this fixation circle, or after 4000 ms, the cross and circle turned yellow to indicate that the search display was imminent. The search display then immediately appeared: it consisted of a diamond and 5 circles distributed evenly around a central circle. Critically, the central circle was a coloured distractor rendered in either orange or blue – all other shapes
were grey. In previous eye-tracking studies reported in this thesis, there was a 150-ms blank screen that intervened between the offset of the fixation cross and onset of the search display. However, in the current study, in line with previous work in our lab on attention disengagement (Watson et al., 2020), the blank screen was removed in this study, and instead the search display appeared immediately following the fixation display. Through this small change to the stimulus presentation, we aimed to maximise the likelihood that participants’ gaze was still at the location of the central distractor when the search display appeared (since there was no intervening period during which they could make an anticipatory saccade away from the central region).

Participants’ task was to move their eyes to the diamond target as quickly as possible; a response was registered when they had accumulated 100 ms of gaze dwell time within a region of diameter 3.5° centred on this target. The colour of the central distractor signalled the type of fruit that was available for a rapid response: for half of the participants, a blue distractor signalled that a rapid saccade to the diamond target would earn a lemon, and an orange distractor signalled a banana; for remaining participants this assignment was reversed. The distractor signalling the high-value fruit (worth 500 points) was the high-value distractor, and the distractor signalling the low-value fruit (worth 10 points) was the low-value distractor. The trial ended immediately after a response was registered, or after a timeout of 2000 ms.

Critically, participants earned the available reward only if their response time (RT) to register a response (i.e., to accumulated 100 ms dwell time on the target) was less than the RT threshold on that trial. In line with previously established protocols for calculating RT thresholds in eye-tracking studies (Watson et al., 2020), for the first block of training, the RT threshold was set at 1500 ms for all participants. From the second block onwards, the RT threshold was the upper quartile of recorded RTs from valid trials (i.e., trials which an eye-movement was registered to be faster than the timeout of 2000 ms) in the previous block. If the participant’s RT on a given trial was less than the RT threshold, then the participant earned a fruit reward: the feedback display stated “Fruit won!”, and showed a picture of the appropriate fruit. If RT was greater than the RT threshold, or if no response was registered before the timeout, feedback stated “Too slow: No Reward. You could have won.”, and presented the fruit overlaid with a red ‘X’. Feedback appeared for 1400 ms; the next trial then began after a blank inter-trial interval of 1400 ms.
Participants completed 16 blocks of trials in the training phase, with each block containing 24 trials (384 trials in total). Half of the trials in each block featured a high-value distractor, and the other half featured a low-value distractor: trial order in each block was random. The location of the target was randomly determined on each trial; it never appeared in the central location. At the end of each block participants took a short break, during which they were informed of the total number of points they had earned so far. During each break, participants also received a reminder of the fruit values (as in Figure 4.2A) which appeared on the screen for at least 10s; following this 10s period, participants opted when to continue with the task.

**Revaluation and Feedback Instructions.** Immediately following the training phase, participants received further instructions depending on their group assignment, with different groups receiving different instructions about the values of the two fruits. As shown in Table 4.1, participants in the NoRev group were simply reminded of the values of the fruits, which were unchanged from training (e.g., if lemons had previously been worth 500 points and bananas 10 points, participants in the NoRev group were reminded of these relationships). Participants in the Dev group were told that the values of the fruits had changed, such that the fruit previously worth 500 points in training was now worth 10 points, while the fruit previously worth 10 points was still worth 10 points. Participants in the Rev group were told that the values of both fruits had changed, such that the fruit previously worth 500 points in training was now worth 10 points, and the fruit previously worth 10 points was now worth 500 points.

In addition, all participants were instructed that they would no longer be told whether they had earned a lemon or a banana on each trial; “instead you will simply be told whether or not you won a fruit. You will still be earning these fruits, but we will keep track of your total. You should still try to earn as many points as possible.” After this instruction, all participants answered check questions to ensure their knowledge of the current fruit values: participants were shown a picture of a lemon and a banana and were asked to select the current value of each fruit. Both responses were required to be correct before they could proceed.
Table 4.1. Design of each phase for one counterbalance condition in Experiment 6. Colours refer to colours of the distractor in the search display; fruits refer to outcomes that could be won for rapid responses. Note: fruit–value and colour–fruit contingencies were counterbalanced across participants.

<table>
<thead>
<tr>
<th>Initial value instructions</th>
<th>Training phase</th>
<th>Group</th>
<th>Revaluation Instructions</th>
<th>Test Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lemon = 500 pts</td>
<td>Blue → Lemon</td>
<td>NoRev</td>
<td>Lemon = 500 pts</td>
<td>Blue → ??</td>
</tr>
<tr>
<td>Banana = 10 pts</td>
<td>Orange → Banana</td>
<td></td>
<td>Banana = 10 pts</td>
<td>Orange → ??</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dev</td>
<td>Lemon = 10 pts</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Banana = 10 pts</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rev</td>
<td>Lemon = 10 pts</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Banana = 500 pts</td>
<td></td>
</tr>
</tbody>
</table>

**Test Phase.** All participants then completed the test phase of the attentional revaluation task under nominal extinction. That is, they continued to earn fruits for responding correctly, but did not receive feedback on the identity of the fruit earned on each trial. If the saccade latency to the target was below the RT threshold on that trial, then feedback stated “Fruit won!”, accompanied by “??” where previously a picture of the fruit had appeared during the training phase. If latency to the target was greater than the RT threshold, or if no response was registered before the timeout, feedback stated “Too slow: No reward”, and displayed “??” overlaid with a red ‘X’. Participants completed 8 blocks of trials in the test phase, with blocks and RT thresholds structured as in the training phase. Participants in all groups were reminded of the current values of each fruit in the short break that followed each block.

**Knowledge Checks.** Following the test phase, participants’ knowledge of the colour–fruit contingencies was assessed. Participants were told that the type of fruit that could be won on each trial depended on the colour of the coloured circle in the search display. They were then presented with an orange and a blue circle, in random order, and for each they were asked to select which fruit (banana or lemon) they could win when that stimulus appeared in the search display. We also included a final knowledge check for the fruit-value contingencies to ensure that participants were aware of the new fruit values following revaluation: participants were presented with each fruit in random order and were asked to select whether it was currently worth 500 points or 10 points.
4.2.1.4 Data Preparation

Data analysis followed our established protocols for eye-tracking data (Experiments 1 to 3; Watson et al., 2020; Pearson et al., 2016). We discarded data from the first two trials after each break, trials that timed out with no response (0.92% of all trials), and trials with less than 25% valid gaze data (as a result of blinks, poor eye-tracking etc.: 0.03% of trials). For the remaining trials, valid gaze data were registered in over 99% of samples from the eye-tracker.

Following Watson et al. (2020), since all trials in this study had the distractor positioned in the centre of the display, our primary measure of interest was the latency of the first saccades that were directed towards the target, as a function of whether the trial featured a high- or low-value distractor; note that we label distractors according to the value they signalled during the training phase. As in saccade analyses featured in Experiments 1 to 3, a velocity-threshold identification algorithm (Salvucci & Goldberg, 2000) identified saccades using raw data from the eye-tracker, sampled at 600 Hz. Gaps in the raw gaze data shorter than 75 ms were first interpolated using linear interpolation. The gaze data were then smoothed using a five-point moving average filter. A saccade was then identified as an eye movement that remained above a velocity criterion of 40° visual angle per second for at least 10 ms. This saccade was classified as moving in the direction of the target if the saccade vector had an angular deviation less than 30° to the left or right of the centre of the target.

For these latency-based analyses, in addition to the exclusions described previously, trials were excluded if saccade latency was below 80 ms (indicating an anticipatory saccade), if the start point of the saccade was not within 100 pixels from screen centre (indicating that the participant’s gaze was outside the fixation point at the start of the trial), if there were gaps in the gaze data which were too large to be interpolated, or if there was insufficient gaze data to identify a saccade. There were 22 participants who had more than 20% of total trials excluded and hence, as in previous work (Watson et al., 2020), were excluded from the saccade-latency analyses. Across the remaining 74 participants (23 in the NoRev group, 25 in the Dev group, 26 in the Rev group), these additional trial-exclusion criteria resulted in removal of a further 5.8% of all trials. Included trial data for each participant were grouped by the phase (training vs. test) and distractor-type (high- vs. low-value).
4.2.2 Results

We first examined the latency of saccades to the target across the training and test phases using a 2 (phase: training vs. test phase) × 2 (distractor-type: high- vs. low-value distractor) × 2 (revaluation group: NoRev, Dev, Rev) ANOVA. Notably, preliminary analysis revealed that the three-way interaction was not significant, $F(2,71) = 1.30, p = .28, \eta^2_p = .035$. Nevertheless, planned analyses focused on the latency of saccades to the target in the training and test phases separately.

4.2.2.1 Training Phase

We examined the latency of saccades that went towards the target during the training phase using a 2 (distractor-type: high- vs. low-value distractor) × 2 (revaluation group: NoRev, Dev, Rev) ANOVA; data are shown in Figure 4.3. This revealed a significant main effect of distractor-type, $F(1,71) = 7.013, p = .01, \eta^2_p = .09$; averaged across revaluation groups, there were slower saccades to the target when the display featured the high-value distractor than the low-value distractor. The main effect of revaluation group was not significant, $F(1,71) = 0.15, p = .86, \eta^2_p = .004$. Importantly, the interaction between these two factors was not significant, $F(2,71) = 0.55, p = .58, \eta^2_p = .015$, suggesting the difference in the speed of disengagement from high- and low-value central distractors did not differ significantly between revaluation groups.
Figure 4.3. Latency of saccades to the target for trials with high- and low-value central distractors A) for the training phase as a whole, and B) across epochs of the training phase, where each epoch contained four blocks, for the NoRev, Dev, and Rev groups. Each point in Panel B was calculated by the mean saccade latency over 4 blocks. All error bars show within-subjects SEMs (Morey, 2008).

Visual inspection of the data in Figure 4.3 suggested some variation in the magnitude of effect of distractor type across groups. Hence we conducted further exploratory analyses examining the simple effect of distractor type in each of the revaluation groups. Paired samples t-tests revealed that slower disengagement from the high-value distractor (versus the low-value distractor) was significant in the Rev group, $t(25) = 2.74$, $p = .011$, $d = .54$, approached significance in the Dev group, $t(24) = 1.80$, $p = .085$, $d = .36$, but was not significant in the NoRev group, $t(22) = 0.58$, $p = .57$, $d = .12$.

Notably, the analyses reported above examine performance in the training phase collapsed across all blocks of training, including the initial training blocks wherein differences in the proportion of distraction trials on high- and low-value distractor trials may have been small (since participants had had little experience of the distractor–reward
relationships during these early blocks). Hence in Figure 4.3, we show graphs of the performance in each of the revaluation groups over the course of 4 epochs of the training phase (each containing 4 blocks of 24 trials), to examine whether an effect of distractor type was more apparent in the latter part of the training phase. In the Dev group (in Panel B) an effect more clearly emerges during the second half of the training phase; a paired samples \( t \)-test restricted to the Dev group’s performance in the second half of the training phase confirmed that saccade latencies were significantly slower on high-value distractor trials than low-value distractor trials, \( t(24) = 2.48, p = .02, d_z = .50 \). On the other hand, the NoRev group did not appear to show an obvious bias in attention towards the high-value distractor across the training phase. Nonetheless it should be noted once again that there was no a priori reason to expect differences in performance between revaluation groups during the training phase – since all groups were treated identically until after training was complete – and collapsed across all participants the main effect of reward (slower disengagement from the high-value than low-value distractor) was significant.

4.2.2.2 Test Phase

Figure 4.4 shows the latency of saccades that went to the target on high- and low-value distractor trials during the test phase, following the revaluation manipulation. A distractor-type \( \times \) revaluation group ANOVA revealed a significant main effect of distractor type, \( F(1,71) = 24.02, p < .001, \eta_p^2 = .25 \), with slower saccades to the target on high-value distractor trials than low-value distractor trials. The main effect of revaluation condition was not significant, \( F(2,71) = 0.02, p = .98, \eta_p^2 < .001 \). Importantly, there was also no significant interaction between distractor-type and revaluation group, \( F(2,71) = 0.40, p = .67, \eta_p^2 = .01 \). That is, the difference in the speed of disengagement from high- and low-value distractors did not differ significantly between revaluation groups.
Figure 4.4. Latency of saccades to the target for trials with high- and low-value central distractors in the test phase. Note that distractors are defined by the value of the outcome they signalled during the training phase (e.g., a high-value distractor was the colour associated with a high-value fruit outcome during training). Error bars show within-subjects SEM (Morey, 2008).

Notwithstanding the non-significant interaction, we conducted planned analyses to examine the simple effect of distractor type in each revaluation group, and to investigate between-group contrasts that were of a priori interest. Paired samples $t$-tests revealed that all three groups demonstrated significantly slower saccade latencies to the target on high-value trials than low-value trials – NoRev, $t(22) = 3.51, p = .002, d_z = .73$; Dev, $t(24) = 2.82, p = .009, d_z = .56$; Rev, $t(25) = 2.27, p = .032, d_z = .45$. That is, all three groups were significantly slower to pull their attention away from the distractor that previously signalled a high-value fruit in the training phase relative to the distractor that previously signalled a low-value fruit.

Next we more closely examined the critical question of whether the pattern of attentional disengagement from high- versus low-value distractors would persist after changes in the values of outcomes. First, comparing the NoRev and Dev conditions provided an index of the effect of the devaluation manipulation on performance during the test phase, since the only difference in treatment between these conditions was whether participants were informed that the high-value fruit had been devalued. Hence, the distractor-type × group interaction during the test phase provided an index of the effect of outcome devaluation on the speed of disengagement. A $2 \times 2$ ANOVA revealed that the revaluation group × distractor-type was not significant, $F(1,46) = 0.72, p = .40, \eta_p^2 = .02$, suggesting that the
magnitude of the bias towards slower disengagement from high-value distractors (over low-value distractors) did not differ significantly between groups. To further analyse this non-significant difference between the NoRev and Dev groups, we calculated *distractor difference scores* for each participant by taking the difference in the latency of saccades to the target on high- and low-value distractor trials; comparing these difference scores between NoRev and Dev groups via a Bayesian independent samples *t*-test (using the default prior in JASP: JASP Team, 2020) yielded a Bayes factor of $BF_{01} = 2.60$ in favour of the null hypothesis (no difference in the speed of disengagement from high- versus low-value distractors between the two groups) over the alternative hypothesis.

Next, comparing the NoRev and Rev conditions provided an index of the effect of the value-reversal manipulation on performance during the test phase, since the only difference in treatment between these groups was whether participants were informed that the high-value fruit was now worth low-value, and the low-value fruit was now worth high-value. As above, a $2 \times 2$ ANOVA revealed that the revaluation group $\times$ distractor-type interaction was not significant, $F(1,47) = 0.53$, $p = .47$, $\eta_p^2 = .01$, suggesting that the magnitude of the bias towards slower disengagement from high-value distractors (over low-value distractors) did not differ significantly between groups. A Bayesian independent samples *t*-test of distractor difference scores between these two groups yielded a Bayes factor of $BF_{01} = 2.82$ in favour of the null hypothesis (no difference in the speed of disengagement from high- versus low-value distractors between the two groups) over the alternative hypothesis.

### 4.2.3 Knowledge Checks

Two participants in the Rev training group reported incorrect colour-fruit associations in the knowledge check at the end of the experiment. One participant in the Dev group and one participant in the Rev group reported incorrect fruit-point associations in the knowledge check at the end of the experiment. We repeated the previously described analyses while excluding these participants. These analyses did not affect the pattern of significant findings.

### 4.2.3 Discussion

Attention is a multifaceted system that can be fractionated into multiple components, including attentional capture and attentional disengagement. Previous research suggests that reward can exert an influence over both these components of attention: stimuli associated
with high reward are more likely to be prioritised by our attentional system than stimuli associated with low reward, imbuing them with a greater ability to elicit attentional capture (i.e., the VMAC effect; Le Pelley et al., 2015) and delay attentional disengagement (Watson et al., 2020). Previous experiments in this thesis have demonstrated that under some conditions, high-reward-related stimuli continue to be prioritised by attention (over low-reward-related stimuli) even after the outcomes signalled by the stimuli have been revalued, suggesting that attentional prioritisation implicated by VMAC may ‘habit-like’ in its independence from the value of (i.e., desire for) the outcomes involved.

Furthering this line of enquiry on the effects of reward on attention, Experiment 6 built on a variant of the VMAC task used by Watson et al. (2020) wherein the distractors were always located at the centre of the search display where participants had already been fixating, therefore requiring participants to disengage from the distractor to make a saccade towards the target in an outer location. This approach provided a statistically robust and pure measure of the ease of attentional disengagement, distinct from attentional capture (relative to previous approaches; Le Pelley et al., 2015; Theeuwes & Belopolsky, 2012). Importantly, as in previous chapters, we introduced outcome elements that mediated between stimuli (colours) and rewards (high- vs. low-value points), allowing us to keep the relationship between a stimulus and an outcome constant while changing the value of that outcome through outcome devaluation or reversal. This procedure allowed us to investigate whether delayed disengagement from reward-related distractors is mediated by retrieval of the outcome signalled by a stimulus, or whether delayed disengagement (like attentional capture) can also become ‘habit-like’ and persistent in the face of changes in outcome value.

During the training phase, prior to the revaluation manipulation, participants (averaged across revaluation groups) showed slower saccades to the target when the central distractor signalled a high-value outcome, relative to a low-value outcome. In line with previous findings (Watson et al., 2020), this suggests that reward can have an influence on attentional disengagement: distractors signalling high-value outcomes are more likely to be prioritised by our attentional systems and hence ‘hold’ our attention at the central location, relative to distractors signalling low-value outcomes. Critically, this pattern of performance was counterproductive for participants’ earnings, in that slower latencies to the target were more likely to result in the cancellation of high-value fruit outcomes (due to failure to meet the RT threshold). There was no interaction between distractor type and revaluation group.
during the training phase, indicating that the pattern of reward-related prioritisation did not differ significantly between groups – which is unsurprising, since all groups had experienced equivalent treatment at this point. It is notable, however, that focused analyses revealed that the simple effect of distractor type did not reach significance in of the revaluation groups during the training phase; in particular, we found that there was no significant effect for participants in the NoRev group. Moreover, this pattern was only marginally significant in the Dev group across the whole training phase; it did emerge when analyses were restricted to the second half of the training phase. These hints of differences in the strength of the training-phase reward effect between groups are presumably a consequence of noise in the data (as indicated by the data from the test-phase; see below).

Of greater interest was the effect of reward on attention disengagement in the test phase, following the revaluation manipulation. For participants in the NoRev group, there was no change in the value of fruit outcomes between the training and test phases: these participants were instructed that the fruit that had high value in the training phase continued to have high value in the test phase, and vice versa. These participants showed significantly slower disengagement from high-value distractors than low-value distractors during the test phase, even though no feedback was provided on the specific fruit earned on each trial. This finding replicated the pattern reported by Watson et al. (2020) wherein slower disengagement from the high-value distractor than the low-value distractor endured even when rewards were removed in a subsequent unrewarded test phase. It is perhaps surprising that a robust effect of reward on disengagement emerged in the test phase for the NoRev group, given that the corresponding effect was nonsignificant during the preceding training phase. The reason for this is unclear – since outcome feedback was not provided during the test phase, it seems unlikely that the effect had grown more pronounced in this phase due to further conditioning of attention. As noted above, one possibility is that noise in the training data had obscured an underlying effect during that phase.

In this experiment, the test-phase data from the NoRev group provided the primary baseline group against which we assessed whether there had been changes in the pattern of reward-modulated attentional disengagement as a consequence of outcome revaluation. First, for the Dev group, the fruit that was of high-value during the training phase was devalued, such that it was of equally low-value to the fruit that was previously of low-value. Despite this change in outcome value, there was no significant effect of outcome devaluation on
patterns of attentional disengagement in the test phase; the Dev group continued to show slower disengagement from the high-value distractor than the low-value distractor, and the magnitude of this difference in the speed of disengagement did not differ significantly from the NoRev group. This finding corresponds to the earlier findings in Experiment 1 of this thesis, wherein despite an outcome devaluation manipulation, participants in the Dev group continued to show greater attentional capture by the high-value distractor than the low-value distractor. Again, as in previous experiments of this thesis, the persistence here cannot be attributed to a failure to update knowledge on outcome values: participants were explicitly informed of the new values, required to pass check questions on the values of the fruits following instructed devaluation, reminded of these values repeatedly at the end of each block, and (all but one participant) were able to correctly report the values of the two fruits in a knowledge check at the end of the experiment. Relating back to earlier findings on the persistence of patterns of attentional capture in Experiments 1 to 4, a possible explanation for the observed persistence of patterns of attentional disengagement is that since there was no relative difference in outcome values, participants were not incentivised to engage in effortful updating of attentional control (updating which will be attempted only if there is incentive to earn a more-valuable outcome over a less-valuable outcome).

To examine this possibility of the importance of a relative difference in outcome value for both attentional capture and disengagement, we turn to the current Rev group: for this group, the fruit that was of high value during the training phase thereafter had low value, and the fruit that was of low value thereafter had high value. Notably, even though this value-reversal manipulation meant that there was still a relative difference in outcome values during the test phase, the Rev group in Experiment 6 showed the same persistence in patterns of attention as the Dev group – i.e., slower disengagement from the high-value distractor than the low-value distractor, with the magnitude of this difference in the speed of disengagement not differing significantly from the NoRev group. Again, this pattern of attentional disengagement in the Rev group occurred even though participants had explicitly encoded the information about the updated values of the outcomes, and (all but one participant) retained that knowledge throughout the test phase. Hence, the insensitivity of attention disengagement to the value reversal manipulation cannot simply be ascribed to a failure to shift explicit knowledge regarding the values of the fruits. Importantly, the pattern of attention disengagement in the Rev group is interesting given findings of Experiment 3 and 4 on attentional capture. In these previous experiments, we found that a reversal of outcome values
resulted in partial updating of attentional priority, suggesting that – unlike the current experiment – the persistence of the bias toward the high-value distractor was influenced by an outcome-mediated component, but the effect of this outcome-mediated updating of attentional control was opposed by an inflexible ‘habit-like’ process. I delay further discussion of the effect of outcome revaluation on attentional disengagement versus attentional capture to the General Discussion.

It is important to note that some caution should be applied to conclusions drawn here. While our primary analyses revealed that the magnitude of the bias towards slower disengagement from high-value distractors (over low-value distractors) did not differ significantly between the NoRev and Dev/Rev groups, Bayesian analyses yielded only anecdotal evidence in favour of this null hypothesis. These Bayes factors suggest that the experiment might have been underpowered, and a future study would benefit from collecting data from a larger sample to improve the power of the study. The implication of these Bayesian analyses is that, at best, conclusions drawn in the current experiment are based upon accepting the null hypothesis. This is an important issue that I will return to in the General Discussion.

What remains to be seen for a future study is whether the pattern of delayed disengagement from the high-value distractor over the low-value distractor would be amenable to retraining through the provision of trial-by-trial feedback on the outcome earned on each trial. This is a particularly interesting question given the findings of Experiment 3 where we saw that outcome feedback was required for the outcome-independent habit-like component to be overcome, resulting in a ‘full’ reversal of attentional patterns following value-reversal. While the current study did not include feedback groups, this remains a possible avenue for future research.

4.3 Chapter Conclusions

Taken together, the findings of this experiment provide important insights into the influence of reward on attentional disengagement. In the training phase, the high-value distractor became imbued with attentional prioritisation, giving it the ability to counterproductively ‘hold’ attention for longer than the low-value distractor. This pattern of attention was learned based on participants’ knowledge of the relationships between the
distractor stimuli, fruit outcomes, and the values of fruits. Once formed, however, this pattern of attentional prioritisation became crystallised and inflexible; participants still took longer to disengage from the high-value distractor than the low-value distractor even after the outcome devaluation or reversal took place, and even when participants were able to report explicit knowledge of these changes in value. In sum, it appears that attentional disengagement can become divorced from a representation of the outcome (and changes in the value of the outcome), and hence it can be considered ‘habit-like’.
Chapter 5

General Discussion

Previous work has demonstrated that stimuli signalling high-value rewards are more likely to be prioritised by our attentional system than signals of low-value rewards, even when this is counterproductive to one’s goals. This modulating influence of reward learning has been termed value-modulated attentional capture (VMAC). Viewed another way, reward-related stimuli can be said to automatically elicit a conditioned attentional prioritisation response, raising the idea that prioritisation of reward-related stimuli may be ‘habit-like’. The VMAC effect shows some features in common with the instrumental habits that are typically studied in the behavioural neuroscience literature: it increases with extended training (Pearson et al., 2015), persists in extinction (Anderson et al., 2011a; Watson, Pearson, Most, et al., 2019), and can operate independently of (and indeed, contrary to) an individual’s goals (Le Pelley et al., 2015). However, at the outset of this thesis, it remained unclear whether attentional prioritisation of reward-related stimuli met the key feature of habits as defined by behavioural neuroscience – that attentional prioritisation becomes divorced from the value of the events involved – or alternatively whether it is mediated by retrieval of a representation of the outcome signalled by a stimulus. Drawing from the behavioural neuroscience literature, one of the critical features of a response that has become divorced from the value of the events involved (i.e., habitual) is that it persists even following changes in desire for the outcomes (de Wit & Dickinson, 2009).

Building on these ideas about instrumental habits and applying them to our understanding of attention, my thesis investigated the question of whether attention can be ‘habit-like’ in three parts: 1) Under what conditions is attentional prioritisation of reward-related stimuli sensitive to changes in the value of the outcome? 2) Does the amount of training have an influence on the extent to which attentional prioritisation will be ‘habit-like’ in its persistence? and 3) Does attentional prioritisation of reward-related stimuli result in ‘habit-like’ delayed disengagement? In a series of 6 experiments, we manipulated participants’ desire for outcomes by incorporating outcome revaluation manipulations into the VMAC procedure. We used both eye-tracking and response-time measures of attention to assess whether participants formed attentional biases towards stimuli associated with high-value outcomes over stimuli associated with low-value outcomes, and whether these biases
persisted following the revaluation manipulations. In this final chapter, I summarise and synthesise the key findings from these studies. I also discuss the theoretical and broader implications of my findings. Finally, I go on to discuss limitations of the thesis and potential directions for future research.

5.1 Summary of Findings

5.1.1 Under what conditions is attentional prioritisation of reward-related stimuli sensitive to changes in the value of the outcome?

Previous research (Pearson et al., 2015; Le Pelley et al., 2015) has demonstrated that in the single-phase VMAC procedure, participants are more likely to have their attention captured by a stimulus that signals a high-value reward (the high-value distractor) than a stimulus that signals a low-value reward (the low-value distractor), even when this is counterproductive to participants’ goals. What remained unclear was the nature of the process by which the reward-related distractors elicit these conditioned attentional responses. In Chapter 2, we investigated whether attentional prioritisation of the high-value distractor (relative to the low-value distractor) is sensitive to acute changes in the values of outcomes, and this allowed us to test whether the attentional prioritisation response is mediated by a representation of the outcome and its current value, or whether it can become divorced from the value of the events involved – that is, whether it can be ‘habit-like’.

In line with previous work on VMAC, data from the initial training phase of Experiments 1 to 4 revealed that the high-value distractor was more likely to elicit attentional prioritisation than the low-value distractor. In Experiments 1 to 3, the high-value distractor stimulus became more likely to elicit rapid oculomotor capture (than the low-value distractor stimulus), and in Experiment 4, participants’ manual responses to the target were slower (and less accurate) when the display contained a distractor that signalled a high reward relative to a low reward. These patterns of results occurred even though looking at the high-value distractor was counterproductive to participants’ goals of maximising their payoff because looking at the distractor led to cancellation of the reward (Experiments 1 to 3) or because response time and accuracy influenced reward magnitude (Experiment 4).

Regarding the critical question of whether the attentional bias towards the high-value distractor would be sensitive to acute changes in the values of outcomes, we found that the
degree of sensitivity to these changes varied depending on the specific revaluation regime. Under conditions of devaluation in Experiment 1 or ‘super-valuation’ in Experiment 2, attentional prioritisation of the high-value distractor persisted. Strikingly, this persistence was seen even in groups who received feedback on the identity of the outcome earned on each trial following revaluation – i.e., further experience of the distractor-outcome pairings. This failure to update patterns of attentional prioritisation cannot be attributed to a lack of explicit knowledge of the new outcome values; participants were explicitly informed of the new values, required to pass check questions on the values of the fruits following instructed revaluation, reminded of these values repeatedly at the end of each block, and able to correctly report the equal values of the outcomes in a knowledge check at the end of the experiment. Hence under these conditions, attentional prioritisation of the high-value distractor appears ‘habit-like’ in its persistence despite the changes in outcome value. Critically, these experiments shared a common feature; following revaluation, the outcomes became of equal value – equally low in the case of devaluation, and equally high in the case of super-valuation.

By contrast, in Experiment 3, the values of outcomes were reversed in the test phase such that the high-value fruit became low-value and vice versa, and similarly in Experiment 4, the colour-point values were reversed (with fruit outcomes being omitted, essentially simplifying the associative structure). Under these conditions, we found that the distractor that had signalled a high-value outcome during initial training was no longer more likely to elicit attentional prioritisation than the low-value distractor, and the magnitude of this high-vs-low difference was smaller than in a control group of participants for whom outcome values remained unchanged from training. Pulling together previous ideas on the importance differences in relative value in the formation of value-modulated attentional biases (Kim & Beck, 2020) and the ‘adaptive view’ of attentional control (Anderson, 2021), our findings in these latter experiments suggest that effortful updating of attentional control will be attempted when there is incentive to earn a more-valuable outcome over a less-valuable outcome. Importantly, however, across both Experiments 3 and 4, knowledge of the instructed revaluation alone was insufficient for participants to completely reverse their pattern of attentional bias so that it reflected the updated outcome values. Instead, there was evidence of the attentional bias being overcome only in the groups of participants who received additional experience of receiving the outcome on each trial of the test phase in the form of feedback. The implication is that reward-related attentional biases, once formed, can
become (at least in part) divorced from explicit knowledge of the current value of the outcome, such that subsequent retraining is required for effective updating – but this updating will be attempted only if there is a relative difference in the values of outcomes.

Overall, findings from Experiments 1 to 4 revealed that value-modulated attentional prioritisation can show elements of both persistence and flexible updating following changes in outcome value. Clearly, the reward structure following outcome devaluation plays a critical role in determining whether participants react in a ‘habit-like’ manner (allowing well-learned attentional patterns to persist in the face of devaluation or super-valuation) or whether more goal-directed control will be recruited to update attentional settings (when there is a relative difference in values). Even in the latter case, it appears that further retraining is required for effective updating, suggesting that attention is partially habit-like even under these reward structures.

5.1.2 Does the amount of training have an influence on the extent to which attentional prioritisation will be ‘habit-like’ in its persistence?

A foundational principle in the study of instrumental habits is that repetition is critical for habit formation: that habits are strengthened as a consequence of extended training. That is, while behaviour may initially be goal-directed, it can transition to be under habitual control after prolonged and stable experience of receiving the outcome for making a particular response in the presence of particular stimuli (de Wit & Dickinson, 2009). This idea has been demonstrated numerous times in animals, wherein extensively trained animals will perform a response for a now-devalued outcome, even despite diminished desire for the outcome (demonstrated in separate consumption tests; e.g., Adams, 1982; Dickinson et al., 1995; Thraillkill & Bouton, 2015). Some translational studies have found that instrumental habits can be experimentally induced in human subjects through extended training (e.g., Tricomi et al., 2009), though this finding has sometimes been difficult to replicate (de Wit et al., 2018). Taking a novel approach to this issue, recent work by Luque et al. (2020) suggests that in humans, instrumental habits may not be overt, but may instead be felt through their interference with goal-directed responding, and this interference can indeed strengthen with overtraining.

When we examined habits beyond the traditional instrumental sense in Chapter 2, we found evidence in line with the idea that ‘habit-like’ processes may not present as overt
behavioural choices; they can operate at the level of attention, depending on the specific revaluation regime. The training procedure (and duration) used in Chapter 2 could result in a pattern of conditioned attentional bias that persisted independent of outcome value (following devaluation, for example) or one that was partly mediated by a representation of current outcome value, and partly independent of this value (following value-reversal). Hence, in Chapter 3, we were interested in establishing whether attentional prioritisation can become increasingly outcome-independent as training proceeds. In Experiment 5, participants received either a moderate or extended amount of training in an online, RT-based version of the attentional revaluation task. We examined how moderate versus extended training influenced the extent to which participants updated their attentional control settings following outcome devaluation or value-reversal.

Consistent with our previous findings in Chapter 2, and in line with previous work on VMAC, a moderate amount of training led to an attentional bias for the high-value distractor over the low-value distractor during the training phase. However, for the groups given extended training, the effect of reward on attention was less clear. While we expected that the VMAC effect would be stable or (if anything) increase with longer training (cf. Le Pelley et al., 2015, Experiment 2; Watson et al., 2019), the magnitude of the VMAC effect in the training phase varied between revaluation groups; most notably we did not find a significant VMAC effect in the NoRev condition (that is, the group of participants who did not experience a subsequent change in outcome values). This finding was unexpected since all revaluation conditions had received equivalent treatment in this initial training phase, and hence we interpreted these findings to likely reflect noise. Patterns of attention in the test phase, following outcome revaluation, were also unclear. In the moderate training condition, the pattern of mean RTs on high- and low-value distractor trials was largely in line with evidence of sensitivity to changes in outcome value: relative to the NoRev condition, participants who experienced devaluation and value-reversal of outcomes showed numerically smaller VMAC effects. However, none of these group differences reached significance, and in fact the NoRev condition no longer exhibited a significant VMAC effect ($p > .08$), which we interpreted again to be a product of noisy data. This lack of a significant test-phase VMAC effect in the NoRev condition was also observed for those who received extended training. These findings for the NoRev groups marred our intended analyses because these conditions were intended as the primary ‘baseline’ conditions against which we would compare changes in attention as a consequence of devaluation and value-reversal.
following extended training. Hence, we were unable to meaningfully assess the effects of extended training on attentional updating following devaluation or value-reversal. We posit that the data in this study largely reflect noise due to poor participant engagement with a repetitive and somewhat demanding task, which was completed by participants online on their own devices without experimenter supervision. For the extended training group in particular, it seems plausible that these participants responded more inaccurately because they became fatigued or bored from the sheer number of training trials they had to complete. Hence, the revaluation instructions and test phase may have come at a time in the experiment where some participants were already disengaged from the task, and this would have meant that they were more likely to respond inaccurately during the test phase.

Future work will benefit from running the experiment again with some changes. I believe the greatest benefit will come from conducting the experiment in-person, which would 1) give more experimenter control over potential distractions, and 2) provide participants with a version of the task (eye-tracking or RT-based) wherein they are incentivised for good performance more directly, earning points/money for each correct response as was the case in Experiments 1, 2, 3, and 6 (rather than a possibility of a bonus for better-than-average performance). In addition, it would be beneficial to conduct testing for the extended condition over multiple testing days (an approach used in previous studies of overtraining in humans – e.g., Le Pelley et al., 2015, Experiment 2; Luque et al., 2020).

However, these changes would have significant resource implications, given the experimenter time required for running a long, eye-tracking study over multiple testing sessions with the number of participants needed for six experimental conditions; such a study was outside the time and resource confines of my PhD. Hence, while Experiment 5 served as a first step towards addressing the question of whether attentional prioritisation can transition to come under habit-like control as a consequence of extended training, at present this issue remains unclear and needs further work.

5.1.3 Does attentional prioritisation of reward-related stimuli result in ‘habit-like’ delayed disengagement?

Previous chapters in this thesis have focused on the idea that reward learning can modulate attentional capture. However, attentional capture is only one facet of a whole set of cognitive processes involved in attention that will influence performance in a search task with salient distractors. Once captured by a salient distractor, attention must disengage from the
distractor in order to move elsewhere (Born et al., 2011; Posner & Petersen, 1990; Wang et al., 2019). Investigating the effect of reward on attentional disengagement, Watson et al. (2020) used a variant of the VMAC procedure that placed a reward-signalling distractor at the centre of the search display on each trial, ensuring a more statistically robust and pure measure of disengagement distinct from attentional capture (relative to previous approaches; Le Pelley et al., 2015; Theeuwes & Belopolsky, 2012). Watson et al. (2020) found that participants were significantly slower to move their eyes away from the high-reward distractor than the low-value distractor, even though this pattern was counterproductive since slower responses meant that participants were less likely to earn reward in this task. This pattern even endured when rewards were removed in a subsequent unrewarded test phase. The implication of these findings is that participants found it harder to disengage attention from the high-reward stimulus than the low-reward stimulus. Hence, in addition to shaping the ability of stimuli to automatically ‘pull’ attention to themselves, reward can also modulate the extent to which stimuli ‘hold’ attention (Watson et al., 2020). This influence of reward on delayed disengagement occurs even when it is counterproductive to participants’ goals, and once learned, will persist even in extinction (wherein rewarding outcomes are removed). Clearly, these features are in line with what would be expected of a ‘habit-like’ response.

In previous chapters, we found that attentional capture can be ‘habit-like’ in its persistence following devaluation, and partially ‘habit-like’ (and partially outcome-mediated) following a value-reversal manipulation. Furthering this line of enquiry, in Chapter 4 we examined whether attentional prioritisation of reward-related stimuli in the form of delayed disengagement would show the key feature of a ‘habit-like’ response. Namely, we examined whether delayed attentional disengagement is mediated by retrieval of the current value of the outcome signalled by a stimulus, or whether instead prioritisation becomes divorced from the value of the events involved. Building on the variant of the VMAC task used by Watson et al. (2020), we introduced outcome elements that mediated between stimuli (colours) and rewards (high- vs. low-value points), allowing us to keep the relationship between a stimulus and an outcome constant while changing the value of that outcome by manipulating how much money each fruit was worth through devaluation or value-reversal.

We found that during the training phase, the high-value distractor counterproductively ‘held’ attention for longer than the low-value distractor. This pattern of attention was learned based on participants’ knowledge of the relationships between the distractor stimuli, fruit
outcomes, and the values of fruits. Once formed, this pattern of attentional prioritisation became inflexible to changes in outcome value, even when participants were able to report explicit knowledge of these changes in value. Specifically, participants still took significantly longer to disengage attention from the high-value distractor than the low-value distractor even after the high-value outcome had been devalued, mirroring findings on the insensitivity of attentional capture to devaluation found in Experiment 1. Strikingly, delayed disengagement from the high-value distractor persisted even following value-reversal of outcomes, which differed from the partial sensitivity of attentional capture to value-reversal observed in Experiment 3. This finding raises the possibility that reward may have distinct effects on attentional capture and disengagement; disengagement in particular may be more susceptible to habit-like persistence following value-reversal given the same amount of experience of reward learning, relative to attentional capture.

So then, what possible processes may give rise to this difference in persistence following value-reversal? We hypothesise that, following the instructed value-reversal, the test phase of the attentional revaluation task may be tapping into the conflict between two task sets. First, an inflexible ‘habit-like’ task set developed through extensive experience in the training phase may be biasing attention towards the high-value distractor over the low-value distractor (in line with the outcome values in the training phase). By contrast, a newer task set based on explicit knowledge of the revised values may bias attention towards the low-value distractor over the high-value distractor (in line with the revised outcome values that these distractors signal following value-reversal). The competition between these two task sets may be a dynamic process that changes over time: the strong, inflexible task set may dominate early attentional processing, only to be later overcome by the weaker task set that is aligned with the new outcome values. Different points in this dynamic process may be emphasised in the attentional capture task versus the attentional disengagement task. Specifically, in the disengagement variant of the attentional revaluation task, participants begin each trial with their attention already at the location of the reward-signalling distractor; by contrast, in the capture variant of the task, participants must move their gaze from the central location to the distractor – which will take a certain amount of time. Consequently, attentional processes will be engaged by the distractor more rapidly in the disengagement variant. As a result, behaviour in this task may be more sensitive to earlier influences of reward on attentional processing; whereas in the capture variant there is more time for slower processes to operate before a saccade is programmed and executed. A dynamic account of
this sort can thus explain why, in the disengagement task, the more inflexible ‘habit-like’ task set would dominate to continue prioritising the high-value distractor in an outcome-independent manner.

However, it is important to note that based on the current findings, we are unable to make definitive conclusions on the interdependence/independence of these two components of attention, since we examined attentional capture and disengagement in separate studies with different sets of participants – and moreover, the Bayesian evidence in favour of the null hypothesis in the study of disengagement (Experiment 6) was only at anecdotal levels. To more clearly assess this possibility, future work could examine the processes of attentional capture and disengagement in a highly powered, within-subjects design – for example, by including trials wherein distractors are located peripherally (to examine attentional capture, as in Experiments 1 to 3) and other trials wherein distractors are located centrally (to examine attentional disengagement, as in Experiment 6). This future study could then examine whether these two facets of attention differ significantly in their sensitivity to outcome revaluation using the attentional revaluation task presented in this thesis. What also remains to be seen is whether delayed disengagement will be amenable to retraining, as was the case in Experiment 3 where we saw a ‘full’ reversal of attentional patterns following value-reversal with feedback; this remains a possible avenue for future research.

5.1.4 ‘Habit-like’ Attentional Prioritisation and Sign-Tracking

Altogether, our findings suggest that the attentional prioritisation response can be both flexible and inflexible to changes in outcome value, depending on the specific revaluation regime and the component of attention that is required by the task (i.e., attentional capture versus disengagement). The idea that these experimental conditions influenced the extent to which attentional prioritisation was sensitive to changes in outcome value could accommodate previous animal work on sign-tracking. As discussed in Chapter 1, there are considerable parallels between sign-tracking and VMAC, wherein the signals of reward (i.e., a lever signalling a food pellet, or a coloured distractor signalling a high-reward outcome) can come to be prioritised over the goal (i.e., the location in which food pellets are delivered, or a diamond-shaped target) despite having no operant value in producing the reward. Notably, previous research by Amaya et al. (2020) found that sign-tracking could reduce in an outcome-mediated manner if the context in which the devaluation was learned was similar to the testing context; this suggests that the updated outcome values must be “online” during
test if behaviour is to update accordingly. By contrast, if the taste aversion to the devalued outcome failed to transfer to the testing context, then sign-tracking persisted in an outcome-independent manner.

At first glance, these findings of context-dependency from the animal literature do not map easily onto the current data on VMAC in humans. The context in which our training and test phases were conducted – in terms of the cover story, way the task was presented to participants, and provision of feedback, for example – was equivalent across studies, and yet we observed sensitivity to outcome revaluation in some cases but not others. That said, one way of reconciling our findings with those of Amaya et al. (2020) is to take a more abstract view of ‘context’, wherein the context for our human participants was shaped by the values of the fruit outcomes – specifically, whether or not the value regime featured a relative difference in outcome values. In cases where revaluation involved a reversal of outcome values (e.g., Experiments 3 and 4), there was a relative difference in values in both the training and test phases, which might constitute a similar ‘context’: here we observed sensitivity to outcome revaluation. By contrast, in cases of devaluation or supervaluation (Experiments 1 and 2), there was a relative difference in outcome values during the training phase but not during the test phase, potentially rendering the context of training and test phases was more distinct: and here we observed insensitivity to outcome revaluation. Clearly, this is a very post-hoc account that attempts to fit our data to the view suggested by Amaya et al. (2020), which is not without its problems. For example, even if we were to accept the account that differences in outcome value can contribute to context, it fails to explain why under conditions of value-reversal, we found sensitivity of attentional capture (Experiment 3) but not attentional disengagement (Experiment 6). Notwithstanding what exactly defines the ‘context’ for our human participants, the studies in this thesis are in line with the idea that sensitivity to outcome revaluation is not a binary measure, but may instead lie on a continuum ranging from a sensitive, outcome-mediated component and an insensitive, outcome-independent ‘habit-like’ component, and specific experimental conditions may shift the relative balance between these two extremes (Daw et al., 2011; Thrailkill & Bouton, 2015).
5.1.5 Overall Summary

In summary, the findings presented in this thesis provide important insights into the nature by which reward-signalling stimuli can elicit attentional prioritisation. In the forms of attentional capture and attentional disengagement, it is clear that attentional prioritisation of reward-related stimuli can persist independently of the value of the events involved. However, the conditions under which this outcome-independent prioritisation will occur are relatively nuanced. In the case of attentional capture, the specific value regime used to change participants’ desire for outcomes will influence the extent to which ‘habit-like’ outcome-independent attentional prioritisation is observed (Chapter 2). If outcomes become of equal value (through devaluation or super-valuation), then greater attentional capture by a stimulus signalling high-reward (over a stimulus signalling a low-value reward) will be allowed to persist; in these cases, there is little incentive for participants to engage in effortful updating of attentional priority, and instead existing settings are allowed to run in an outcome-independent manner. By contrast, if there is a relative difference in the values of outcomes (through value-reversal), then attentional settings controlling capture will be updated to reflect these new values given incentive to earn a relatively more valuable outcome over another. But the effect of this outcome-mediated updating of attentional control will be opposed by an inflexible ‘habit-like’ process that continues to prioritise the (previously) high-value distractor in an outcome-independent manner. On this account, further training (i.e., provision of specific outcome feedback in the test phase) is required for updating of this ‘habit-like’ component of conditioned attention, and hence full reversal of the pattern established in training. Whether the relative balance between persistence versus updating of attentional control settings shifts in favour of ‘habit-like’ persistence with more training remains a direction for future work (Chapter 3). Interestingly, the specific value regime used to change participants’ desire for outcomes may exert less of an influence on the persistence of delayed attentional disengagement from high-reward signalling stimuli, relative to attentional capture (Chapter 4). Slower attentional disengagement from a stimulus signalling high reward (over a stimulus signalling a low-value reward) seems to persist following both devaluation and value-reversal, suggesting attentional prioritisation in the form of disengagement may be outcome-independent in a more general fashion than attentional capture.
5.2 Implications of the Current Findings

While the above theoretical insights are interesting in their own right, it is also important to consider the potential broader implications of the findings in this thesis. In this section, I will discuss what the current findings mean for our understanding of the relationship between reward and attention, and its role in influencing behaviour and psychopathology.

5.2.1 Attentional Economics

As discussed at the very beginning of my thesis in Chapter 1, what we pay attention to will have knock-on effects for our behaviours. Almost every step of our lives can be seen as a decision, from the trivial (What should I wear to work today?) to those with far-reaching consequences (What is the next step in my career?). A number of information-processing steps are required to reduce a potentially overwhelming number of options in order to eventuate in a final decision. One of the first steps in any decision will be information-gathering about what our options are, and this is where attention is likely to play a central role. Since attention is tasked with the role of narrowing and filtering the vast number of stimuli around us, it will mould the choice alternatives in any decision (Krajbich, 2019; Pearson et al., 2022; Rangel et al., 2008). Stimuli that are prioritised by our attention (e.g., our phone notification sounds) are likely to feature in our set of behavioural options, and become more likely to be chosen as targets of our behaviour (to reply to calls and emails). Conversely, stimuli that are not prioritised by our attention (e.g., the withering plant in the corner of your office) are less likely to be targets of action. In line with this idea that attention plays a crucial role in decision-making, previous work has demonstrated that choice alternatives that receive more attention are more likely to be selected as targets of choice (Newell & Le Pelley, 2018; Shimojo et al., 2003). This role of attention in shaping our choice alternatives, and hence biasing the likelihood that stimuli will be eventually selected (or not) as targets of overt behavioural choice has been termed attentional economics (Pearson et al., 2022).

According to this view, understanding why particular stimuli come to be the targets of our behaviour necessitates an understanding of how and why particular stimuli are prioritised by our attention early on in the decision-making process. Findings in the current thesis are in line with the idea that reward learning will exert an influence over which stimuli come to be
prioritised by attention. As demonstrated by studies of VMAC (reviewed in Chapter 1) and replicated in the training phases of experiments in this thesis, reward-related stimuli elicit rapid prioritisation by attention, particularly stimuli signalling high-value outcomes (relative to those signalling low-value outcomes). Hence, these stimuli associated with (high) value will be more rapidly and easily highlighted by our attentional systems for further processing and consideration down the line of decision-making (see Gluth et al., 2018, 2020).

Importantly, findings in my thesis highlight that once trained, attentional prioritisation of reward-related stimuli can become ‘habit-like’ in its independence from the value of the events involved. This outcome-independent nature of attentional prioritisation means that (under some conditions) it will be insensitive to acute changes in our desire for the outcomes. That is, even after outcomes change in value, stimuli associated with the outcomes may continue to be prioritised by our attention, and hence may continue to be given increased weighting as options for behavioural selection.

Consider an example I often experienced throughout my candidature. I often walk through the main walkway of UNSW Sydney, where I am surrounded by bright colours, movements, sounds, and the general commotion of other students. However, what I often find is that my attention is captured by the cafés that line the walkway, and this prioritisation of the cafés occurs by virtue of their association with the rewarding stimulation that coffee has provided during my candidature. I may find that my desire for coffee changes one day because I believe I have become over-reliant on it, it has become too expensive, or my favourite barista has moved elsewhere, for example. The findings in this thesis suggest that despite a change in my desire for coffee, the coffee-signalling café signs lining the university walkway will continue to be prioritised by my attention over other stimuli. According to the attentional economics account, through receiving additional weighting and significance through rapid attentional prioritisation early on in a decision-making cascade, the cafés will continue to be considered as a prominent option for behavioural selection, even if my desire for coffee has reduced. By shedding insight into how and why our attentional systems prioritise some stimuli over others, the current findings ultimately add to a fuller picture of how and why our attention shapes which stimuli come to be the targets of our behavioural choice.
5.2.2 Clinical Implications

The influence of ‘habit-like’ attentional prioritisation on decision-making also has potential clinical relevance. As discussed earlier in Chapter 1 of this thesis, we live in an era of constant access to rewards, such as food, alcohol, drugs, and gambling. With each reward, there are inevitably a growing number of stimuli that come to be associated with these rewards through learning: for example, drug-related stimuli (e.g., certain places, people, or paraphernalia associated with drug use) can become associated with the pleasurable feelings of the drug itself. Through repeated pairings of stimuli and rewarding outcomes, these reward-related stimuli can become imbued with incentive salience, wherein these stimuli gain salience through a change in their perceptual representation (Berridge & Robinson, 1998). It has been suggested that the increased incentive salience of drug-related stimuli may play a role in reward-seeking behaviours implicated in drug addiction. Animal research on sign-tracking behaviour – which, as discussed in Section 5.1, is characterised by preoccupation with signals of reward that have no operant value in producing the reward – has found that the propensity for an animal to sign-track has been associated with neurobiological and behavioural markers of addiction (when compared to goal-tracking; Flagel et al., 2007, 2008, 2009; Robinson & Flagel, 2009).

Lab-based VMAC tasks, including the attentional revaluation task featured in this thesis, have been used to investigate this idea in humans. These tasks generally investigate attentional prioritisation of stimuli that signal monetary rewards, since it has been suggested that the same cognitive processes underlie the increased attentional priority given to both drug-related and monetary-related stimuli (Anderson, 2016). Existing work on VMAC has shown that reward-related attentional biases are associated with maladaptive drug-seeking behaviours, such as illicit drug use (Krajbich, 2019; Pearson et al., 2022; Rangel et al., 2008) and risky alcohol use (Albertella, Watson, et al., 2019). Moreover, the magnitude of attentional biases to both monetary- and drug-related cues have been shown to predict the likelihood of abstinence success (Albertella et al., 2021; Marissen et al., 2006; Waters et al., 2003). Together these findings suggest a relationship between drug-seeking behaviours implicated in addiction, and attentional processing that acts to highlight reward-related stimuli more generally. One possibility is that VMAC reflects a predisposition – i.e., people who are more likely to have their attention captured by reward are more susceptible to developing and/or maintaining addictive behaviours. Another possibility is that addictive
behaviours result in sensitisation of the reward system more generally (Anderson et al., 2013).

The current findings add to this picture, providing insight into the nature of attentional biases to drug-related stimuli. As discussed above, with experience of the relationship between stimuli and rewarding outcomes, attentional prioritisation of reward-related stimuli can become ‘habit-like’ in its independence from the value of the events involved, and hence persistent even in the face of changes in the desire for the outcome. By this account, someone who smokes may over time associate smoking-related cues (like the sign for the tobacco store down the street) with the pleasurable feelings associated with smoking. Our findings suggest that even when one explicitly decides and knows that cigarettes are no longer of value or aligned with one’s goals, attentional prioritisation of smoking-related stimuli may persist; smoking-related objects in the environment can continue to capture attention and delay attentional disengagement. As a consequence of this bias in the initial process of information-gathering, the scales may be tipped in favour of action tendencies toward smoking (heading into the tobacco store). While this account does not imply that one cannot simply discount the prioritised stimuli, it is plausible that this would be more difficult to do under particular state and trait variables. Indeed, the magnitude of VMAC has been found to be associated with variables such as trait compulsivity (thought to be a transdiagnostic contributor to the maintenance of addictive behaviour; Albertella, Le Pelley, et al., 2019), as well as a higher working memory load (a moment-to-moment state variable; Watson, Pearson, Chow, et al., 2019), suggesting that under these conditions, attentional biases may render people more susceptible to potentiating drug-seeking behaviours.

The idea that reward-related stimuli may continue to be prioritised by our attentional systems despite a change in the value of the outcome also accommodates the theoretical view that drug-seeking is driven by excessive goal-directed behaviour, rather than excessive ‘habitual’ behaviour. As discussed in Chapter 1, drug choice has been found to be reliably increased by motivation to alleviate aversive withdrawal states (e.g., Hogarth et al., 2017; Hutcheson et al., 2001) and negative mood (e.g., Hardy et al., 2017; Hogarth et al., 2017; Hogarth & Hardy, 2018), and willingness to work to pay for drugs (e.g., Gray & MacKillop, 2014; MacKillop & Murphy, 2007, 2007; Murphy et al., 2011). Our findings suggest that ‘habit-like’ processes may exist at the level of attentional prioritisation of drug-related stimuli, and this outcome-independent prioritisation may increase the likelihood that
individuals will choose to engage in (potentially) goal-directed drug-seeking behaviour as the end point of the decision-making cascade. In effect, this argument suggests that habit-like attentional processes may act to prime the goal of drug-seeking, which then elicits corresponding goal-directed behaviours (cf. Cushman & Morris, 2015): e.g., a smoker whose attention is automatically captured by the sign outside a tobacco store is more likely to entertain the goal of smoking (despite a conflicting desire to abstain) than a smoker who does not notice the store.

While on the one hand, it is possible that ‘habit-like’ conditioned attention may bias decision-making in favour of drug-seeking behaviour, our findings also highlighted that (under certain conditions) conditioned attention can be retrained and overcome. In Chapter 2, when participants received a value-switch manipulation and additional experience of the stimulus-outcome pairings, the attentional bias towards the reward-signalling distractors adjusted in line with their current values. What this suggests is that individuals may be able flexibly adapt their attentional control settings to reflect the values of outcomes if there is sufficient incentive (with outcomes of relative difference) and retraining. This idea is in line with an explosion of interest in the addiction literature towards modifying attentional biases implicated in drug-related cues (review by Wiers et al., 2013). Some have found that training-based interventions can reduce additional biases to alcohol-stimuli within laboratory settings (Field et al., 2007; Schoenmakers et al., 2007), generalise to other untrained alcohol-stimuli, and result in longer periods of sobriety (relative to no intervention; Schoenmakers et al., 2010), although notably, some have cast doubt on the efficacy of these interventions (e.g., review by Christiansen et al., 2015). Notwithstanding the ongoing debate on this topic, our findings suggest that these interventions would benefit from incorporating alternative (more desired) reward-related stimuli which attention may prioritise over (less desired) drug-related stimuli.

5.3 Limitations and Future Directions

As discussed above, our findings on ‘habit-like’ attention have theoretical implications for our understanding of the relationship between reward and attention, and also potentially broader implications for our understanding of decision-making and psychopathology. However, there are some limitations to the findings in this thesis. Moreover, there is also much to be understood about outcome-independent processes in
attention, and how these processes can influence behaviour. Here, I will discuss limitations of the work in this thesis, and provide some directions for future work that may further this line of research.

5.3.1 Habits as Null Effects

In the outcome devaluation procedure, habitual behaviour in the test phase is defined by insensitivity to the devaluation. For example, in the classic rodent paradigm of lever pressing for a food-pellet that has been devalued, behaviour is said to be goal-directed if the rat reduces lever-pressing following devaluation, in line with the reduced value of the food pellet outcome; conversely, behaviour is said to be habitual if the rat fails to suppress responding for the now-devalued outcome, despite no longer desiring the food (as demonstrated in a separate consumption test). In the attentional revaluation task used in this thesis, conceptually based upon the outcome devaluation procedure, we relied on a null effect between a ‘baseline’ condition who experience a change of outcome values (the not-revalued condition) versus devalued, supervalued, or value-reversal conditions. For all outcome devaluation procedures, it has been recently highlighted that it is problematic to rely on a null effect when comparing responding for devalued versus still-valuable outcomes as an indicator of habits (Balleine & Dezfouli, 2019; Robbins & Costa, 2017; Watson et al., 2022).

One of the main issues with using a null effect to define habits is that there are likely many situations where sensitivity to acute changes in outcome value are observed that are not reflective of a habit; perhaps the participant has not yet learned task contingencies or is confused (Watson et al., 2022; De Houwer 2018). We aimed to address this issue by ensuring that participants passed check questions on the new outcome values before proceeding to the test phase, and including explicit knowledge checks at the end of each experiment; these were used as guardrails against the possibility that a failure to update attentional control settings was simply a consequence of poor contingency knowledge. In addition to all participants passing the check questions, we found that (the vast majority of) participants in each experiment reported correct explicit knowledge of the changes in outcome value at the end of the experiment, suggesting that participants had explicitly encoded the information about the updated values of the outcomes, and retained that knowledge throughout the test phase. Moreover, when we simplified the associative structure of the task (by removing the mediating fruit outcomes in Experiment 4) we obtained a similar pattern of findings –
consistent with the idea that a failure to update was not simply a consequence of participants’ poor understanding of the task contingencies.

One of the other issues with defining habits as null effects is that non-significant differences between conditions may reflect statistical weaknesses (such as being underpowered) as opposed to habits per se. We tried to control for this possibility by using Bayesian analyses of non-significant findings in order to quantify the strength for the null (that there were no between-group differences). Overall, this issue of habits as a null effect is inherent to the outcome devaluation procedure (not specifically the studies presented in this thesis) and remains an ongoing source of discussion in the field (reviews by Balleine & Dezfouli, 2019; Watson et al., 2022).

5.3.2 Individual Differences in Sensitivity to Outcome Revaluation

As discussed in Chapter 2, and in Section 5.1.1 of the current chapter, a key finding was that attentional settings controlling capture will be updated if there is a relative difference in the values of outcomes (through value-reversal), since under these conditions there is incentive to earn a relatively more-valuable outcome over a less-valuable outcome. However, the effect of this outcome-mediated updating of attentional control will be opposed by an inflexible ‘habit-like’ process that continues to prioritise the (previously) high-value distractor over the low-value distractor in an outcome-independent manner; full reversal of the established attentional pattern requires further training (i.e., provision of specific outcome feedback in the test phase).

An alternative possibility is that the observed pattern of equivocal attentional capture by the high- and low-value distractors following value-reversal reflects a mixture of participants: perhaps some participants showed a ‘full’ reversal of their attentional settings, while others did not. Studies in this thesis lacked the power to examine these potential individual differences in patterns of sensitivity to the value-reversal, as data collection for the number of participants required to answer this question was beyond the scope of my PhD. Future work could collect data from larger sample to examine whether distribution of sensitivity following value-reversal is normal (with participants showing partial sensitivity) or bimodal (with individual differences in a ‘full’ reversal versus insensitivity).
5.3.3 Knowledge Tests versus Consumption Tests

As briefly discussed above, and in Chapter 2, we used knowledge tests following the test phase, wherein participants were asked to select the current value of each fruit outcome. We used these knowledge tests to verify that participants had explicit knowledge of the updated outcome values that persisted through the whole of the test phase. However, previous studies of instrumental habits in humans have often instead used a ‘consumption test’ to assess the efficacy of devaluation (based on the approach used in animal studies that assesses willingness to consume a now-devalued food outcome: e.g., Adams & Dickinson, 1981). In a typical consumption test in human studies, participants are shown two outcomes (one still- valuable, the other devalued) and are asked to select the outcome they would prefer to receive – on the rationale that, if devaluation has been effective, participants will choose the still-valuable outcome – with these consumption test trials being interspersed with ‘standard’ test trials (assessing conditioned responding) during the test phase (e.g., Gillan et al., 2015; Luque et al., 2017).

In this thesis, knowledge tests were used instead of consumption tests because we believe that using explicit knowledge tests has some advantages over this previous approach: 1) the explicit knowledge check is arguably a more conservative test of knowledge of outcome values, since it requires participants to reflect on and report this knowledge for each outcome individually and directly, whereas the consumption tests may be open to additional (potentially more automatic) influences that are invoked by the requirement to choose between items, that go beyond explicit knowledge about the values of those outcomes, and 2) by assessing knowledge after the test phase – rather than at points during that phase – our procedure allows us to verify that participants’ knowledge spanned the whole of this critical phase. Nonetheless, future work could incorporate additional consumption test trials into the attentional revaluation task for a closer parallel to existing work on outcome devaluation.

5.3.4 Instrumental Attentional Habits

As discussed in Chapter 1, the experiments in my thesis used the VMAC procedure originally developed by Le Pelley et al. (2015), wherein the critical associative relationship between the distractor stimuli and the delivery of reward is Pavlovian in nature. That is, distractors signal the availability of reward, but participants are not required to respond to the
distractors themselves in order to earn this reward. ³ By contrast, however, habits are considered in the behavioural neuroscience literature as involving instrumental relationships, wherein responses are directly reinforced with the delivery of a rewarding outcome; that is, there is a causal relationship between the response and the outcome. For example, in a typical rodent study of habits, lever pressing (an instrumental response) is trained to be reinforced by the delivery of a food pellet (a rewarding outcome). To reflect this distinction, I referred to the attentional patterns observed in my thesis experiments as being ‘habit-like’, rather than reflecting “true” instrumental habits as they are traditionally understood.

Future work in this area could investigate whether habits in a traditional instrumental sense can operate at the level of attention. An instrumental version of the attentional revaluation task may require participants to make a saccade towards the reward-signalling colour-singleton shapes, in order to earn high- or low-value outcomes: that is, the type of response (saccade to blue, versus saccade to orange) would determine the outcome that was delivered. Thereafter, the values of the outcomes can be changed (through devaluation or value-reversal, for example) in order to test whether the response has become divorced from the representation of the outcome it produces. Using an instrumental procedure may also allow for an examination of the effect of reinforcement schedules on the development of attentional habits. As noted in Chapter 3, the transition from goal-directed to habitual behaviour following extensive instrumental training has been observed primarily under specific experimental parameters. Animals trained on interval schedules of reinforcement (wherein behaviour is reinforced after a certain period of time) are more likely to demonstrate behaviour insensitive to outcome devaluation, relative to animals trained on ratio schedules of reinforcement (wherein behaviour is reinforced after a certain number of responses; Dickinson et al., 1983). Moreover, behaviour appears to remain sensitive to devaluation even after extensive training if animals have been trained on two responses for different outcomes (Colwill and Rescorla 1985; Kosaki and Dickinson 2010). Hence, future work using an instrumental variant of the attentional revaluation task may be able to elucidate the effect of

---

³ The issue is slightly complicated here because the VMAC task does involve an instrumental component – participants must respond to the target in order to earn reward – with the distractor colour acting as a Pavlovian signal of what the outcome of this instrumental response will be. In this sense, the distractors in the current tasks are perhaps most accurately described as discriminative stimuli that modulate the instrumental contingency between target and reward.
these training schedules on the propensity for human habit formulation at the level of attention.

5.3.5 Assessing Habitual Attention Using the Belief Criterion

As discussed in Chapter 1, the relative balance between goal-directed and habitual systems can influence reward-related behaviour (Dickinson, 1985; Wood & Rünger, 2016). In order for a response to be considered goal-directed (as opposed to habitual), the response is performed when the agent desires a valued outcome, and when the agent believes that the response will obtain the outcome (de Wit & Dickinson, 2009; Heyes & Dickinson, 1990). The desire criterion captures the idea that goal-directed actions are controlled by the current motivational value of the outcome, in that goal-directed actions are performed when the expected outcome is desired or valuable. The outcome devaluation procedure is used as a test to establish whether responses meet the desire criterion; this procedure was adapted to the attentional revaluation task in my thesis, wherein following training on the task, outcomes were devalued, super-valued, or switched in value.

According to the belief criterion on the other hand, goal-directed behaviour is based upon knowledge of the causal relationship between the action and the outcome — i.e., that the response will cause the goal (de Wit & Dickinson, 2009; Dickinson & Balleine, 1994). To diagnose whether behaviour is goal-directed or habitual according to the belief criterion, researchers have used a contingency degradation procedure. In a typical contingency degradation procedure, rodents are trained to press two levers, each which delivers a rewarding outcome (for example, pressing a left lever delivers sucrose, and pressing a right lever delivers food pellets; e.g., Corbit & Balleine, 2000). After this initial training phase, one of the instrumental contingencies is degraded by delivering the outcome in the absence of action (making it ‘noncontingent’); for example, sucrose is delivered to the rodent regardless of whether it presses the left lever or not (non-contingent), whereas food pellets continue to be delivered only if the animal presses the right lever (contingent). If behaviour is goal-directed by virtue of being mediated by a representation of the causal relationship between the action and the outcome, rodents will selectively decrease responding on the lever that delivered the non-contingent outcome (since they no longer ‘believe’ that this action causes delivery of the outcome), while lever pressing for the contingent outcome will remain intact. By contrast, if response rates on the two levers remain similar, this indicates that the performance was insensitive to the degradation of the contingency and hence behaviour is not
mediated by a belief in the relationship between action and outcome – and so may be under habitual control.

A future study could use this contingency degradation procedure to assess whether conditioned attention can be habit-like according to the belief criterion; that is, whether it will occur independently of knowledge of the causal relationship between the action and the outcome. This would require only a relatively simple change to the existing attentional revaluation task: rather than devaluing outcomes in the test phase, the task would instead be changed so that one outcome (e.g., a lemon) was sometimes delivered regardless of whether the corresponding distractor was present in the search display, hence degrading the Pavlovian contingency between distractor and outcome. Likewise in an instrumental version of the task (see Section 5.3.3) one could arrange for lemons to be delivered regardless of whether participants make a saccade to the appropriate colour, degrading the instrumental contingency between response and outcome. This line of enquiry would provide insight into whether attention may be ‘habit-like’ or habitual in its failure to meet both the desire criterion (as demonstrated in this thesis, under some conditions) and the belief criterion, or whether attentional habits are more specifically insensitive to changes in desire only.

5.3.6 Impact of ‘Habit-Like’ Attention on Choice

As discussed with regard to attentional economics in Section 5.2.1, reward has an influence over which stimuli in our environment receive rapid attentional prioritisation, and hence can bias the likelihood that stimuli will be eventually selected (or not) as targets of overt behavioural choice. The findings in my thesis suggest that stimuli associated with (high) value will be more rapidly and easily highlighted by our attentional system for further processing. And once trained, attentional prioritisation of reward-related stimuli can become ‘habit-like’ in its independence from the value of the events involved – even after outcomes change in value, stimuli associated with the outcomes will continue to be prioritised by our attention, and hence may continue to be given increased weighting as options for behavioural selection. Notably, the idea that habit-like attention can have an impact on behavioural choice has not been formally tested in the current thesis, and this is a potentially fruitful area for future research. This future study could look at the effect of changing outcome values on the flexibility or inflexibility of actual behavioural choice, and aspects of choice such as accuracy and decision time (an area already receiving attention in decision-making literature; Gluth et al., 2018).
5.3.7 ‘Habit-Like’ Attention in Clinical Populations

As discussed in Section 5.2.2, the concept of ‘habit-like’ attention has potential relevance for our understanding of maladaptive behaviours implicated in clinical disorders. Habitual behaviour has, in recent years, been a rapidly growing avenue of interest in the clinical domain, where it has been suggested that a propensity for habits underlies the compulsive behaviours implicated in a range of psychopathologies, particularly drug addiction. Given the potential importance of the distinction between goal-directed actions and habits in understanding motivated behaviour, more recent translational work has attempted to investigate this distinction in human behaviour. Existing work on habits in humans has primarily examined overt behavioural ‘slips’ of goal-directed action (such as keypresses for devalued outcomes) as a measure of habitual behaviours. These kinds of tasks have found reduced sensitivity to outcome devaluation in clinical populations characterised by compulsive behaviours, such as cocaine-dependence (Ersche et al., 2016), alcohol-dependence (Sjoerds et al., 2013), Tourette’s syndrome (Delorme et al., 2016), OCD (Gillan, Apergis-Schoute, et al., 2015; Gillan et al., 2011), and Parkinson’s disease (de Wit et al., 2011). It remains to be seen whether ‘habit-like’ processes operating earlier on in the decision-making process – at the level of attention – are more pronounced in these clinical populations, relative to non-clinical controls. Future work in this area may compare performance between clinical versus non-clinical populations on the attentional revaluation task. If it is found that clinical populations show greater persistence of attentional prioritisation following outcome devaluation (in the form of capture or disengagement) then this would suggest that habit-like attention may indeed play a role in tipping the scales in favour of drug-seeking behaviour, and hence the maintenance of maladaptive behaviours implicated in these populations.
5.4 Concluding Remarks

Stimuli signalling high-value rewards are more likely to be prioritised by our attentional system than signals of low-value rewards, a phenomenon termed value-modulated attentional capture (VMAC). Viewed another way, reward-related stimuli can be said to automatically elicit a conditioned attentional prioritisation response. In the current thesis, I aimed to explore whether this conditioned attentional prioritisation response could be ‘habit-like’, in the sense that it becomes divorced from the value of the events involved. The results reported here demonstrated that, in line with previous work on VMAC, our attentional system prioritises stimuli signalling high-value outcomes over stimuli signalling low-value outcomes. Under some conditions, this pattern of attentional prioritisation – in the forms both attentional capture and attentional disengagement – can persist despite acute changes to the values of (and hence participants’ desire for) outcomes. These findings suggest that attentional prioritisation of reward-related stimuli can occur independently of the value of the events involved and hence it can be considered – at least in part – ‘habit-like’. In turn, our findings provide a novel framework to further understand the nature of the process by which stimuli elicit attentional prioritisation, and how this outcome-independent prioritisation may play a role in shaping overt behaviour.
References


https://doi.org/10.1523/JNEUROSCI.20-11-04233.2000


https://doi.org/10.1073/pnas.1506367112


de la Tourette syndrome. *Brain, 139*(2), 605–615.

https://doi.org/10.1093/brain/awv307


https://doi.org/10.1016/j.bbr.2007.07.022

https://doi.org/10.1007/s00213-006-0535-8

https://doi.org/10.1080/17437199.2013.876238


Hogarth, L. (2020). Addiction is driven by excessive goal-directed drug choice under negative affect: Translational critique of habit and compulsion theory. *Neuropsychopharmacology, 45*(5), Article 5. https://doi.org/10.1038/s41386-020-0600-8


