2024

A NOVEL MEASURE OF HABITS AND GOAL-DIRECTED CONTROL

Wood, Katie

https://pearl.plymouth.ac.uk/handle/10026.1/22608

---

http://dx.doi.org/10.24382/5241
University of Plymouth

COPYRIGHT STATEMENT

Copyright and Moral rights arising from original work in this thesis and (where relevant), any accompanying data, rests with the Author unless stated otherwise[1].

Re-use of the work is allowed under fair dealing exceptions outlined in the Copyright, Designs and Patents Act 1988 (amended)[2], and the terms of the copyright licence assigned to the thesis by the Author.

In practice, and unless the copyright licence assigned by the author allows for more permissive use, this means,

- that any content or accompanying data cannot be extensively quoted, reproduced or changed without the written permission of the author / rights holder; and
- that the work in whole or part may not be sold commercially in any format or medium without the written permission of the author/rights holder.

Any third-party copyright material in this thesis remains the property of the original owner. Such third-party copyright work included in the thesis will be clearly marked and attributed, and the original licence under which it was released will be specified. This material is not covered by the licence or terms assigned to the wider thesis and must be used in accordance with the original licence; or separate permission must be sought from the copyright holder.

The author assigns certain rights to the University of Plymouth including the right to make the thesis accessible and discoverable via the British Library's Electronic Thesis Online Service (EThOS) and the University research repository, and to undertake activities to migrate, preserve and maintain the medium, format and integrity of the deposited file for future discovery and use.

---

[1] *E.g.* in the example of third-party copyright materials reused in the thesis.

[2] In accordance with best practice principles such as, *Marking/Creators/Marking third party content* (2013). Available from: https://wiki.creativecommons.org/wiki/Marking/Creators/Marking_third_party_content [accessed 28th February 2022]

A NOVEL MEASURE OF HABITS AND GOAL-DIRECTED CONTROL

By

KATIE M. J. WOOD

A thesis submitted to the University of Plymouth
in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

School of Psychology

December 2023

Acknowledgments

I would firstly like to thank the School of Psychology at Plymouth University for providing the funding to facilitate the research reported in this thesis.

I also want to thank my supervisor, Professor Chris Mitchell, for his support and guidance throughout my PhD. I will miss our coffees in the Hairy Barista! Thank you also to Dr Tina Seabrooke for all your help – particularly with programming the E-Prime experiments (I couldn’t have done it without you). Thanks also to the MSc and undergraduate students who helped to collect the data reported in Experiments 3, 4, 6 and 8. A big thank you also to the tech office for their invaluable support getting my experiments online throughout the pandemic lockdowns. Without, Mark Cooper and Anthony Mee, it would not have been possible to continue my data collection, so I am incredible grateful – thank you Mark and Anthony for all your patience, determination, and programming skills!

Thanks also to my friends and family for their continuous support and understanding, despite my absence over the past four years. A special thank you to Gabriel for putting up with the endless conversations about my thesis and providing me with copious amounts of tea to get me through it. Thank you to my lovely family: Mum, Dad, Jessie, Granny, Andrew and Josie, and everyone else who always believed I could do it.

Author's Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Doctoral College Quality Sub-Committee.

Work submitted for this research degree at the University of Plymouth has not formed part of any other degree either at the University of Plymouth or at another establishment.

This study was financed with the aid of a studentship from the University of Plymouth.

Related publications:

Wood, K. M., Seabrooke, T., & Mitchell, C. J. (2023). Action slips in food choices: A measure of habits and goal-directed control. *Learning & Behaviour*, *50*, 1-13. https://doi.org/10.3758/s13420-023-00573-5

Word count of main body of thesis: 34,420

Signed:

Date: 22/12/2023

Abstract

A novel measure of habits and goal-directed control

Katie M. J. Wood

Repetitive habitual behaviour is thought to occur even when this directly conflicts with goals. This is termed an action slip. The current research aimed to present a simple procedure that captures action slips and is easy to implement and interpret. Chapter 1 reviews the current literature. Chapters 2-4 then report nine experiments that aimed to test our novel procedure. A congruency effect, where participants' performance was good on congruent trials but comparatively poor on incongruent trials, was consistently found in our experiments (with one exception), providing evidence of action slips. Chapter 2 details three experiments that used an Outcome (O)-Stimulus (S) delay manipulation. Although evidence for the effect of O-S delay was not strong, the numerical pattern was consistent across experiments, where shorter delays showed a bigger congruency effect on accuracy. Chapter 3 explores a load manipulation in three experiments, which aimed to reduce participants' working memory capacity during the experiments, and the effects of time pressure in a fourth experiment. The results provide some evidence that the congruency effect can be increased by reducing working memory capacity. Chapter 4 explores a devaluation version of the experiment, followed by a further experiment that manipulates the amount of training. When the devaluation procedure was used, the congruency effect was not observed. When we returned to the standard testing procedure and manipulated the amount of training, we observed a congruency effect, even after a short amount of training. This is inconsistent with the S-R account and dual process theory. Chapter 5 explores whether individual differences are associated with performance in our experiments: self-reported habitual behaviour (COHS), goal-directed control (HSCQ), impulsivity (BIS-11), perceived stress (PSS), depression (PHQ-9), and anxiety (GAD-7). Our results found no association between action slip scores and self-reported habitual behaviour, nor any of our other individual differences measures. Participants with high non-planning impulsivity scores performed less accurately across the experimental trials, which could be interpreted as a lack of goal-directed control. Chapter 6 discusses the question of whether the congruency effect seen in our Experiments is evidence to support a dual-process account of habitual behaviour.

Table of Contents

List of Tables

List of Figures

Chapter 1 - Introduction

In an ever-changing world, flexibility and the ability to adapt to a new environment is extremely valuable for both humans and animals alike. Performing familiar responses automatically can be beneficial as it frees up cognitive capacity for achieving new goals. However, repetitive, habitual behaviour is thought to occur even when this behaviour directly conflicts with intended actions (Wood et al., 2022) – this is termed an action slip (Reason, 1979). In a longitudinal study, Ji and Wood, (2007) found that participants repeated habitual behaviours even if they reported intentions to do otherwise. For example, when participants had formed strong habits to take the bus, they continued to do so regardless of how they personally framed their intentions to reduce their carbon footprint. A further example of how action slips can occur in everyday life is driving to work. The action of driving to a regular place of work becomes so automatic after repetition that it no longer requires a great deal of conscious effort, freeing up cognitive capacity to plan the day ahead. When the driver reaches a traffic light, they automatically stop when the light turns red, as this behaviour is so engrained through repetition that it requires little conscious thought. When the traffic light switches from red to green, this prompts the driver to turn left towards their place of work. However, when the individual starts a new job, which requires turning right at the traffic lights, they may accidentally turn left when the light turns green, reverting to their habitual route instead of their new goal. The current thesis focuses on this kind of situation, in which a putatively habitual response (turning left to the old job) is set against the goal (turning right to the new job) - an action slip.

Habitual behaviour can sometimes have negative implications on an individual. Examples of medical implications of habitual behaviour which directly contradict with living a healthy lifestyle include binge eating, smoking, and drinking excessively. Furthermore, existing research has found that individuals high in depression demonstrate greater habitual behaviour in the face of stress (Heller et al., 2018). Repetitive negative thinking (RNT) is a widely studied transdiagnostic mechanism that has been identified as a risk factor for depression and anxiety (Ehring & Watkins, 2008). RNT is defined as thinking about negative content in a way that is repetitive, passive or difficult to control

(Bell et al., 2023). Research exploring the emergence of RNT (Watkins & Roberts, 2020) suggests that these thought patterns initially begin as a goal-directed response to stress-inducing cues, but over time the repetition of these patterns of thoughts can lead to the automatic association of the RNT cues, such as negative mood, with engaging in RNT as a habitual response style (Roberts et al., 2021).

While there are varying definitions of habitual behaviour, experts in the field of associative learning differentiate between goal-directed actions influenced by the present motivational value of outcomes and ingrained habitual behaviours triggered directly by the surrounding environment (Watson & de Wit, 2018). An important question to consider is how habitual responses can be distinguished from genuine mistakes. One possible way of defining the difference between a habitual response and a mistake is by simplifying behaviour into responses made to stimuli. An example of a mistake could be failing to process the stimulus, and consequently generating a random response to that stimulus, generating a response to a different stimulus, or making no response at all. Another example of a mistake could be where the stimulus is processed, but the correct response to that stimulus was too weak to be generated. Therefore, the response is either random, or there is no response at all. An example of a habit, on the other hand, is where the stimulus is processed and the correct response is available, but the correct response is not executed because an inappropriate response is generated instead. That inappropriate response may have been appropriate at a previous point in time or in a different context.

The aim of this thesis is to explore evidence of habitual behaviour in humans. This thesis therefore begins by discussing theories of habitual verses goal-directed behaviour, existing evidence for these theories, current methods of measuring habitual behaviour in a lab-based environment and evidence for factors that influence habitual behaviour. The design of our new computer-based task is then described. Chapters 2-4 report nine experiments that test whether our computer-based task shows evidence of habitual behaviour. This thesis also aims to explore whether the amount of habitual behaviour observed varies with a range of individual differences. Therefore, Chapter 5 investigates whether there are correlations between performance on our computer-based task (reported in Chapters 2-4) and scores on a number of self-report individual difference measures: Creature of Habits Scale (COHS), Habitual Self Control Questionnaire (HSCQ), Barratt Impulsiveness Scale (BIS-11), Perceived Stress Scale (PSS), Patient Health Questionnaire (PHQ-9), Generalised Anxiety Disorder Assessment (GAD-7). The results described in the empirical chapters 2-5 are discussed in Chapter 6. To foreshadow the results, Chapter 6 concludes that whilst our experiments appear to show evidence of action slips, there is some indication that our congruency effect, showing good performance on congruent trials but comparatively poor performance on incongruent trials, may be due to competing goals in a single goal-directed system, rather than two distinct systems.

Habitual behaviour verses goal-directed action

To understand habitual behaviour, it is important to firstly introduce the existing theories surrounding habitual behaviour verses goal-directed actions. According to Heyes and Dickinson (1990) a goal-directed action is an action this is focused on achieving a desired outcome. These actions often involve cognitive effort and motivation to focus actions on a desired outcome (Marien et al., 2015). Habitual behaviour, on the other hand, may allow an individual to perform responses with less cognitive effort and may be beneficial for freeing up cognitive capacity. Functional imaging studies suggest that automatic habitual behaviour is associated with less uncertainty, less brain activity and more efficiency (Saling & Phillips, 2007). However, habits can also be detrimental to individuals (Hogarth & Chase, 2011), preventing behavioural change in the face of new situations and goals (e.g., habitual binge eating whilst attempting to lose weight). Therefore, it is important to understand the factors that determine when goal-directed actions might be undermined by habits. This will allow us to identify ways to promote positive behaviour that help us to achieve goals like living a healthy lifestyle, rather than slipping back into unhelpful habits, such as binge eating, smoking and drinking excessively. It is also important to consider the empirical question of how long it takes to train a habit. Lally et al (2010) conducted a study aiming to find to out how long it takes to form a habit. They found the average modelled time to reach a plateau of automaticity was 66 days, and the range was from 18 to 254 days. This indicates that it can take a substantial amount of time to form a habit. However, if the definition of a habit is the execution of a previously learned response that is inappropriate in the current context, amount of training may not be a necessary condition for a habit.

Therefore, the current research aims to address the question of whether amount of training influences habitual behaviour.

Goal-directed action

An intuitive explanation of goal-directed actions is that they are actions that are performed to achieve a desired goal or outcome. A general theory is that the goal-directed system is slow to respond and effortful (Heyes & Dickinson, 1990), but highly sensitive to changes in the environment. An example of this could be picking up your mobile phone with the purpose of sending a message to someone. The action of picking up the phone is done due to the belief that this will lead to the desired outcome of sending a message. These behaviours, therefore, rely on a combination of belief and desire (Heyes & Dickinson), where an individual desires an outcome (sending a message) and they believe an action (picking up their phone) will achieve that outcome. According to this belief-desire model, if you reduce the desire for the outcome, then the action will cease. Similarly, if you reduce the belief that the action will result in the outcome, the action will also cease. You can reduce either or both. It does not matter which one of them is reduced. Reduction of either of them will reduce goal-directed responding, but not automatic habitual responding.

A way of testing for the presence of goal-directed action is by using the outcome revaluation procedure (de Wit & Dickinson, 2009; Dickinson, 1985). This procedure was first demonstrated using rats as subjects. Adams (1982) provided a good example of the outcome revaluation procedure, where evidence of goal-directed action is shown when reducing the value of the outcome also reduces responding for that outcome. Conversely, the absence in a reduction in responding would implicate non-goal-directed behaviour. In Adams' first experiment, rats received either moderate or extensive training to press a lever to obtain sucrose pellets. For half of the rats, the sucrose pellets were then devalued by pairing them with lithium chloride-induced sickness. For the remaining animals, the sucrose pellets remained valuable. The number of lever-press responses performed by each group was then assessed in an extinction test (where lever presses did not produce pellets). This extinction test was used to assess whether information about the relationship between the lever press and the sucrose pellets – an action-outcome relationship – had been encoded. The rats that received moderate training performed fewer lever press responses when the sucrose was devalued than when it was still valued. Thus, the rats demonstrated goal-directed control after moderate instrumental training. Conversely, rats given extended training did not show this sensitivity to outcome value, a finding that will be examined below.

The theory behind the outcome revaluation procedure is that goal-directed actions are driven by the intention to achieve a particular outcome. If an individual believes that a particular action will achieve the desired outcome, then they will perform that action. If the outcome is no longer desirable, or no longer believed to be available through this specific action, the individual will stop making the action. In the moderate training groups of Adams' experiments, rats no longer desired the outcome and therefore reduced their responding for that outcome.

Habit theory

The dominant theory of habits is the Stimulus-Response (S-R) theory. This theory received a great deal of attention between 1920 and 1950 (e.g., Hull, 1933; Miller, 1935; Spence, 1950; Tolman, 1932; Watson, 1924) and focuses on the relationship between external stimuli (e.g., lever) and an individual's behavioural responses (e.g., lever press). This theory, which is closely associated with behaviourism, proposes that behaviour is primarily a result of the direct influence of environmental stimuli on an individual's responses (Treisman, 1960). According to S-R theory, the association between a stimulus (e.g., lever) and a response (e.g., lever press) strengthens if the response is reinforced with an outcome (e.g., food reward). The more a response is performed with a desired outcome, in the presence of the stimuli, the greater the stimulus-response association will be. After sufficient reinforcement, the stimulus will later trigger a response. This is predicted to occur even if the subject does not believe that the outcome is available, or if the subject no longer desires the outcome. The presentation of S triggers the R due to the link that allows activation to travel from the S to the R (Holland, 2008). This S-R association means that the response has become automatic in the presence of the stimulus. Automatic responses can be defined as unintentional, uncontrolled, goal-independent, autonomous, purely stimulus driven, unconscious, efficient, and fast (Moors & De Houwer, 2006). In contrast to goal-directed actions, therefore, habitual responses are quick and effortless. However, habitual behaviour is likely to make errors when the environment changes and old responses are no longer appropriate. When a change in goal requires performing a different action, the pre-commuted S-R will persist, leading to habitual selection of outdated responses (Hardwick et al., 2017). For instance, an S-R response could relate to the example given above of automatically driving to an old place of work. When the new goal is to drive to the new place work, the individual may still drive to the old place of work, as this is the habitual response. Hence, an action slip occurs.

As well as evidence for goal-directed action, the outcome revaluation procedure described above can be used to test for S-R associations, as these should be insensitive to devaluation. Adams' (1982) outcome devaluation experiments appear to provide evidence for habitual control. Moderate training resulted in fewer responses for the devalued outcome compared to the still-valued outcome. Extensive training, by contrast, resulted in no devaluation effect on responses. Thus, overall, moderate training demonstrated goal-directed control, but extensive instrumental training demonstrated habitual control. One way to think about this is that the S-R association has become so strong that the response was performed automatically (we will discuss limitations of this theory below).

Dual-process theory

To capture both goal-directed actions and S-R behaviour, the dual process theory was developed. Dual-process theories of instrumental learning propose that instrumental behaviours are controlled by two distinct systems: the goal-directed system and stimulus-triggered habitual system (e.g., Dickinson & Balleine, 1994; Verplanken, 2018). As discussed above, the goal-directed system is suggested to facilitate deliberate, reward-driven actions, while the habitual system is suggested to produce comparatively automatic responses based on stimulus-response (S-R) associations. Habits can be helpful or unhelpful (depending on whether the environmental contingencies change). In this way, dual-process theories of instrumental learning capture a wide range of both adaptive and counterproductive reward-based behaviours.

As discussed above, Adams (1982) used the outcome revaluation procedure to provide evidence for both the goal-directed and habitual components of instrumental behaviour. This, therefore, could be considered evidence for dual-process theory. Adams' findings, in which moderately trained but not extensively trained rats showed sensitivity to outcome revaluation, have since been replicated (Dickinson etl., 1995; Thrailkill & Bouton, 2015). It is traditionally regarded as evidence of a habitual response in the extensively trained groups that is driven by an S-R association.

Adams' (1982) outcome devaluation procedure has since been translated for use with human participants. Valentin et al.'s (2007) participants, for example, learnt to perform two instrumental actions to obtain different drink rewards. One of the rewards was then devalued by allowing unlimited consumption. In a subsequent extinction test, participants responded more for the still-valued outcome than the devalued outcome, a clear demonstration of goal-directed control. This finding has since been conceptually replicated many times (e.g., de Wit et al., 2012, 2018; Liljeholm et al., 2015; Piray et al., 2016).

In contrast to the evidence for goal-directed control in humans, clear evidence for habits in humans has been hard to come by. Tricomi et al. (2009) used a procedure very similar to that of Adams (1982) to obtain initial promising evidence of habits. Their participants learnt to press one of two buttons to obtain pictures representing food which was to be consumed after the experiment. An instrumental button press (R1) in the presence of one fractal stimulus (S1) produced a picture of chocolate (O1), while a different instrumental button press (R2) in the presence of a second fractal stimulus (S2) produced a picture of crisps – O2 (S1: R-O1, S2: R-O2). After completing either moderate or extensive training of these contingencies, one of the outcomes was devalued through selective satiation (participants ate a lot of chocolate or crisps). In a subsequent extinction test, participants who had received moderate training performed significantly fewer instrumental responses for the devalued outcome than the valued outcome, showing evidence for goal-directed behaviour. Participants who received extensive training, by contrast, showed no significant difference in response rates for the still valued and devalued outcomes. Therefore, the devaluation did not reduce responding in the extensively trained participants. Similar to Adams' findings in rodents, extended training promoted habitual behaviour.

Although Tricomi et al. (2009) provide exciting evidence for the dual-process theory of instrumental learning, subsequent research has not been so convincing. In particular, de Wit et al. (2018) reported five failures to induce behaviour that was insensitive to an outcome devaluation manipulation through overtraining. Their final two experiments were replications of Tricomi et al.'s study (barring minor details). As de Wit et al. acknowledged, one can always argue that more extensive training may have produced evidence of habitual control. The main conclusion from these studies, however, is that it is by no means easy to experimentally induce habits in humans through overtraining.

de Wit et al. (2007; see also de Wit et al., 2013) also tried to show evidence for insensitivity to devaluation but using quite a different procedure. Here, participants first learned to perform different instrumental responses (R1 and R2) to earn points. Each trial began with a picture of a closed box and a fruit icon (denoted here as fruits A-H) on the door lid. Once the participant performed an instrumental response, the door opened to reveal the contents of the box. If the participant selected the correct response, a fruit stimulus was presented within the box, and the participant received 1-5 points, depending on the speed of their response. If the participant selected the wrong response, however, the box was empty, and the participant received zero points. Importantly, there were three trial types. On *congruent* trials, the discriminative stimulus (the fruit on the front of the box) matched the outcome (the fruit inside of the box, i.e., A: R1-A, B: R2-B). For example, a strawberry on the lid, if followed by the correct response, would be followed by a strawberry inside the box (and participants would earn points for the correct response). On *incongruent* trials, in contrast, the discriminative stimulus was paired with an outcome that served as a discriminative stimulus for a different response (C: R1-D, D: R2-C). For example, a banana on the lid (cue C) would be followed by an apple in the box (outcome D) and an apple on the lid (cue D) would be followed by a banana in the box (outcome C). Finally, on biconditional (*control)* trials, the discriminative stimuli and outcomes were unique (E: R1-F, G: R2-H), where the outcomes did not serve the role of discriminative stimuli, and vice versa. For example, a grape (cue E) might lead to a plum (outcome F) and a peach (cue G) might lead to a cherry (outcome H).

de Wit et al. (2007) tested the dual-process theory by incorporating an outcome revaluation test after the learning phase. The aim of this was to explore whether insensitivity to outcome devaluation could provide evidence for habitual control. This was quite different to the outcome devaluation procedure used by Tricomi et al. In de Wit et al.'s test, participants were presented with two open boxes at the start of each trial. Inside each box was a fruit, one of which had been the outcome associated with a left key response. The other had been an outcome associated with a right key response (i.e., outcomes A and B were presented on congruent test trials, C and D on incongruent test trials, and F and H on biconditional test trials). Importantly, during the test trials, there was a cross superimposed on one of those fruit outcomes. The cross signalled that the instrumental response that previously produced that fruit outcome no longer earned points. Hence, participants were to perform the other instrumental response; the one that previously produced the still-valued outcome (without a cross). Other than this, the test phase procedure was the same as the procedure during training trials. Once participants made their response, the box opened, revealing how the number of points earned.

de Wit et al. (2007) found that participants performed above chance on congruent and biconditional trials; they responded more for the valued outcome (on which no cross had appeared at the start of the trial) than the devalued outcome (with a cross). de Wit et al. suggested that the participants solved the congruent and biconditional tests using their goal-directed system (since their behaviour was governed by the current value of the outcomes). However, importantly, incongruent trials produced quite different behaviour; participants failed to respond more for the valued than the devalued outcome on incongruent trials. Hence, performance on these trials was suggested to be driven by the habitual system, which did not incorporate any representation of the outcomes.

de Wit et al.'s (2007) explanation of the failure to respond in a goal-directed fashion on incongruent trials is as follows. Crucially, the argument goes, the goal-directed system is subject to response conflict in the incongruent condition. Specifically, each fruit (C and D) was associated with both responses (R1 and R2), either as an outcome or as a cue. Hence, when fruit C was presented as a stimulus, it activated R1 directly (via a C-R1 link). But fruit C was also associated with, and would therefore activate, outcome D, which would then activate R2 (because D was the outcome following execution of R2). Activation of response R2 by stimulus C, via activation of outcome D, is referred to as an S-O-R (in this case C-D-R2) process. In other words, the presentation of C and D at the start of each incongruent test trial, one with a cross over it, did not provide the participant with any useful information; they could not base their responding on the current value of C and D because they may not have known with which responses these outcomes were associated. Hence, if they were to attempt to respond in a goal-directed fashion, their responding would have been at chance. All this complexity can be resolved, suggested de Wit et al, by responding simply on the basis of the S-R (e.g., stimulus C = R1) associations, which do not incorporate any representation of the outcomes. That is, the habit system took over on incongruent trials, and participants therefore performed R1 and R2 in the presence of C and D, respectively. Hence, the fruit icons served as discriminative stimuli rather than being associated with outcomes, and primed both instrumental responses equally, leading to the chance performance observed. The pattern of results observed by de Wit et al. (2007) has since been replicated many times (de Wit et al., 2009, 2011, 2012, 2013; Sjoerds et al., 2013).

Although de Wit et al.'s (2007) discrimination findings are robust, the extent to which the incongruent condition gives rise to habitual control has recently been questioned. De Houwer et al., (2018) argued that there are two kinds of outcomes on every trial: the fruit outcome and the points. They were concerned that the points may have been more salient than the fruit outcomes and so, during the learning phase, participants would learn which responses produced points, rather than which responses earned particular fruit outcomes. Therefore, De Houwer et al. argued that the chance performance seen on incongruent test trials (but not the congruent or control trials) was not evidence for non-goal-directed behaviour, but of behaviour directed at the goal of points (rather than specific fruit). In other words, participants may have been behaving in a goal-directed way, despite the observed result.

To explore whether an effect of outcome devaluation can be found in the Fabulous Fruit Task, De Houwer et al. (2018) replaced de Wit et al.'s (2007) outcome revaluation test with tests that devalued the *points* (in Experiment 1), rather than the fruits, to assess whether participants would show goal-directed control on incongruent trials under these circumstances. If the behaviour is goal-directed (and the goal is the points, not the fruit), then devaluing the points would reduce responding (relative to the other response). The test followed the same procedure as the discrimination learning task (but without feedback). However, the participants were either instructed that the points that were previously produced by each instrumental response would now be deducted from their score (Experiment 1), or that the instrumental contingencies were now reversed (Experiments 2 and 3). In each case, the participants successfully switched their response (above chance) on all trial types (congruent, biconditional and incongruent). Hence, when the test targeted knowledge of the points, rather than the fruit outcomes, participants demonstrated goal-directed control, even on incongruent trials. Therefore, one needs to be careful when inferring from the absence of an effect of outcome devaluation that the outcome (that participants are responding for) has been correctly identified. Overall, the Fabulous Fruit Task does not appear to provide any evidence for habitual/S-R behaviour. Only when the goal is correctly identified will an effect of devaluation be observed.

In summary, the traditional devaluation approach – the canonical assay of goal-directed versus habitual control – has, in two paradigms, failed to provide compelling evidence in humans for the habitual S-R controller that is proposed by the dual-process theory of instrumental learning. Tricomi et al. (2009) provided initial evidence of habitual control following overtraining in the traditional outcome devaluation task (e.g., Adams, 1982). This finding has, however, proven challenging to replicate (de Wit et al., 2018). In another line of evidence, de Wit et al. (2007, 2013) found promising signs of habitual control in their discrimination learning studies (the Fabulous Fruit Task), but subsequent research suggests that those studies lacked sensitivity to detect goal-directed control (De Houwer et al., 2018). Another issue is that both of these paradigms rely on null effects, or *absence* of differences between conditions, to provide evidence of habitual control. Tricomi et al.'s (2009) overtraining paradigm requires no difference in rates of responding for the valued and devalued outcomes, while de Wit et al.'s (2007, 2013) Fabulous Fruit Task requires responding to be no different from chance on incongruent trials. As De Houwer et al. have argued, such null effects *could* reflect habitual control, but they could equally reflect other factors such as poor learning. Thus, even when demonstrations of habitual control in humans have been reported, those demonstrations are subject to alternative interpretations.

The failures to experimentally observe habits in humans (e.g., de Wit et al., 2018), and the concerns around interpretations of null effects (e.g., De Houwer et al., 2018), prompted Luque et al. (2020) to recently develop a new assay of habitual control. They argued that investigating changes in overt responding after overtraining was impractical in experiments with human participants, because too much training would be required to *change* response choice. Instead, they argued that a more viable approach would be to explore evidence of habits in changes to *reaction times* (RTs). In Luque et al.'s (2020) experiments, participants initially learned different stimulus-response-outcome (S-R-O) relationships, where participants performed two instrumental responses (R1 and R2) to earn different types of fictitious diamonds, which translated into points of different value. In the presence of each stimulus, one response was optimal, in that it earned more points than the other response. In the presence of S1, for example, R1 responses might have earned diamonds that were worth 100 points, while R2 responses earned diamonds worth five points. Thus, both responses produced points on each trial, but the number of points received differed based on the response performed. On subsequent devaluation tests, the participants could continue earning the diamonds, but some of those diamonds were devalued, such that they now earned zero points. When the diamonds associated with R1 were devalued from 100 points to 0 points, the optimal response would be to switch to the alternative response (R2) to earn five points. In this scenario, participants switched their responses, showing sensitivity to outcome devaluation (in contrast to Tricomi et al.'s results). However, Luque et al. observed an *RT switch cost*, where choosing the alternative response for the non-devalued outcome came with slowed responding, particularly with extended training and when participants were placed under time pressure. Luque et al. therefore argued that, even when overtraining does not reveal evidence of habits in changes to overt response selection, habits can be revealed through slowed RTs. Unlike previous experiments, Luque et al.'s (2020) work suggests that habits can be revealed in humans, following outcome devaluation, through slowed RTs rather than response selection.

Wider literature

There are several examples of research exploring the topic of automaticity where the learning takes place outside of the laboratory. In these experiments, the learning has already happened when the participants arrive at the experiment. For instance, research has shown that automatic processes can be engaged in attentional tasks. The Posner Cueing task (Posner, 1980) is a classic paradigm used to study attentional processes (Anderson, 2018). It involves presenting a cue that directs attention to a specific location, followed by a target stimulus. The task measures how quickly and accurately participants respond to the target, depending on whether the cue was valid or invalid. A valid cue is where the cue correctly indicated the target location. An invalid cue is where the cue incorrectly indicated the target location. The Posner Cueing task is considered an example of automaticity in attentional tasks, as participants often show faster response times to valid cues, indicating that attention can be automatically directed by the cue without conscious effort (Anderson, 2018; Neumann, 1984).

The Stroop task is another paradigm often used to investigate learnt automatic processes (Stroop, 1935) and how these automatic processes can interfere with controlled tasks. In a Stroop task, participants are shown a list of colour words (e.g., “red,” “blue,” “green”) printed in various ink colours. The task is to name the colour of the ink, not the word itself. The reaction time is measured for both congruent stimuli (where the word’s meaning matches the ink colour, e.g., the word “red” is printed in red ink) and incongruent stimuli (where the word’s meaning does not match the ink colour, e.g., the word “red” is printed in blue ink). The Stroop effect refers to participants’ delay in reaction time when naming the ink colour of incongruent stimuli compared to congruent stimuli. This delay occurs because reading the word is an automatic process that interferes with the task of naming the ink colour and highlights the strength and automatic nature of reading, showing how difficult it is to suppress this automatic response (Augustinova & Ferrand, 2014).

Schneider and Shiffrin’s (1977) work on automatic and controlled processing is a further example of an approach which aimed to train automatic responses, where their studies showed that automatic processing develops with consistent practice and exposure. Schneider and Shiffrin conducted a series of experiments where participants were trained under consistent mapping conditions, where specific stimuli were consistently paired with specific responses. Over time, this consistent pairing led to the development of automatic detection of the target stimuli. In contrast, varied mapping conditions involved changing the pairing of stimuli and responses across trials. This required participants to use controlled processing, as they could not rely on automatic responses. Participants became faster and more accurate in detecting target stimuli under consistent mapping conditions, indicating that the responses had become automatic. Schneider and Shiffrin’s work highlights the importance of practice and consistency in developing automatic responses, providing valuable insights into how we learn and perform tasks efficiently (Schneider & Chein, 2003).

However, their procedure is complex and takes a considerable amount of time to train. Therefore, it would be beneficial to develop a procedure that is more easily implemented to investigate these processes.

A further example of tasks that have played an important role in exploring theories surrounding instrumental cue-driven response choice are Pavlovian Instrumental Transfer (PIT) tasks (Colwill & Rescorla, 1988; Corbit & Balleine, 2005). In a PIT task, two responses (R1 and R2) are initially trained to predict two outcomes (O1 and O2), resulting in R1-O1 and R2-O2 associations. In a subsequent test phase, instrumental response choice (R1 vs R2) is then assessed in the presence of Pavlovian stimuli (S1 and S2) that are associated with each of the two outcomes (S1-O1, S2-O2). The common result of these tasks is that, in the presence of Pavlovian stimulus S1, subjects are more likely to make response R1 than R2. Conversely, in the presence of S2, subjects are more likely to make response R2 than R1 (Watson et al., 2014). This has been considered evidence for insensitivity to outcome value (Corbit et al., 2007). However, there are other ways to explain these findings. Seabrooke and colleagues (2017) conducted three experiments using an outcome-selective Pavlovian-instrumental transfer (PIT) design and an outcome devaluation procedure. In Experiment 1, participants learned to perform one response to earn crisps points and another to earn popcorn points. Following this, one outcome was devalued by making it taste unpleasant. The results showed that participants' choices were biased towards the non-devalued outcome, indicating goal-directed control. Importantly, the extent of this bias was unaffected by the devaluation manipulation. In the second experiment, when stimuli indicated an equal likelihood of both outcomes, the cue-driven response choice was influenced by the devaluation manipulation. The third experiment reinforced these results by showing a selective avoidance of the cued, devalued outcome. Overall, Seabrooke et al.'s research supports a goal-directed model of PIT, where the expected probability and value of outcomes independently influence response choices.

In the current research, we were keen to pursue the idea that a habit, learned in a simple instrumental-learning task, could produce more than just a slowing of an optimal response, but also result in non-optimal/incorrect responding. Our approach was slightly different to that usually taken in the learning field. It is common to recognise the devaluation method as evidence for habitual behaviour. However, we did not manipulate outcome value. Rather, we used a simple interference procedure in which a putative habit, possibly S-R in nature, was set against a goal-directed action.

Current research

Our main aim is to present a simple procedure measuring habitual behaviour that is easy to implement and interpret and captures the behaviour described in the driving to work example above. In our experimental design, participants first learn two stimulus-response-outcome (S-R-O) relationships, as shown in Table 1. The stimuli are the background colours of the screen (green or blue), the responses are joystick movements (left or right), and the outcomes are pictures of food rewards (Pringles or jellybeans) presented on screen. The participants are instructed to move the cursor as quickly as possible to earn points towards the food rewards. On S1 trials, R1 responses produce O1 points (e.g., if green then left responses produce jellybean points), while R2 responses produce no points. Conversely, on S2 trials, R2 responses produce O2 points (e.g., if blue, then right responses produce Pringles points), while R1 responses earn no points.

Table 1. Experiment design: 2-way congruency (congruent, incongruent)

Instrumental Training			**Test**			
Stimulus	**Response**	**Outcome**	**Outcome**	**Stimulus**	**Correct Response**	**Test Trial**
S1	R1 →	O1	O1	S1	R1	Congruent
	R2 →	Ø		S2	R1	Incongruent
S2	R2 →	O2	O2	S2	R2	Congruent
	R1 →	Ø		S1	R2	Incongruent

Note. S1 and S2 represent the screen colour (green/blue), R1 and R2 represent instrumental responses (left/right), and O1 and O2 represent food outcomes (jellybeans/Pringles). Ø represents “no outcome”.

In the test phase, each trial begins with the presentation of either outcome O1 or O2. The participants are instructed to perform the response that, during the training phase, produced the presented outcome (i.e., O1 → R1, O2 → R2). Hence, the presentation of the outcomes at the start of each test trial communicates to participants which outcome they could earn on that trial. Importantly, participants are not allowed to respond immediately. They are required to withhold responding until the screen changes to the colour of one of the stimuli S1 or S2 (therefore, S1 and S2 play the same role as the traffic lights in the driving to work example above). The presented stimulus (which signals to the participant they can now respond) could be either congruent or incongruent with the outcome that had been presented at the start of the trial. Therefore, it is also congruent or incongruent with the correct response on that trial. As demonstrated in Table 1, on a congruent trial, O1 might be followed by S1, both of which are associated with R1 in training. Conversely, on an incongruent trial, O1 might be presented first, indicating that R1 is the correct response but S2, associated with R2, would signal that a response could now be executed. Hence, any S-R link that might form in training would promote the incorrect response on incongruent trials on test. Importantly, participants are informed that they should ignore the colour of the stimulus on test, and only use the presentation of S1 and S2 to signal *when* to respond, not *how* to respond. We chose screen colour as the stimulus to ensure participants process the stimulus. All participants were screened for colour blindness before each experiment and screen colour was counterbalanced for S1 and S2. We thought there could be an issue with participants being able to detect the stimulus presentation without processing the colour (e.g., they may just notice an event occurs in the corner of the screen). Lastly, the number of jellybean and Pringle points earned depends on the speed of the correct responses to increase motivation to respond quickly. In this way, we expect incongruent trials to impair response accuracy relative to congruent trials. The difference in accuracy on the two trial types can be interpreted as the extent to which the stimulus (inappropriately) controls responding on incongruent trials – a measure of habit learning.

Our procedure is similar to that used by Brass et al., (2001) to investigate automatic imitation. In their experiment, Participants were presented with a number 1 or 2 on each trial, in the centre of a computer screen. On trials where the number 1 was presented, they were required to lift their left index finger off a keyboard as quickly as possible. On number 2 trials, they were required to lift their right index finger. Also on screen, in addition to the numbers, a pair of hands were present with all fingers pointing at the participant. On each trial, at the same time as the number 1 or 2 appeared, the index finger of one hand would rise. Participants were required to ignore this movement when making their own movement. The observed rising finger could be on the same side as that required to be moved by the participant (i.e., the finger of the hand on the left moved when a 1 appeared, and the finger of the hand on the right moved when a 2 appeared). These were congruent trials - the observed (to-be-ignored) finger movement matched the participant's finger movement. On incongruent trials, the observed moving finger was on the opposite hand to that of the correct response. (i.e., the finger of the hand on the left moved when a 2 appeared, and the finger of the hand on the right moved when a 1 appeared). Brass et al. found that participants responded faster and more accurately when the spatial location of the finger on screen matched the response they needed to make. For example, observing one finger moving on incongruent trials can lead to participants moving that finger, rather than the "correct" finger on that trial – an effect that could be described as "automatic imitation". Similar to this, evidence for habitual behaviour would be seen in our experiments when responses are determined by an S-R relationship, or habit, that undermines goal-directed action. In both cases, two responses are set against each other, one of which is 'correct' and the other of which is better practised in the sense that the S-R was trained, rather than the O-R.

Chapter 2 – O-S delay

Chapter 1 outlined the theory of habitual behaviour. It also described the design of our new experimental procedure for measuring habitual behaviour. In this new procedure, participants are told which outcome to respond for on each test trial but are not allowed to respond for that outcome until a stimulus is presented. The main manipulation was the nature of the stimulus that signalled that a response could now be made. On congruent test trials, that stimulus had been associated in training with the correct response. On incongruent test trials, that stimulus had been associated with the incorrect response. We anticipated that we would observe action slips/higher levels of incorrect responding on incongruent than congruent test trials. It was necessary to adapt the design of our experiments to allow for the most appropriate procedure. Therefore, we designed two separate general methods one for face-to-face testing and one for remote testing. The order of presentation of experiments is not chronological but chosen for logic and flow. Hence, the two methods will appear interchangeably throughout the following chapters.

Method 1 (in-person) was administered both pre- and post-pandemic lockdowns. The experiment was programmed using E-Prime (https://pstnet.com/) and was presented on a 22-inch desktop computer in a University of Plymouth laboratory, with an experimenter present. Instrumental responses were performed using a joystick. The reason a joystick was used was to ensure participants were not able to physically prepare their left/right response. For example, if a keyboard was used, they could simply take their finger off the key for the incorrect response. With a mouse, if the correct response was a right swipe, they could simply place the mouse right over to the left, so that a left swipe was impossible.

Method 2 (online) was administered during the pandemic lockdowns. We developed a web-based instrumental task to carry out our standard experimental procedure. The experiment was programmed by the University of Plymouth Psychology technicians using ASP.NET web pages, with a C# server-side code structure mixed with some in-page JavaScript. The programme ran on the University of Plymouth Net IIS internal web server with an MS SQL database hosted internally.

Participants completed the experiment online using either a desktop computer or a laptop. Left and right movements were made either with a mouse or a laptop trackpad.

Experiment 1 – O-S delay

Experiment 1 implemented our experimental design, but with an additional O-S delay manipulation. According to the dual-process theory, goal-directed control will be disrupted when there is less time to prepare responses in advance, as goal-directed behaviour is time consuming and effortful (Heyes & Dickinson, 1990), in comparison with automatic habitual behaviour, which is fast and effortless. Therefore, we aimed to see if manipulating the amount of time participants had to prepare their response would influence the number of action slips made. We added a long delay between presentation of the O and the S on half of the trials, and a short O-S delay on the other half of the trials. We anticipated that any congruency effect would be exaggerated by the short delay between the outcome and stimulus presentations. Shorter delays should present less opportunity for participants to prepare their response in advance and keep the goal in mind (Mittner et al., 2014).

In addition to applying a response delay manipulation, we created a baseline measure to compare against the congruent and incongruent trials. To do this, we included a control stimulus in the presence of which neither instrumental response produced points. If performance is hindered on incongruent trials by the priming of an incompatible response (by the stimulus), we would expect to see poorer performance in the incongruent trials than in the neutral baseline condition.

Method

Participants. Forty-one undergraduate psychology students (36 females, 5 males) from University of Plymouth participated in exchange for course credit. The participants were aged between 18 and 45 years (*M* = 20.68 years, *SD* = 5.25 years).

Design. The experiment followed a 3 (congruency: congruent, incongruent, baseline) by 2 (delay: short, long) repeated-measures design, as shown in Table 2. The dependent variables were response accuracy and RT during the test phase. The O-S delay manipulation was applied to test trials only.

Stimuli. As shown in Table 2, as well as the two main discriminative stimuli (S1 and S2), a third discriminative stimulus (S3) was used as the control stimulus. Green/blue screens were randomly assigned to S1 (Stimulus 1) or S2 (Stimulus 2) for each participant. S3 was the same screen colour for all participants – yellow. Pringles/jellybeans were randomly assigned to O1 (Outcome 1) or O2 (Outcome 2) for each participant. Left/right responses were also randomly assigned to R1 (Response 1) or R2 (Response 2) for each participant.

Table 2: Experiment 1 design: 3 (congruency: congruent, incongruent, baseline) by 2 (delay: short, long)

Instrumental Training			**Test**				
Stimulus	**Response**	**Outcome**	**Outcome**	**Stimulus**	**Correct Response**	**Congruency**	**Delay**
S1	R1 →	O1	O1	S1	R1	Congruent	Short delay Long delay
	R2 →	Ø		S2	R1	Incongruent	Short delay Long delay
S2	R2 →	O2	O2	S2	R2	Congruent	Short delay Long delay
	R1 →	Ø		S1	R2	Incongruent	Short delay Long delay
S3	R1 →	Ø	O1	S3	R1	Baseline	Short delay Long delay
	R2 →	Ø	O2	S3	R2	Baseline	Short delay Long delay

Note. S1, S2 and S3 represent the screen colour (green/blue/yellow), R1 and R2 represent instrumental responses (left/right), and O1 and O2 represent food outcomes (jellybeans/Pringles). Ø represents "no outcome".

Procedure. Experiment 1 was conducted using our in-person method (Method 1). At the beginning of the experiment, each participant was presented with a series of self-report measures of individual differences on-screen to be completed directly before the experiment. Details of these self-report measures will be discussed in Chapter 5. A joystick was then positioned in front of each participant. Participants were instructed to hold onto the joystick throughout the experiment and to keep the cursor central (unless they are making their responses). Before the instrumental training phase, participants were told that they could earn Pringles and jellybeans points by moving the cursor left and right, and that they would need to learn the relationship between the screen colour, response, and food picture. They were also told to respond as quickly as possible on each trial because the number of points earned are related to the speed to which they respond, as demonstrated in Table 3.

Table 3: Number of points earned during training, based on reaction time (RT)

Reaction Time (ms)	Points
Less than or equal to 200	10
201-400	9
401-600	8
601-800	7
801-1000	6
1001-1200	5
1201-1400	4
1401-1600	3
1601-1800	2
1801-2000	1
More than 2000	0

The participants completed six practice training trials (two of each training trial type). At the start of each practice training trial, the target zones for the instrumental responses were highlighted on the screen, as shown in Figure 1. These target zones were two rectangles, one on the far-left and one on the far-right, both the full height of the screen, which indicated to participants where they needed to move the cursor (to the left or right) for a response to be registered. Each training trial also began with a fixation cross, presented centrally and in grey, with the statement "*Please centralise the cursor*". Once the cursor was centralised, the fixation cross turned black (500ms), before being replaced by one of the discriminative stimuli S1 or S2 (background screen colour). Participants then moved the cursor as quickly as possible to make their left or right response, using the joystick.

During training, correct responses were reinforced with points. The number of points that the participants earned on any given training trial depended on how quickly they responded, as shown in Table 3 detailing the RT-points conversion rate. Once the response was made, a picture of the reward (jellybeans or Pringles) was presented, along with the points earned on that trial. An exception to this rule was that, if the correct response was registered after 2 s of the stimulus appearing onscreen, a message appeared stating that the response was too slow.

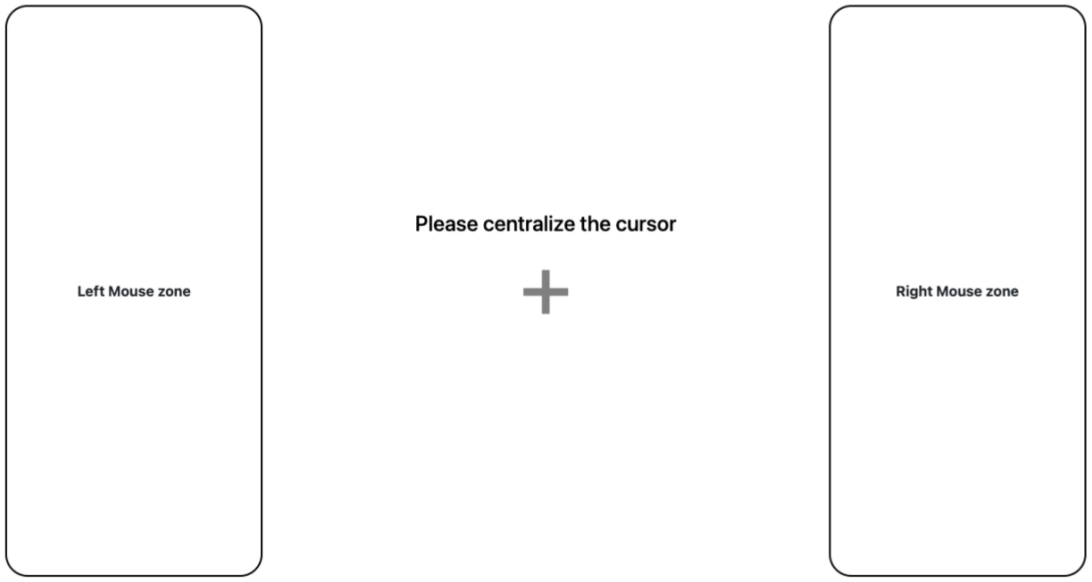


Figure 1: Example of screen layout showing target zones where participants make their instrumental responses during practice training trials.

The practice training trials were followed by 252 instrumental training trials. This training phase was separated into three blocks of 84 trials, with 28 trial types in each block, as shown in Table 2. Each block was separated by a two-minute break. The only difference between practice and training trials was that the screen no longer showed target response zones on the training trials.

At the end of the training phase, participants completed four blocks of 16 test trials. These 16 test trials consisted of four presentations of each baseline trial type (O1-S3 and O2-S3), and two presentations of each congruent trial type (O1-S1 and O2-S2) and incongruent trial type (O1-S2 and O2-S1). The interval between the presentation of the outcome (O) and the stimulus (S) on test was manipulated to create two conditions: short delay (2-3 seconds) verses long delay (10-11 seconds). The precise delay on any given trial was randomly determined between the upper and lower limits. Half of each trial types within each block were presented with a short O-S delay, the remainder with a long O-S delay. Each test trial began with the fixation cross from the training phase, indicating that participants should centralise their cursor. Once the cursor was centralised, one of the outcome pictures from the training phase (jellybeans or Pringles) was presented centrally for 2000ms to indicate which outcome the participant should respond for on that trial. Participants had to then perform the correct instrumental response, shown in Table 2, as quickly as possible once the screen changed colour and were told they would lose points if they moved the cursor before the screen changed colour. Participants were not told how many points they earned on each test trial, but they were told that points were still accumulating. The trial ended if the cursor was moved at any point before the onset of the stimulus, and the participants were informed that they had moved the cursor too early. At the end of the test phase, participants were given a randomly determined (range 3-6) number of each Pringles and jellybean reward for consumption.

Results

Exclusions. We chose to exclude participants who achieved less than 80% accuracy on the final test, because we were concerned that these participants may have misunderstood the instructions. Specifically, participants who performed poorly on the final test may have thought that they were to respond to the stimulus (screen colour) that was presented on each test trial, rather than the outcome. This would produce a large congruency effect (good performance on congruent trials but comparatively poor performance on incongruent trials), but it would not be a true demonstration of habitual responding. Participants who were simply making the same response for all trials would have scored 50% overall accuracy, which would be an indication of lack of understanding or effort rather than habitual behaviour. Similarly, those who misunderstood the instructions and responded consistently for the stimulus instead of the outcome picture would have scored 50%. In order to be cautious, we therefore took 80% accuracy during the test as a conservative cut-off for differentiating participants who showed a lack of understanding or effort from those showing habitual behaviour. Any demonstration of a congruency effect under these circumstances would be against the backdrop of good overall performance, which would suggest that the participant had understood the test instructions. Five participants were removed for achieving less than 80% accuracy on the test trials, leaving 36 participants (31 females, 5 males).

Training. As can be seen in Figure 2, participants achieved over 90% correct responses after the initials 42 trials and maintained that accuracy throughout the remaining practice trials. This indicates that they learnt the S-R-O contingencies quickly during the training phase. It is also important to note that they were training well beyond the point where they had reached asymptote.

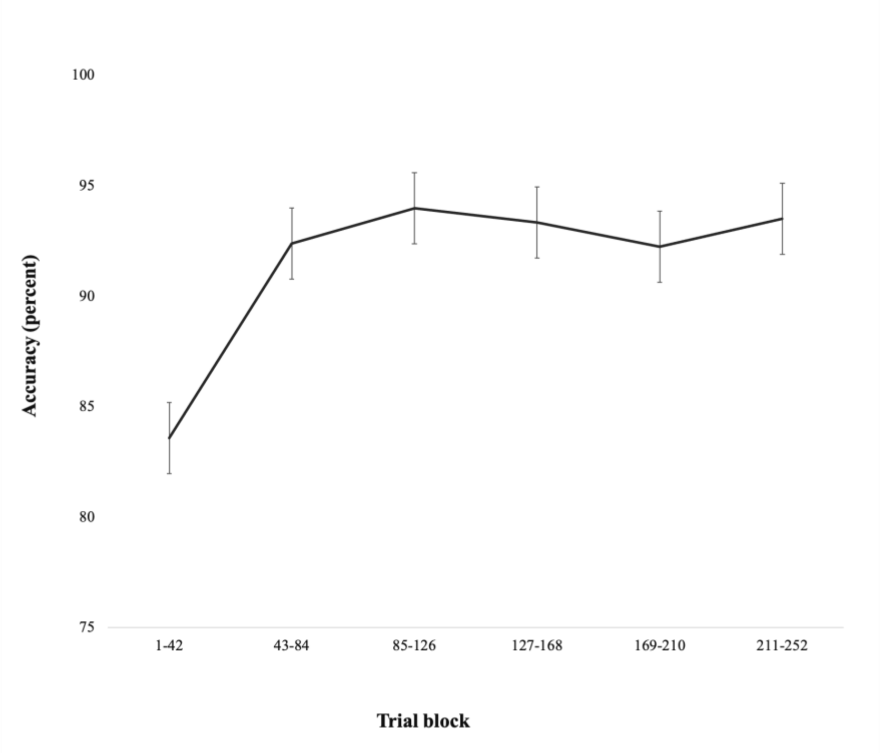


Figure 2: Line graph of mean accuracy (%) of training trials for S1 and S2 in Experiment 1 (excluding trials with baseline stimuli). Over 90% correct responses are shown after the initials 42 trials. Error bars represent the standard error of the mean for each training trial interval.

Test accuracy. Figure 3 shows the mean percentage of correct responses in the test, separated by congruency and delay condition. Error bars represent difference-adjusted, 95% within-subjects confidence intervals (Baguley, 2012). The graphs shows that accuracy was higher during congruent trials than baseline trials and incongruent trials. This was supported by a 3 (congruency condition: congruent, incongruent, baseline) × 2 (delay condition: short, long) repeated-measures ANOVA, which revealed a significant main effect of congruency condition F(2,70) = 9.93, p < .001, generalised eta squared (n_g^2) = .06. The Bayes Factor analysis provided extreme evidence for this congruency effect, BF_{10} > 100. In order to unpack the differences found in the ANOVA, a series of paired t-tests were run to compare the effects of two congruency conditions. Participants performed more accurately on congruent trials than baseline trials. t(71) = 3.05, p = .003, d_z = 0.47, BF_{10} = 8.76, and incongruent trials, t(71) = 3.72, p < .001, d_z = 0.69, BF_{10} = 59.05, as shown in Figure 3. Accuracy on baseline and incongruent trials did not differ, t(71) = 1.29, p = .20, d_z = 0.24, BF_{10} = 0.29. There was a significant main effect of delay, with higher accuracy on long trials than short trials, F(1, 35) = 4.62, p = .04, n_g^2 = .01, although these results should be treated with caution as the Bayes Factor was inconclusive, $BF_{10} = 1.04$. Figure 3 suggests that delay had a greater effect on accuracy during incongruent trials than congruent or baseline trials, with accuracy higher after a long delay than a short delay. However, the congruency by delay interaction was not significant, $F(2, 70) = 2.26$, $p = .11$, $n_g^2 = .02$. A Bayes Factor was inconclusive, $BF_{10} = 0.59$, which therefore, did not provide evidence of a null hypothesis.

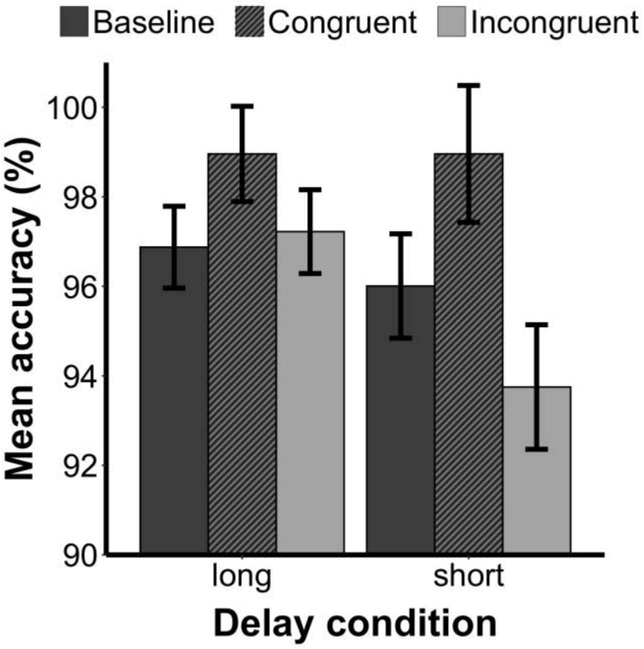


Figure 3: Bar graph of mean accuracy (%) for congruency and delay conditions in Experiment 1 (n=36). Higher accuracy is shown in the congruent trials than the incongruent or baseline trials, for both delay conditions. Error bars represent difference-adjusted, 95% within-subjects confidence interval.

Test reaction times. Figure 4 shows the mean reaction times during the test phase, for trials in which the correct response only was performed. The graph shows that reaction times were similar for congruent and incongruent trials and were slower overall for baseline trials. A 3 (congruency condition: congruent, incongruent, baseline) × 2 (delay condition: short, long) repeated-measures ANOVA revealed a significant main effect of congruency, $F(2, 70) = 6.60$, $p = .002$, $n_g^2 = .01$. However, the Bayes Factor was inconclusive, $BF_{10} = 1.73$. Therefore, these results should be treated with caution. Subsequent simple-effect comparisons showed that participants responded significantly more slowly on baseline trials than congruent trials, $t(71) = 3.88$, $p < .001$, $d_z = 0.71$, $BF_{10} = 95.00$ and incongruent trials, $t(71) = 2.82$, $p = .006$, $d_z = 0.44$, $BF_{10} = 4.94$. No significant difference in reaction times was observed between congruent and incongruent trials, $t(71) = 0.48$, $p = .63$, $d_z = 0.08$, $BF_{10} = 0.14$.

There was a significant main effect of delay, with participants performing more quickly after a long delay than a short delay, $F(1, 35) = 32.07$, $p < .001$, $n_g^2 = .04$. The Bayes Factor analysis provided extreme evidence for this RT difference, $BF_{10} > 100$. However, there was no congruency by delay interaction, $F(2, 70) = 0.50$, $p = .61$, $n_g^2 = .0006$, $BF_{10} = 0.12$.

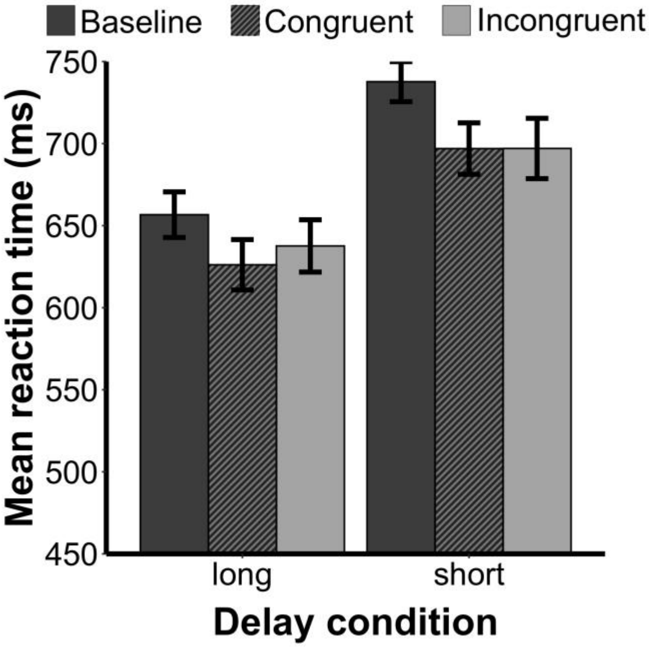


Figure 4: Bar graph of mean reaction times (ms) for congruency and delay conditions in Experiment 1 (n=36). Faster responses are shown in the long delay condition than the short delay condition, for all congruency conditions. Error bars represent difference-adjusted, 95% within-subjects confidence interval.

Discussion

Experiment 1 aimed to explore whether our instrumental task with a congruent verses incongruent trial manipulation could get a congruency effect, with the use of a baseline condition for further comparison. Additionally, Experiment 1 incorporated a O-S delay manipulation during test trials to see if this would have an impact on the main congruency effect. Participants were given instrumental S:R-O training and were then required to execute goal-directed actions to earn jellybeans and Pringles reward points on test. The results suggest that in Experiment 1, goal-directed actions (e.g., R1 to obtain O1) were affected by the presentation of stimuli that signalled that a response could now be made. In particular, accuracy on incongruent test trials (e.g., when the signal that responses for O1 could now be made was the presentation of S2) was lower than accuracy on congruent trials (e.g., when the signal that responses for O1 could now be made was the presentation of S1). There was no difference between performance on incongruent and baseline trials in Experiment 1. This suggests that performance on congruent trials was *enhanced* by the presentation of the compatible stimulus (e.g., S1 following O1), relative to the other two conditions; the *compatible* stimulus on congruent test trials appeared to *help* participants to respond accurately. Therefore, the results of Experiment 1 appear to show evidence of ‘action promoters’ rather than ‘action slips’.

While Experiment 1 seems to provide evidence of ‘action promoters’, it is worth considering the differences between the baseline stimulus, S3, and the other stimuli, S1 and S2. This is a reason to question that the current congruency effect is evidence for a stimulus-supported enhancement of performance in the congruent condition (i.e., that the effect of stimuli on congruent trials is faciliatory), rather than action slips in the incongruent condition. Stimulus S3 was never followed by reward, whichever response was executed during training. This may have led to some frustration for the participants; there was no ‘correct’ response to the baseline stimulus S3 during training. Perhaps then, stimulus S3 became an aversive stimulus, and this negative value interfered with performance on baseline test trials (e.g., O1 followed by S3). Alternatively, perhaps the salience of S3 was maintained throughout training, in a way that the salience of S1 and S2 was not. One speculative possibility is that S3 was more salient on test than S1 and S2 because it was not followed by consistent responses in training (e.g., Pearce & Hall, 1980); there was no correct response to S3, so responses would have been chosen at random. This increased salience of S3 may then have distracted the participant and prevented fast/accurate responding on test. Consistent with this idea, presentation of S3 slowed responding (longer RTs were seen to S3 than to S1 and S2) on test. It would appear that it was difficult to execute the chosen response on test in the presence of S3. Of course, there is the possibility that it was the actual colour (yellow) that made this stimulus more salient, given that it was not counterbalanced.

The analysis above points to a more general issue about appropriate control conditions in experiments such as the one presented here. Stimuli S1 and S2 and outcomes O1 and O2 are exactly equated in the current design. It is not, however, possible to create a baseline condition that is equal in its motivational and attentional properties to the two experimental conditions. A novel S3 on test might be more salient, and a familiar S3 that was not involved in instrumental training might be aversive. An S3 that was followed by reward regardless of the response made may also become more salient. In line with this, Le Pelley et al. (2016) propose a simple model where salience is increased when associative strength is high. Hence, wherever the observed baseline response accuracy lies relative to the experimental congruent/incongruent conditions, it may have been artificially increased or decreased by factors that are irrelevant to the experimental manipulation. It would probably be possible to produce a baseline condition that generates high accuracy and, therefore, apparent evidence that incongruent trials interfere with responding (e.g., using a low salience colour, varying the training parameters until you produced a low-salience, non-aversive stimulus), but what would that actually show? On reflection, we feel that perhaps the baseline condition implemented here is of limited utility. In the subsequent experiments, we removed the baseline condition, leaving the simpler and more elegant bidirectional control design. The issue of whether there is response facilitation on congruent trials or interference on incongruent trials will, therefore, be left open.

Our results showed that the delay manipulation did not affect the number of action slips observed (short versus long delay between presentations of outcome O and stimulus S). Our expectation was that, in the short delay condition (2-3s between O and S), participants would find it harder to think of the correct response before the stimulus S was presented. This might lead to more stimulus-consistent responses (errors) on incongruent trials. Whilst the congruent-incongruent difference in accuracy was larger for the short delay condition (around 5%) than the long delay condition (around 2%), the analysis did not show a significant congruency by delay interaction. One explanation for the absence of any effect of delay on action slips is simply that the delay manipulation was not strong enough in this experiment. Perhaps no delay at all is needed to see an increase in action slips.

A second, more complex, possible reason why we saw no effect of delay on action slips might have been that there were two processes in operation, one facilitated by the short delay and one by the long delay, which cancelled one another out. For example, it may be that the short delay condition allows easy maintenance of the goal in working memory (Schooler et al., 2011), but the long delay condition allows greater time for the participants to prepare their response. That is, the short delay (2-3s) may not always have been enough time for participants to prepare the appropriate response R for that trial (given the presented outcome O). This would then lead to an increase in action slips on short incongruent trials. Some support for this hypothesis comes from the observation that overall responding was faster on long delay trials than short delay trials, as shown in Figure 4. Hence, the failure to see a difference in the size of the action slip effect between the long and short delay conditions may have been due to better response preparation in the long delay condition and lower mind-wandering in the short delay condition.

The main finding from Experiment 1 was that performance on incongruent trials produced poorer performance than on congruent trials, which provides prima facie evidence for action slips. These findings are similar to those in automatic processing tasks, such as the Stroop task (Stroop, 1935), but with the incorrect response trained during the instrumental task. It is possible that the salient yellow stimulus in the baseline condition, which we decided was unlikely to be helpful in identifying the source of the congruent-incongruent difference in accuracy, may have nevertheless influenced participants' performance in the two crucial conditions. In a subsequent experiment, we sought to replicate the congruency effect, but without a baseline condition.

Experiment 2 – O-S delay online

In Experiment 2, we aimed to see if the congruency effect could also occur when the experiment is run online in a less controlled environment. We also aimed to further explore the delay manipulation. We used the same task as Experiment 1, but this time applied the online method (Method 2).

Method

Participants. Forty-one undergraduate psychology students (35 females, 6 males) from the University of Plymouth participated in exchange for course credit. The participants were aged between 18 and 52 years (*M* = 25.27 years, *SD* = 7.34 years).

Design and Procedure. The experiment followed a 2 (congruency condition: congruent, incongruent) by 2 (delay condition: short delay, long delay) repeated-measures design. The design and procedure were the same as Experiment 1, except in the following respects. We omitted the baseline condition, so only green and blue background screen colours were used, as shown in Table 1 illustrating Experiment 2's design. Participants completed four blocks of eight test trials. These eight trials within each block consisted of presentations of each congruent trial (O1-S1 and O2-S2) and each incongruent trial (O1-S2 and O2-S1) presented both with and without the delay task.

Results

Exclusions. The exclusion criteria from Experiment 1 were applied. Twelve participants were removed for achieving less than 80% accuracy on the final test, leaving 29 participants (25 females, 4 males).

Training. Figure 5 shows, as in Experiments 1, the participants learnt the contingencies quickly, with a long period during which participants were asymptote.

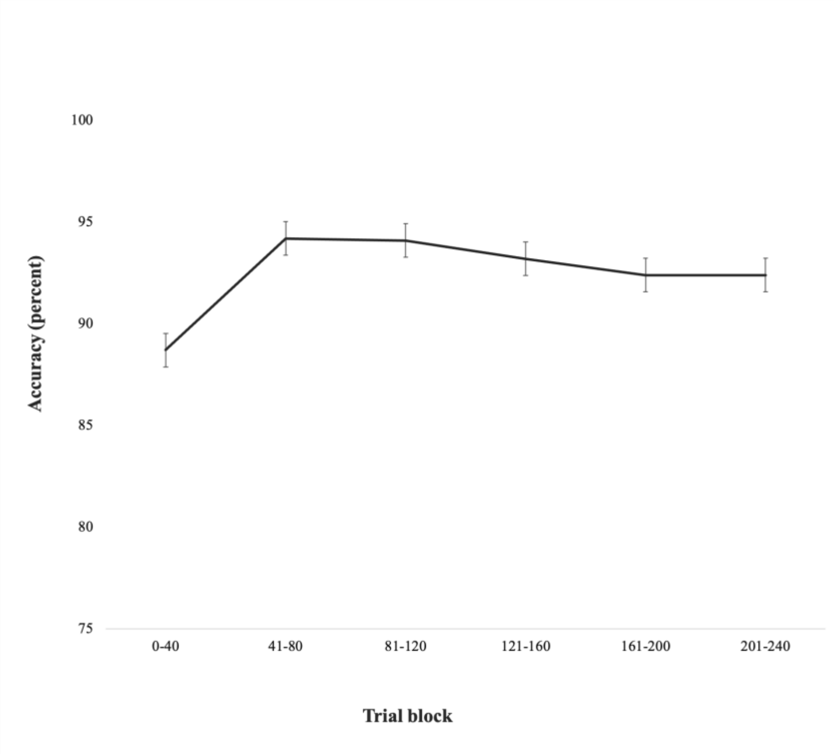


Figure 5: Line graph of mean accuracy (%) of training trials for S1 and S2 in Experiment 2. Over 90% correct responses are shown after the initials 40 trials. Error bars represent the standard error of the mean for each training trial interval.

Test response accuracy. Figure 6 shows the mean percentage of correct responses during the test, separated by congruency and delay conditions. The results shown in Figure 6 suggest that the participants were more accurate on congruent trials than incongruent trials, and that this pattern was not affected by the delay manipulation. A 2 (congruency condition: congruent vs. incongruent) by 2 (delay condition: short delay vs. long delay) repeated-measures ANOVA confirmed this conclusion. Most importantly, there was a significant main effect of congruency on response accuracy, $F(1, 28) = 7.71$, $p = .010$, $d_z = 0.52$, $BF_{10} = 7.74$, with higher accuracy during congruent trials than during incongruent trials. There was no significant effect of delay, $F(1, 28) = 0.38$, $p = .541$, $BF_{10} = 0.23$, or a congruency × delay interaction, $F(1, 28) = 0.38$, $p = .541$. A Bayes Factor analysis confirmed that there is moderate evidence that accuracy did not differ between long and short delay trials, $BF_{10} = 0.31$.

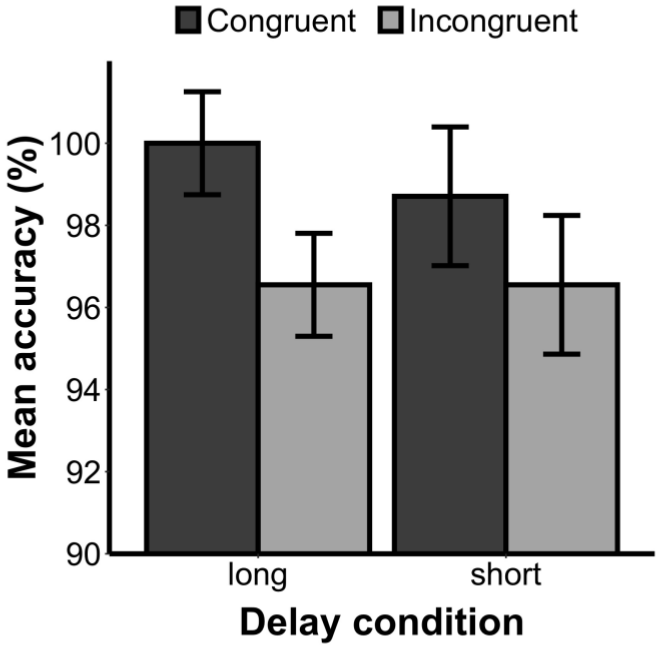


Figure 6: Bar graph of mean accuracy (%) for congruency and delay conditions in Experiment 2 (n=29). Higher accuracy is shown in the congruent trials than the incongruent trials, for both delay conditions. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Test reaction times. As in Experiments 1, a comparable analysis was conducted on the RTs for correct trials only at test. Figure 7 shows that reaction times were slower during incongruent trials than congruent trials. This was supported by ANOVA and Bayes Factor analyses, which showed a significant main effect of congruency on RT, $F(1, 28) = 8.05$, $p = .008$, $d_z = 0.53$, $BF_{10} = 5.37$. There was also a main effect of delay condition on RT, *F*(1, 28) = 25.42, $p < .001$, $d_z = 0.94$, and the Bayes Factor provided extreme support for this, $BF_{10} > 100$, with participants responding more quickly on long delay trials than short delay trials. No significant congruency × delay interaction, *F*(1, 28) = 0.42, $p = .524$, $BF_{10} = 0.30$, was observed.

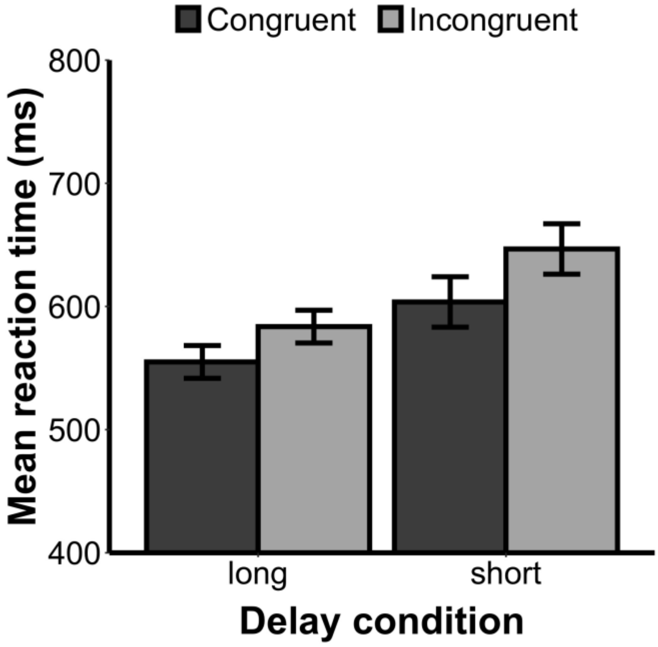


Figure 7: Bar graph of mean reaction times (ms) for congruency and delay conditions in Experiment 2 (n=29). Faster responses are shown in the long delay condition than the short delay condition, for both congruency conditions. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Discussion

Experiment 2 was a replication of Experiment 1 but using our online method (Method 2) to see if the congruency effect could be replicated in a less controlled setting. Additionally, the baseline condition was removed as it appeared to be of limited utility. In line with Experiment 1, we found a highly significant congruency effect on accuracy in Experiment 2. These results suggest that even when we apply a less controlled online procedure, we find evidence of action slips in our experiments. Furthermore, the RT pattern suggests that the accuracy data were not due to a speed accuracy trade off. Rather, the data suggests that the incongruent trials were harder, with lower accuracy and higher RTs. This finding has methodological implications, as it shows that the experiment can be run remotely, which allows greater scope for reaching a wider target population. Running online experiments can also be particularly beneficial in times where it is not possible to conduct face-to-face research, such as during the pandemic.

Whilst Experiment 1 found a significant main effect of delay on accuracy, we found no significant main effect of delay in Experiment 2. Furthermore, there was no delay x congruency interaction in Experiment 1 or Experiment 2. If the delay manipulation was influencing the congruency effect, we would expect to see a congruency x delay interaction, which was not the case for either Experiments. In Experiment 1, the action slips were, if anything, larger in the short delay condition. However, in Experiment 2, it was the reverse. As previously mentioned, Experiment 1 was conducted in a lab rather than online, with a joystick rather than a mouse/trackpad. These factors could have contributed to the differing delay effects between Experiment 1 and Experiment 2. As previously mentioned, one account of the greater errors on short delay trials in Experiment 1 is that goal-directed control is disrupted when there is less time to prepare responses in advance; goal-directed behaviour – preparing the response that is appropriate to earn the outcome – is time consuming and effortful (Heyes & Dickinson, 1990), in comparison with automatic habitual behaviour (e.g., S-R responding left to the green screen), which is fast and effortless. Therefore, when there was less time to prepare a response (short O-S delay), more action slips are made. However, the absence of the delay effect in Experiment 2 can also be explained using the dual-process theory. In the online procedure, participants may have found it easier to physically prepare their responses, even in short O-S delay trials, as they were using a trackpad or mouse to make their responses, rather than the joystick used in the face-to-face procedure. However, during long O-S delays, there is more opportunity for participants to get distracted from the task, and therefore revert to habitual behaviour when the goal is no longer held in working memory. Previous research has shown that during a high-demand task, mind wandering is associated with less accuracy and increases over time (Thomson et al., 2014). Whilst this could be the case during long O-S delays in both Experiment 1 and Experiment 2, Experiment 2’s online procedure is less controlled, and therefore there are higher risks of external distractions interfering with participants goal-directed behaviour during these long O-S delays.

Experiment 3 – no O-S delay

As neither Experiments 1 or 2 found a significant congruency x delay interaction, but there was a hint from Experiment 1 that accuracy is worse after short delays, we felt it was important to replicate these experiments with an even shorter O-S delay. The aim of this was to see if there was an effect on congruency when participants had no time to prepare their response at all. With this in mind, Experiment 3 shortened the short O-S delay condition to almost no delay (500ms to 1 second) with the expectation that this would increase action slips, in line with the dual process theory. We kept the long delay the same as Experiments 1 and 2 (10-11 seconds). Experiment 3 used the face-to-face procedure (Method 1).

Method

Participants. Forty-four undergraduate psychology students (35 females, 9 males) from the University of Plymouth participated in exchange for course credit. The participants were aged between 18 and 37 years (*M* = 21.95 years, *SD* = 4.47 years).

Design and Procedure. The experiment followed the same 2x2 (congruency x delay) design as Experiment 2, demonstrated in Table 1, and used the same face-to-face procedure as Experiment 2 (Method 1), with the following exceptions. The interval between the presentation of the outcome (O) and the stimulus (S) on ‘short delay’ test trials was decreased to 500ms to 1 second. We also doubled the number of test trials that had previously been used to 64 test trials. This was done to allow a more stable estimate of participant performance.

Results

Exclusions. The exclusion criteria from the previous experiments were applied. Nine participants were removed for achieving less than 80% accuracy in test trials, leaving 35 participants (29 females, 6 males).

Training. Figure 8 shows participants learnt the contingencies quickly, with a long period (although slightly shorter than previous experiments) during which participants were at asymptote.

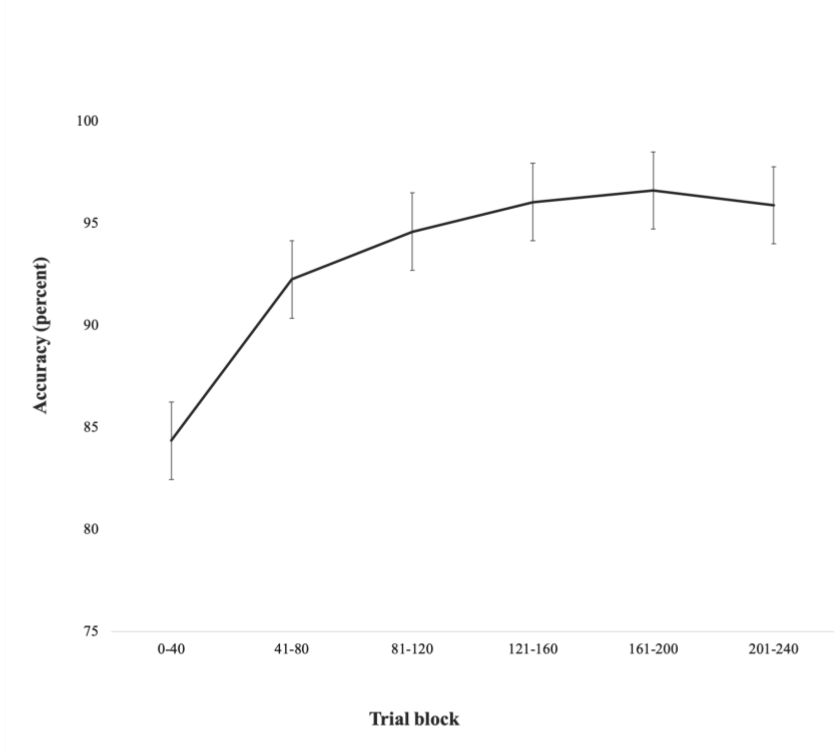


Figure 8: Line graph of mean accuracy (%) of training trials for S1 and S2 in Experiment 3. Over 90% correct responses are shown after the initials 40 trials. Error bars represent the standard error of the mean for each training trial interval.

Test response accuracy. Figure 9 shows the mean percentage of correct responses during the test, separated by congruency and delay conditions. The graph suggests that the participants were more accurate on congruent trials than incongruent trials, and that this pattern was affected by the delay manipulation. A 2 (congruency condition: congruent vs. incongruent) by 2 (delay condition: short delay vs. long delay) repeated-measures ANOVA confirmed this conclusion. There was a significant main effect of congruency on response accuracy, $F(1, 34) = 12.14$, $p = .001$, $d_z = 0.59$, and the Bayes Factor provided extreme support for this, $BF_{10} > 100$, with higher accuracy during congruent trials than during incongruent trials. The main delay effect was also significant, $F(1, 34) = 8.03$, $p = .008$, $d_z = 0.48$, with participants responding more accurately in long delay trials than short delay trials. However, the Bayes Factor for the main delay effect was inconclusive, $BF_{10} = 1.79$, therefore, these results should be treated with caution. There was also a significant congruency × delay interaction, $F(1, 34) = 5.59$, $p = .024$, but again, the Bayes Factor was inconclusive, $BF_{10} = 0.80$, and actually pointed towards support for the null hypothesis, although anecdotal.

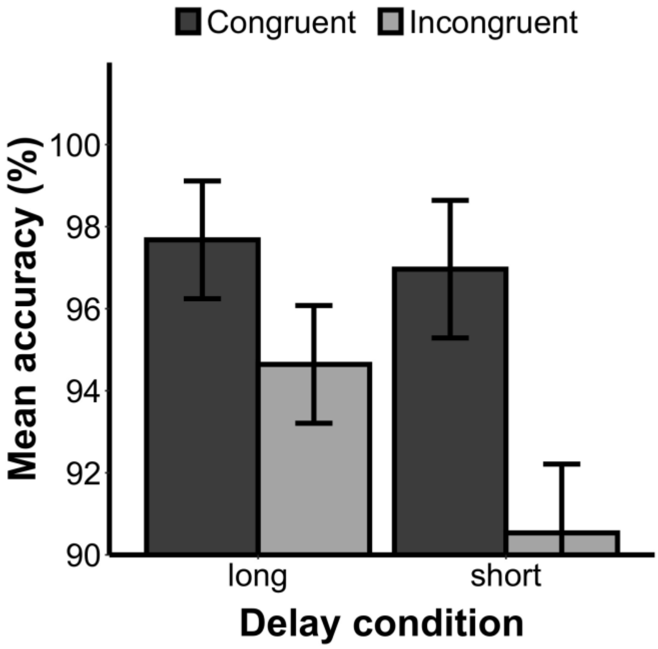


Figure 9: Bar graph of mean accuracy (%) for congruency and delay conditions in Experiment 3 (n=35). Higher accuracy is shown in the congruent trials than the incongruent trails, with a greater congruency effect in the long delay condition. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Test reaction times. Figure 10 shows, as in previous experiments, a comparable analysis was conducted on the RTs for correct trials only at test. There was a significant main effect of congruency condition on RT, $F(1, 34) = 8.66$, $p = .006$, $d_z = 0.50$, $BF_{10} = 5.15$, with participants responding more quickly on congruent trials than incongruent trials, as shown in the graph. There was also a main effect of delay condition on RT, $F(1, 34) = 17.12$, $p < .001$, $d_z = 0.70$, with participants responding more quickly on long delay trials than short delay trials. A Bayes Factor analysis also provided extreme evidence for this effect, $BF_{10} > 100$. No significant congruency × delay interaction on RT was observed, $F(1, 34) = 1.60$, $p = .215$, $BF_{10} = 0.36$.

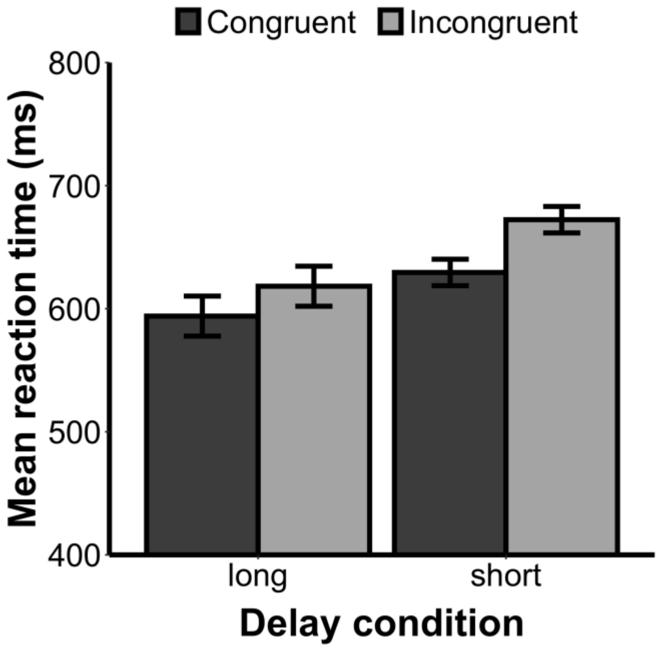


Figure 10: Bar graph of mean reaction times (ms) for congruency and delay conditions in Experiment 3 (n=35). Faster responses are shown in the long delay condition than the short delay condition, for both congruency conditions. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Discussion

Experiment 3 replicated the main procedure of Experiment 2 but applied the more controlled environment of Experiment 1 (in a laboratory, using a joystick) and decreased the ‘short delay’ condition to almost no delay (500-1000 ms). Again, a main congruency effect was found, with participants responding more accurately on congruent trials than incongruent trials. Experiment 3 also replicated the main effect of delay found in Experiment 1 (not seen in Experiment 2). Again, participants responded more quickly and more accurately on long delay trials than short delay trials, suggesting that action slips are more likely to occur when the participant has no time prepare their response. Most importantly, Experiment 3 found a significant congruency by delay interaction on accuracy, which has not been found in previous experiments. This suggests that delay has a significant effect on accuracy, but mostly on the incongruent trials. These results indicate that habitual behaviour is more prominent where there is no delay between the presentation of the outcome and the stimulus.

General discussion

In all three experiments, we appear to see evidence for habitual behaviour. Responding was more accurate when the presented stimulus was congruent with the available outcome (i.e., O1 followed by S1, or O2 followed by S2) than when the stimulus was incongruent with the outcome (i.e., O1 followed by S2, or O2 followed by S1). These results indicate that the stimuli S1 and S2 presented on incongruent trials disrupted goal-directed behaviour (Dickinson & Balleine, 1994); participants' accuracy appeared to be influenced by the stimuli presented, even though they were instructed explicitly to ignore those stimuli.

While we have taken the difference in accuracy between the congruent and incongruent trials as evidence of habitual behaviour, it is important to acknowledge that accuracy rates did not significantly differ between the incongruent and baseline trials in Experiment 1. If accuracy had been significantly poorer in the incongruent condition than the baseline condition (similar to de Wit et al.'s biconditional control trials), it would have provided even stronger evidence of habitual control. That said, it is difficult to interpret the function of the baseline trials, which is why we removed them from subsequent experiments. A direct comparison of congruent and incongruent trials allows for a cleaner experimental design, but we acknowledge that it is unclear at this stage whether the incongruent stimuli interfere with goal-directed action, or whether the congruent stimuli *facilitate* goal-directed action (or both).

In addition to the congruency manipulation, the delay between O and S on test was manipulated in all three of the experiments in this chapter. This manipulation did not affect the number of errors observed in Experiment 2 but did have a significant effect on the number of errors observed in Experiments 1 and 3. The main difference here was that Experiments 1 and 3 were carried out in a controlled lab environment, where participants used a joystick to make a response, whereas participants completed Experiment 2 remotely online using either a mouse or trackpad. In Experiments 1 and 3, participants made more errors on short delay trials than long delay trials. These results suggest that participants' opportunity to prepare their response in advance in long O-S delay improved their accuracy. However, in trials with short O-S delays, participants did not appear to have enough time to prepare the correct response, and consequently may have reverted to S-R responses (Hardwick et al., 2019). The online nature of Experiment 2 may have meant participants had additional distractions during the long O-S delay which disrupted their goal-directed behaviour, which could explain why the results differed in terms of the delay manipulation.

In Experiment 3, the short delay was decreased further to almost no delay. It is here that we found a congruency x delay interaction. An interpretation of this is that by decreasing the short delay to almost no delay, we made it even harder for participants to prepare their responses in advance on these trials, reverting them to habitual behaviour. It is important to note that the Bayes Factor was inconclusive, so these results should be treated with caution.

In addition to these findings, in all three experiments, shortening the delay increased participants' reaction times. It is worth bearing in mind that these reaction times are for accurate responses only, so in order to respond accurately, it took participants more time to respond on short delay trials than long delay trials. A dual process explanation of these results is that on long delay trials, participants were able to hold the correct response in mind after presentation of the outcome and made their response quickly after presentation of the stimulus. In short delay trials, on the other hand, the sudden presentation of the stimulus did not allow participants enough time to prepare the correct response in advance of stimulus presentation, causing a longer response time after the stimulus was presented to make the correct response.

To summarise, Chapter 2 has presented clear congruency effects in our experiment, providing evidence of action slips. The evidence for the effect of O-S delay is mixed at best. Decreasing O-S delay had a tendency to increase the congruency effect. Although evidence for this effect was not strong. We wanted to explore whether we could increase action slips through other experimental manipulations. In addition to the congruency manipulation, Chapter 3 focuses on introducing working memory manipulations to our experiments to explore this potential effect on action slips.

Chapter 3 – **Load and time pressure**

As discussed in Chapter 1, goal-directed processes are thought to take up working memory capacity, whilst automatic processes do not. Therefore, when there is competition between goal-directed actions and habitual responses, increasing cognitive load is likely to increase automatic behaviour (de Wit et al., 2013; Le Pelley et al., 2005). One way to test for automatic behaviour in experiments is, therefore, to use a secondary cognitive task that requires participants' working memory. Cognitive load tasks, for example, are recognised as an effective way to exhaust working memory capacity and, therefore, increase the potential to detect automatic behaviour (De Houwer & Beckers, 2003; Wills, et al., 2011). An example of this is seen in De Houwer and Beckers' forward cue competition experiment. In a computer game task, participants first learnt that that firing a particular weapon (A) was followed by the destruction of a tank. Later, weapon A was fired simultaneously with a new weapon (B), which also led to the destruction of the tank. Participants were then asked to rate the extent to which weapon B caused the destruction of the tank. The causal status of B was uncertain, as it always occurred in the presence of A, which was known to destroy the tank independently of B. However, the authors found that participants gave higher ratings for the extent to which B was considered to cause the tank's destruction (a reduced blocking effect) when they were under additional working memory load, compared to no load. Therefore, applying additional working memory load interfered with participants' ability to reason that the causal status of B was uncertain. Instead, they responded based on the contiguity between weapon B and the destruction of the tank. The authors argued that the load task prevented the operation of the reasoning processes that would otherwise have reduced the ratings of the efficacy of the tank (cue B).

To explore whether reducing participants' working memory capacity leads to more action slips in our experiments, we embedded an additional task during the O-S delay in half of the trials in Experiments 4-6. This additional task was designed to create a distraction or increase cognitive load on half of the trials, reducing participants' working memory capacity and, therefore, increasing action slips in these trials.

Applying time pressure in experimental tasks is another manipulation that is thought to increase action slips (Buabang et al., 2023). According to dual-process theories, goal-directed processes operate relatively slowly (Moors & De Houwer, 2006). Therefore, stimulus-driven responses may become apparent when examining responses that are generated under strict time constraints. One example of existing research demonstrating this effect of time pressure is Hardwick et al. (2019). They conducted an experiment where participants were trained to make specific responses when presented with specific visual stimuli. Participants received either moderate or extensive training. In a subsequent reversal phase, two of the trained stimulus-response associations were reversed. The test phase then involved four different levels of time restraints. The results showed that more action slips were made by participants in the extensive training group when they were under greater time pressure. Experiment 7 explores whether time pressure has an effect on action slips in our experiments, using a between-subject design. Half of the participants are presented with the original procedure from Chapter 1, where they are told that their points depend on the speed of their response and that they lose points if they respond too slowly. The remaining participants are told to focus on responding as accurately as possible, with no mention of speed.

In line with Chapter 2, we expected to find a congruency effect in each of Chapter 3’s experiments, with more action slips on incongruent trials. The idea behind applying a cognitive load task to our experiments during the O-S delay is that it requires more of participants’ working memory capacity, reducing their ability to keep the goal in mind across the O-S delay. Therefore, we expected any congruency effect to be exaggerated by increasing distraction or cognitive load tasks in Experiments 4-6. In Experiment 7, we expected to see a larger congruency effect for participants in the ‘time pressure’ condition used in previous experiments, compared with those in the ‘no time pressure’ condition. The reasoning behind this is that removing the time pressure is thought to free up working memory capacity. Therefore, it will be easier for participants to keep the goal in mind.

Experiment 4 – distraction task

Experiment 4 applied the ‘in-person’ experimental design detailed in Chapter 1. A distraction task was embedded between O and S presentation on half of the test trials. We used a moderate O-S delay of 3-4s for all test trials, to allow time to implement the distraction task between presentation of the outcome (O) and stimulus (S) on test trials. On distraction trials, a photograph of a person was presented, and participants were asked to rate verbally how much they liked each photograph. This was meant as a mild distraction to trigger mind wandering (Schooler et al., 2011), as a larger distraction may also affect participants' performance on the congruent trials, reducing all responding to chance. Our hypothesis was that the distraction task might disrupt the maintenance of the goal in working memory and therefore render the participants more susceptible to influence of the incompatible stimulus on incongruent trials.

Overall, the experiment utilised the same training procedure as Experiments 1-3 and manipulated O-S congruence and distraction on test orthogonally. We predicted that performance would be poorer on incongruent trials than congruent trials, and that this difference would be more marked when a distraction was present.

Method

Participants. Sixty-six undergraduate psychology students (57 females, 9 males) from the University of Plymouth participated in exchange for course credit. The participants were aged between 18 and 34 years (*M* = 20.41 years, *SD* = 2.87 years).

Design and Stimuli. The experiment followed a 2 (congruency condition: congruent, incongruent) by 2 (distraction condition: distraction, no distraction) repeated-measures design. Thirty photographs of famous people (faces only) were used for the distraction manipulation, as shown in Appendix 1.

Procedure. The procedure was the same as Experiment 3, with the following exceptions. On test, there was a 3-4s delay between the presentation of the outcome and the stimulus on every test trial. On distraction trials, if the joystick was kept central throughout the delay, a photograph of a person was presented (also throughout the delay). Each photograph was individually presented with a white background and placed in the centre of the screen. Participants were asked to verbally rate how much they liked each photograph, using a scale between 1 ("Don't like at all") and 7 ("Like very much"). On 'no distraction' trials, the screen remained blank for an equivalent duration.

Participants completed four blocks of eight test trials, as shown in Table 1. These eight trials consisted of presentations of each type of congruent trial (O1-S1 and O2-S2) and each type of incongruent trial (O1-S2 and O2-S1), each with and without the distraction task.

Results

Exclusions. The exclusion criteria from previous experiments were applied. Nineteen participants were removed for achieving less than 80% accuracy in test trials, leaving 47 participants (39 females, 8 males).

Training. Figure 11 shows, as in previous experiments, participants learnt the contingencies quickly, and mean accuracy during training remained above 95% after the first 40 trials.

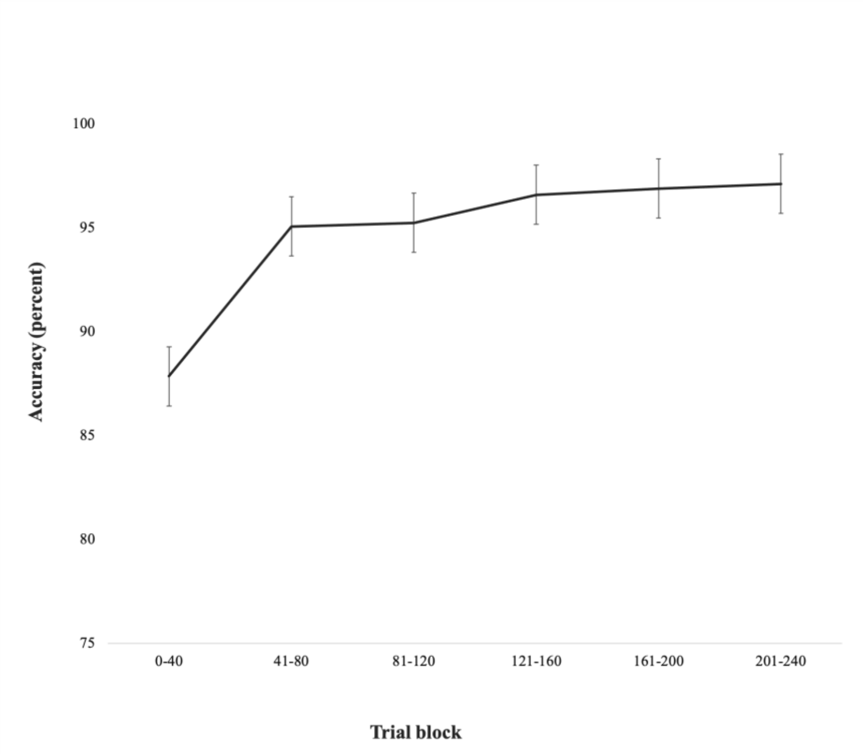


Figure 11: Line graph of mean accuracy (%) of training trials for S1 and S2 in Experiment 4. Over 90% correct responses are shown after the initials 40 trials. Error bars represent the standard error of the mean for each training trial interval.

Test response accuracy. Figure 12 shows the mean percentage of correct responses on test, separated by congruency and distraction conditions. The graph suggests that the participants were more accurate on congruent trials than incongruent trials, and that this pattern was not affected by the distraction manipulation. A 2 (congruency condition: congruent vs. incongruent) by 2 (distraction condition: distraction vs. no distraction) repeated-measures ANOVA confirmed this conclusion. Most importantly, there was a significant main effect of congruency on response accuracy, $F(1, 46) = 10.49$, $p = .002$, $d_z = 0.47$, $n_g^2 = .06$, with higher accuracy on congruent trials than on incongruent trials. A Bayes Factor analysis provided extreme support for this congruency effect, $BF_{10} > 100$. There was no effect of distraction on accuracy, $F(1, 46) = 0.61$, $p = .44$, $n_g^2 = .002$, $BF_{10} = 0.18$, or congruency × distraction interaction, $F(1, 46) = 0.09$, $p = .76$, $n_g^2 = .0004$, $BF_{10} = 0.22$, with both Bayes Factors providing evidence for the null hypothesis.

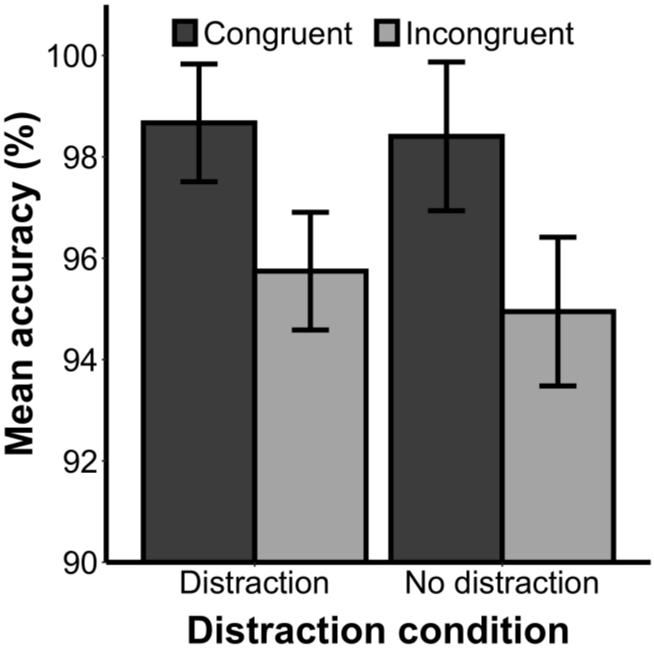


Figure 12: Bar graph of mean accuracy (%) for congruency and distraction conditions in Experiment 4 (n=47). Higher accuracy is shown in congruent trials than the incongruent trials, for both distraction conditions. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Test reaction times. Figure 13 shows, as in previous experiments, a comparable analysis was conducted on the RTs for correct trials only at test. There was a main effect of distraction condition, $F(1, 46) = 18.65$, $p < .001$, $d_z = 0.63$, $n_g^2 = .04$, with participants responding more quickly on no distraction trials than distraction trials. A Bayes Factor analysis provided extreme support for this distraction effect, $BF_{10} > 100$. No effect of congruency condition, $F(1, 46) = 2.92$, $p = .09$, $n_g^2 = .004$, $BF_{10} = 0.48$, nor congruency × distraction interaction, $F(1, 46) = 0.10$, $p = .75$, $n_g^2 = .0001$, $BF_{10} = 0.23$, was observed.

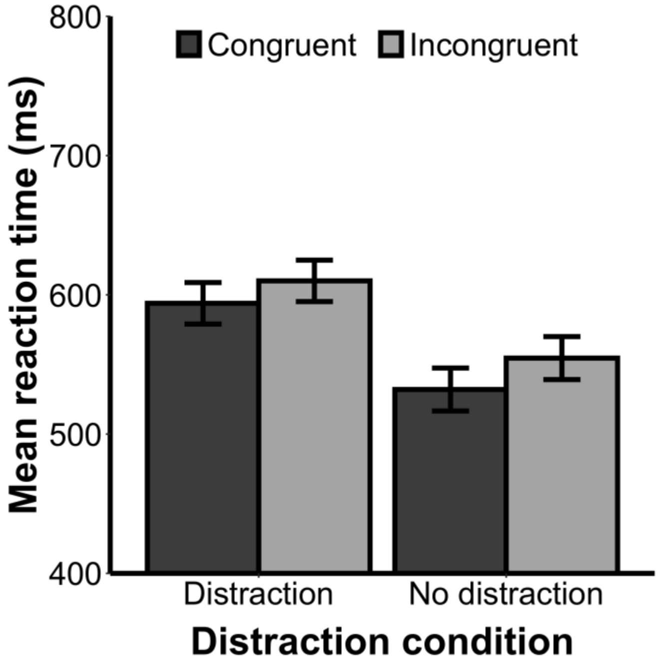


Figure 13: Bar graph of mean reaction times (ms) for congruency and distraction conditions in Experiment 4 (n=47). Faster responses are shown in the no distraction condition than the distraction condition, for both congruency conditions. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Discussion

As in previous experiments, participants made more errors on incongruent trials than congruent trials, supporting the evidence for action slips found in Chapter 2. We predicted a reduction in working memory capacity would increase the number of errors on incongruent trials. However, our distraction manipulation had no impact on the number of action slips observed. The faces task was only a light distraction; therefore, it may have been insufficient to affect participants' accuracy. Having said this, the distraction task slowed participants' responses, suggesting it had some impact on performance in the experiment. A simple explanation for this, based on Luque et al.'s (2020) findings, is when trials became more difficult, participants compensated by taking longer to respond. Thus, accuracy was not affected by the distraction task. It is reassuring that Experiment 4 replicated the action slip effect seen in the experiments from Chapter 2, and that the distraction task had an impact. However, participants seem to have been able to maintain accuracy despite this impact. In order to further explore the effect of working memory on action slips, we implemented a task that imposes a higher cognitive load than the faces task of the current experiment.

Experiment 5 – online load

Wills et al. (2011) used a more challenging load task to reveal perhaps more automatic associative processes in a contingency learning task. Their load task requires participants to remember number sequences during a transfer test where working memory is occupied at the point of responding. They found that this concurrent load task generated a switch from rule-based to a simpler feature-based generalisation in their causal learning task. Experiment 5 aimed to further explore the effect of cognitive load on action slips by embedding Wills et al.'s cognitive load task into half of the test trials between the O and S presentation. As our previous experiments found a significant main effect of congruency on accuracy, we expected to find the same effect in Experiment 5. Whilst Experiment 4 found no effect of distraction on accuracy, we were investigating whether a cognitive load manipulation would have this effect in Experiment 5, as it was designed to impose higher cognitive load on participants than Experiment 4.

Method

Participants. Sixty-one participants (27 females, 34 males) were recruited via Prolific and took part in the experiment. The participants were aged between 20 and 70 years (*M* = 38.30 years, *SD* = 12.62 years).

Design, Stimuli and Procedure. Experiment 5 used the 'online' (Method 2) experimental design described in Chapter 1, with the same 2 (congruency condition: congruent, incongruent) by 2 (load condition: load, no load) repeated-measures design used in Experiment 4. To allow time to implement the cognitive load task, all test trials had a delay of 10-11 seconds between presentation of O and S. Half of the test trials involved a working memory task for participants to complete during the delay, creating a 'cognitive load' manipulation. The imposed 'load' was a digit-span test. Six random numbers were played sequentially, through headphones. Participants were instructed to remember these numbers in sequence. After they heard the 6 numbers, they had to keep the cursor centred whilst they waited for the screen to change colour to make their response. After making a response, participants were presented with a number on screen and were asked to respond by selecting the number they heard next in the sequence via the keyboard. Hence, if they heard "3, 4, 6, 1, 5, 2" during the O-S delay, and were presented with "6" after they had made their response, they were required to select "1" using the keyboard. As the cognitive load manipulation was reliant on participants hearing the numbers played through the headphones, an audio check was added to the beginning of the experiment. During the audio check, participants were unable to continue to the next screen until they correctly typed the two words played to them through the headphones, using the keyboard.

Results

Exclusions. The exclusion criteria from previous experiments were applied. Eight participants were removed for achieving less than 80% accuracy on the final test, leaving 53 participants (25 females, 28 males).

Training. As in previous experiments, Figure 14 shows participants learnt the contingencies quickly, reaching around 90% accuracy by training trial 40. However, there does tend to be a slower pattern for online experiments, compared to face-to-face experiments.

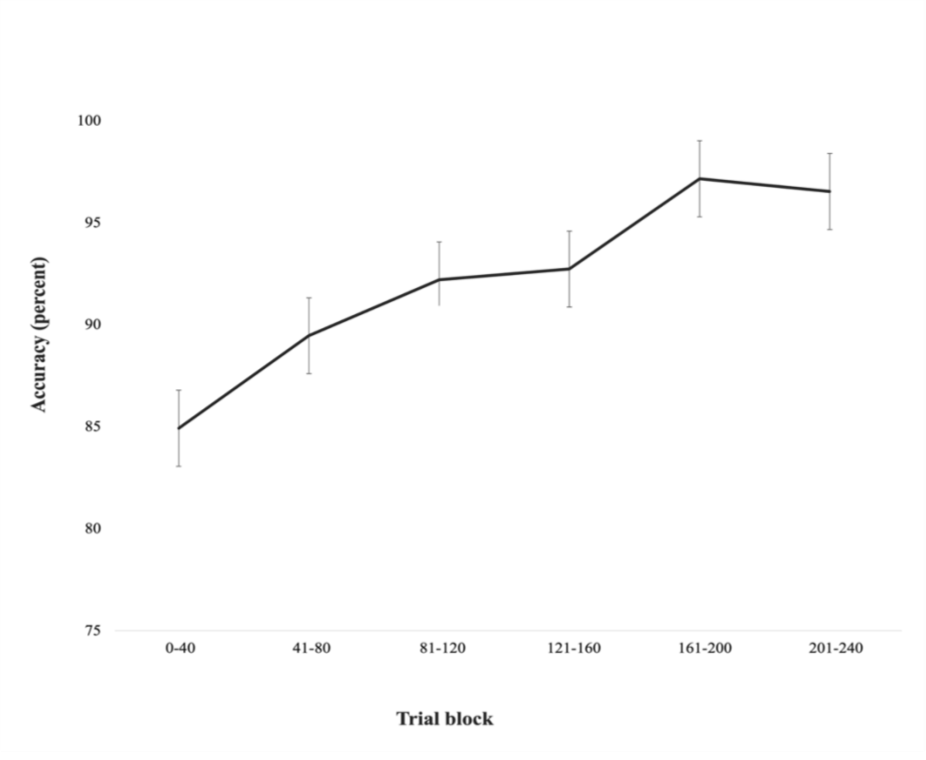


Figure 14: Line graph of mean accuracy (%) of training trials for S1 and S2 in Experiment 5. Over 90% correct responses are shown after the initials 40 trials. Error bars represent the standard error of the mean for each training trial interval.

Test response accuracy. Figure 15 shows the mean percentage of correct responses during the test, separated by congruency and load conditions. The graph suggests that the participants were more accurate on congruent trials than incongruent trials and made more errors on load trials than no load trials. A 2 (congruency condition: congruent vs. incongruent) by 2 (load condition: load vs. no load) repeated-measures ANOVA confirmed that there was a significant main effect of congruency on response accuracy, $F(1, 52) = 10.36$, $p = .002$, $d_z = 0.44$, with higher accuracy during congruent trials than during incongruent trials. A Bayes Factor analysis provided extreme support for this congruency effect, $BF_{10} > 100$. Whilst there was lower accuracy on load trials than no load trials, this effect was not significant and the Bayes Factor was inconclusive, $F(1, 52) = 3.39$, $p = .071$, $d_z = 0.25$, $BF_{10} = 0.61$. There was also no congruency × load interaction, $F(1, 52) = 1.48$, $p = .229$, $BF_{10} = 0.24$.

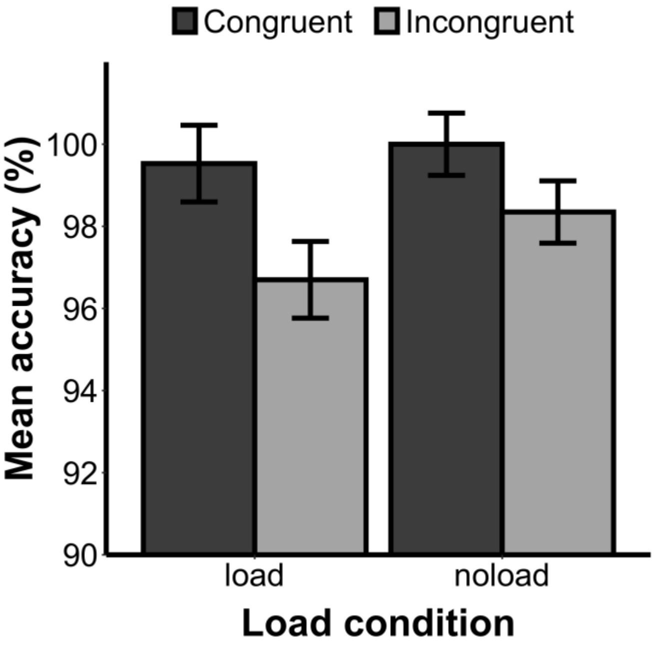


Figure 15: Bar graph of mean accuracy (%) for congruency and load conditions in Experiment 5 (n=53). Higher accuracy is shown in congruent trials than incongruent trials, and more errors in the load condition than the no load condition. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Test reaction times. Figure 16 shows, as in previous experiments, a comparable analysis was conducted on the RTs for correct trials only at test. There was a large main effect of load condition, $F(1, 52) = 69.94$, $p < .001$, $d_z = 1.15$, $BF_{10} > 100$, with participants responding more quickly on no load trials than load trials. No significant effect of congruency condition, $F(1, 52) = 1.62$, $p = .208$,

BF_{10} = 0.21, was observed, and the congruency × load interaction, $F(1, 52) = 3.41$, $p = .071$, BF_{10} = 0.66, was inconclusive.

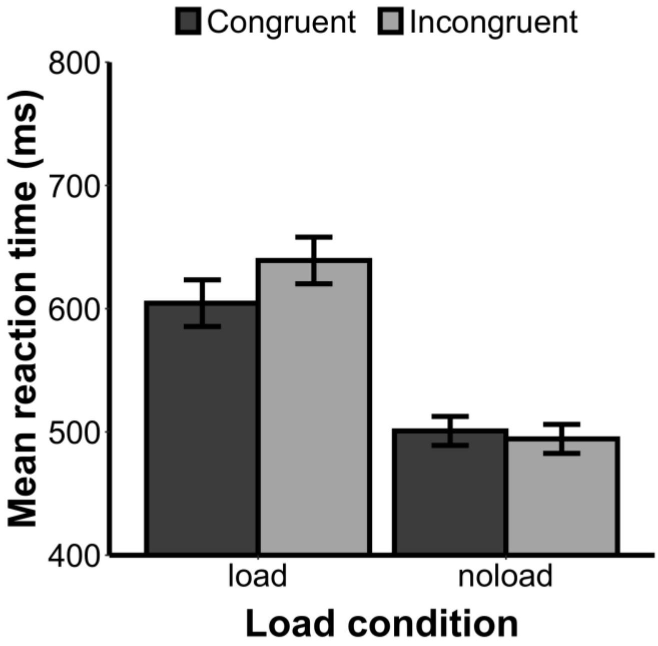


Figure 16: Bar graph of mean reaction times (ms) for congruency and load conditions in Experiment 5 (n=53). Faster responses are shown in the no load condition than the load condition, for both congruency conditions. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Discussion

Consistent with previous experiments, there was a significant congruency effect on accuracy. Whilst accuracy appeared to be lower on incongruent trials in the load condition than in the no load condition, there was no main effect of load or interaction between the two factors. This suggests that our O-R procedure does not require a great deal of working memory capacity. However, there was a large effect of load on RTs for correct responses. As in Experiment 4, a simple explanation for this is that participants prioritise accuracy over response time and compensate on the more difficult load trials by responding more slowly to ensure high accuracy. There was no significant difference in RTs between congruent and incongruent trials, suggesting that main effect of congruency on accuracy was not due to a speed accuracy trade off. Most importantly, there was no congruency x load interaction. Therefore, there was no evidence for the effect of cognitive load on habitual behaviour. One explanation for this is that environmental conditions outside of experimenter control could have affected participants' working memory throughout the experiment, regardless of trial condition. As participants completed Experiment 5 remotely online, we were unable to monitor the conditions of their surrounding environment. Although participants were instructed to complete the experiment in a quiet setting, this experiment was conducted during a 'lockdown' phase of the pandemic. Therefore, potential distractions in participants' home environments may have been unavoidable and may have reduced working memory capacity even on 'no load' trials. It felt important to replicate this experiment using our in-person method (Method 1) to see if we would observe a load effect on action slips when there is less risk of external distractions.

Experiment 6 – lab load mouse

Experiment 6 aimed to explore the same cognitive load manipulation of Experiment 5 under more controlled conditions. The same 'online' programme from Experiment 5 was used. However, Experiment 6 ran the programme in the lab with an experimenter present. The only adjustment to the programme itself was the number of test trials, which were doubled to allow more opportunity to observe action slips, as accuracy in all previous experiments had been consistently high. Responses were made using a mouse rather than a joystick, as a joystick was not compatible with this web-based version of the experiment.

Method

Participants. Eighty undergraduate psychology students (55 females, 22 males, 3 other) from the University of Plymouth participated in exchange for course credit. The participants were aged between 18 and 41 years (*M* = 20.28 years, *SD* = 3.56 years).

Design, Stimuli and Procedure. The experiment followed the same 2 (congruency condition: congruent, incongruent) by 2 (load condition: load, no load) repeated-measures design as Experiment 6. This experiment, however, was run in the laboratory for a more controlled environment (Method 1). Two additional changes were made: a mouse was used instead of a joystick and the number of test trials was doubled to 64.

Results

Exclusions. The exclusion criteria from previous experiments were applied. Fourteen participants were removed for achieving less than 80% accuracy in test trials, leaving 66 participants (47 females, 16 males, 3 other).

Training. Figure 17 shows, as in previous experiments, participants learnt the contingencies quickly, with a long period during which participants were asymptote.

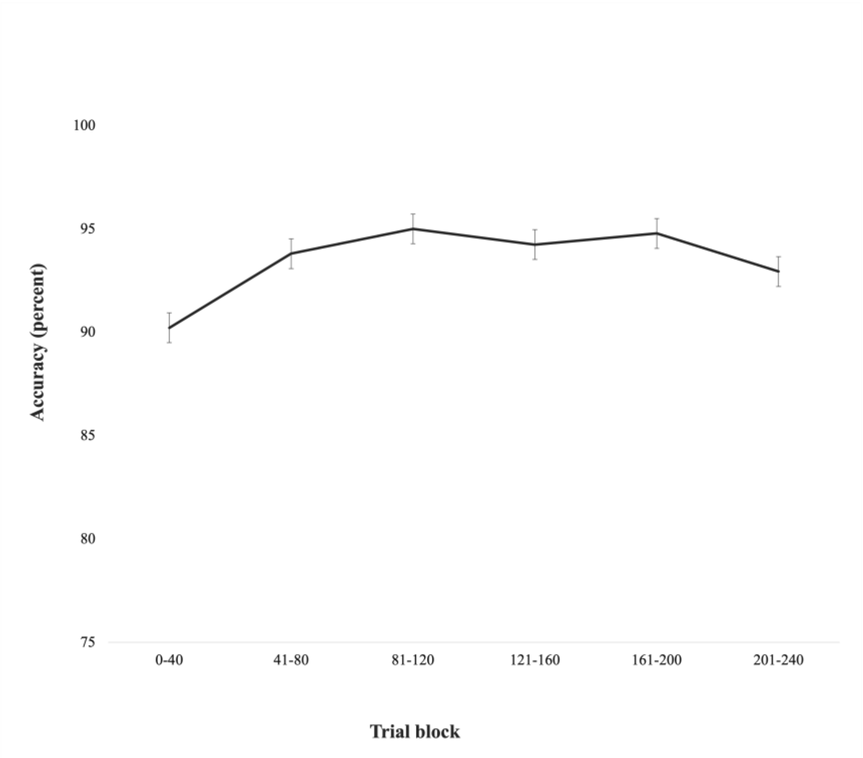


Figure 17: Line graph of mean accuracy (%) of training trials for S1 and S2 in Experiment 6. Over 90% correct responses are shown after the initials 40 trials. Error bars represent the standard error of the mean for each training trial interval.

Test response accuracy. Figure 18 shows the mean percentage of correct responses during the test, separated by congruency and load conditions. The graph suggests that the participants were more accurate on congruent trials than incongruent trials and made more errors on load trials than no load trials. A 2 (congruency condition: congruent vs. incongruent) by 2 (load condition: load vs. no load) repeated-measures ANOVA confirmed that there was a significant main effect of congruency on response accuracy, $F(1, 64) = 16.89$, $p < .001$, $d_z = 0.51$, with higher accuracy during congruent trials than during incongruent trials. A Bayes Factor analysis provided extreme support for this congruency effect, $BF_{10} > 100$. There was no significant main load effect, $F(1, 64) = 3.23$, $p = .077$, $BF_{10} = 0.22$.

There was a significant congruency × load interaction, *F*(1, 64) = 4.46, *p* = .039, where load appeared to have an effect on incongruent trials but not congruent trials. However, the Bayes Factor was inconclusive, BF_{10} = 0.53. Therefore, this result should be treated with caution.

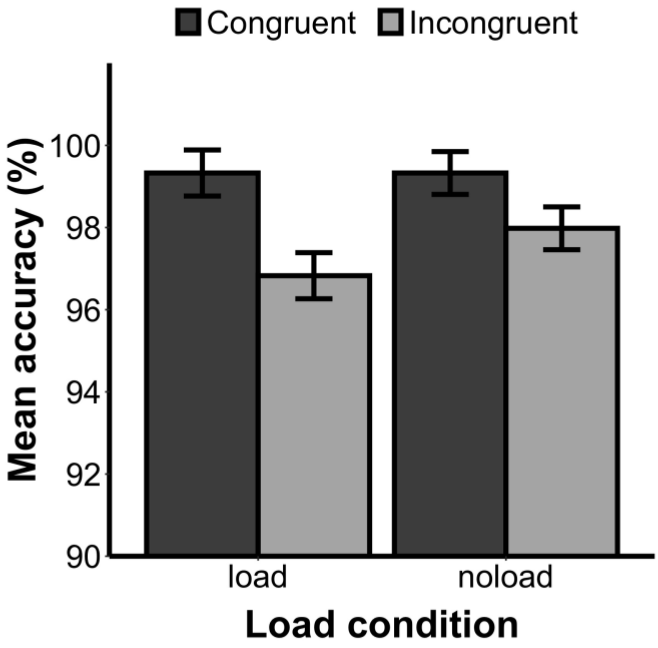


Figure 18: Bar graph of mean accuracy (%) for congruency and load conditions in Experiment 6 (n=66). Higher accuracy is shown in congruent trials than incongruent trials, and more errors in the load condition than the no load condition. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Test reaction times. Figure 19 shows, as in previous experiments, a comparable analysis was conducted on the RTs for correct trials only at test. The graph suggests that RT were similar for all trial conditions, where all responses appear to be relatively fast (between 420 and 450 ms). No main effect of load condition, *F*(1, 64) = 0.49, *p* = .485, BF_{10} = 0.21, or congruency condition, *F*(1, 64) = 0.86, *p* = .358, BF_{10} = 0.17, was observed, and the congruency × load interaction, *F*(1, 64) = 3.53, *p* = .065, BF_{10} = 0.45, was inconclusive.

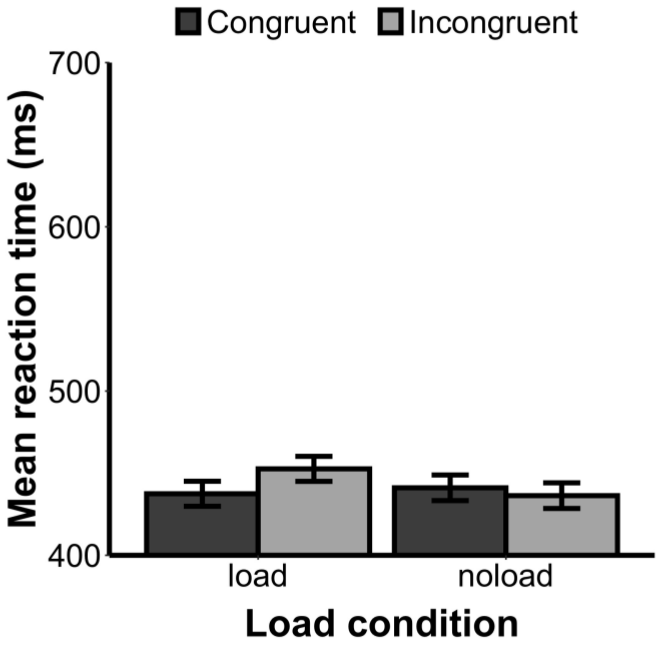


Figure 19: Bar graph of mean reaction times (ms) for congruency and load conditions in Experiment 6 (n=66). Similar response times are shown in all conditions. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Discussion

Experiment 6 replicated the main effect of congruency on accuracy found in our previous experiments. Again, there was no main effect of congruency on RT, indicating that our congruency effect on accuracy was not due to a speed-accuracy trade-off. We used the same procedure as Experiment 5, with the only exceptions being that the current experiment was conducted in-person (Method 1), whereas Experiment 5 was conducted online (Method 2), and the current experiment used double the number of test trials than Experiment 5. When investigating effects of load, results differed across the two experiments. In Experiment 5, there was a main effect of load on RT, and no load effect on accuracy. Conversely, Experiment 6 found no main effect of load on RTs (participants responded quickly on all trials, regardless of load or congruency condition). When exploring effects on accuracy, there was a significant load by congruency interaction in Experiment 6, where the difference between accuracy during congruent and incongruent trials was greater under the load condition. These results indicate that the congruency effect was more pronounced when participants had reduced working memory capacity, providing us with some evidence for automatic processes in our experiment, although the Bayesian evidence was anecdotal.

Although Experiment 5 and 6 applied the same cognitive load manipulation (Wills et al.'s digit span task), the effect of this manipulation varied between experiments. There was a significant congruency × load interaction in Experiment 6, but no interaction in Experiment 5. As previously mentioned, the key difference between these experiments was that Experiment 5 applied the online procedure, where responses were made with a trackpad/mouse (Method 2) and Experiment 6 applied the in-person procedure (Method 1), but responses were made with a mouse (rather than a joystick). It is possible that participants found it easier to prepare their response in advance using a trackpad, which could explain why we saw a greater effect of load on action slips in Experiment 6, where only a mouse was used. Another possibility is that there were more external distractions when participants completed the experiment online in Experiments 5. Therefore, this could have affected participants' working memory throughout the experiment, regardless of trial condition. In order to explore whether the response method (trackpad verses mouse) or environment (lab verses online) had an effect on the results, we combined the data from Experiments 5 and 6, giving us 118 participants in total, and ran post-hoc analyses.

Post-hoc analysis of response method. Firstly, we investigated the effects of congruency and load for all participants in these two experiments when adding an additional between-subject factor of response method (mouse verses trackpad). 95 participants used a mouse to make their response, and 23 participants used a trackpad. As shown in Figure 20, this analysis revealed no main effect of response method on accuracy, $F(1, 116) = 1.07$, $p = .303$, $BF = 0.30$. There was also no response method x congruency interaction $F(1, 116) = 0.04$, $p = .843$, BF = 0.19, nor response method x load interaction $F(1, 116) = 0.67$, $p = .413$, $BF = 0.24$.

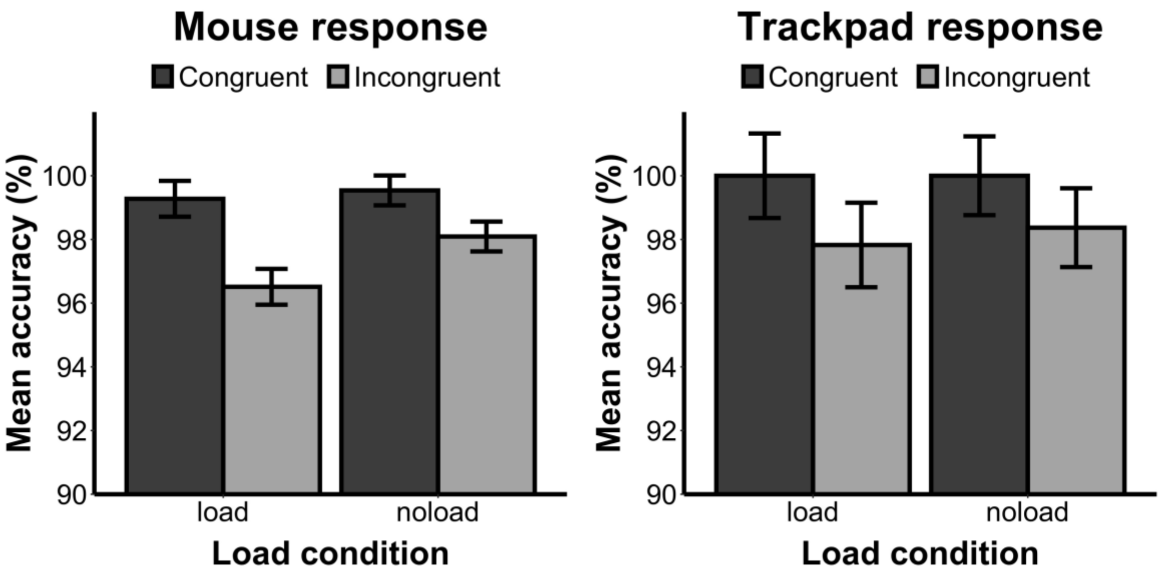


Figure 20: *Bar graph of mean accuracy* results from a post-hoc analysis of response method by congruency and load conditions in Experiments 5 and 6 *(n=118). Higher accuracy is shown in the congruent trials than the incongruent trials* for both response methods. The graph on the left shows the results of participants who responded using a mouse *(n=95)*, whilst the graph on the right shows the results of participants who responded using a trackpad *(n=23)*. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Whilst response method had no effects on accuracy, Figure 21 shows, there was a significant main effect of response method on RT, *F*(1, 116) = 14.74, *p* < .001, d = , *BF* > 100), with participants responding faster with a mouse (mean RT = 472 ms) than a trackpad (mean RT = 589 ms). There was no response method x congruency interaction *F*(1, 116) = 0.20, *p* = .653, *BF* = 0.22. However, there was a response method x load interaction *F*(1, 116) = 9.59, *p* = .002, *BF* > 100, where the difference in RTs between load conditions was larger when a trackpad was used. Participants appeared to respond much slower on load trials than no load trials when they were using a trackpad. However, there was a much smaller difference in RTs when participants used a mouse.

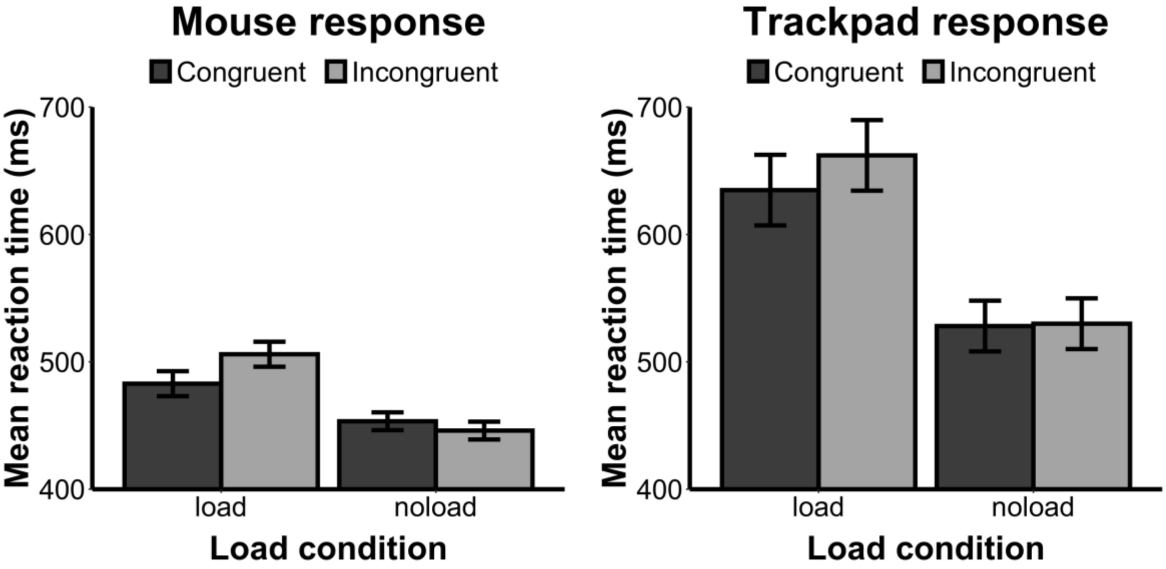


Figure 21: Bar graph of mean reaction times (ms) from a post-hoc analysis of response method by congruency and load conditions in Experiments 5 and 6 (n=118). Faster mouse response time is shown than trackpad response time, and the difference in response times between load conditions was larger when a trackpad was used. The graph on the left shows the results of participants who responded using a mouse (n=95), whilst the graph on the right shows the results of participants who responded using a trackpad (n=23). Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Post-hoc analysis of in-person/online condition. As mentioned above, we felt it was important to explore the effect of conducting the experiment in person (Method 1) or online (Method 2) when comparing results between Experiments 5 and 6. To ensure the procedure was the same for all participants except for the online versus lab environment, we excluded participants who responded using a trackpad, leaving only those who responded using a mouse. This left a sample of 95 participants: 65 who conducted the experiment in-person in a lab and 30 who conducted the experiment online. We ran another ANOVA with online verses laboratory environment as a between-subject factor. As shown in Figure 22, there was no main effect of environment on accuracy, *F(*1,93) = 0.01, *p* = .961, *BF* = 0.18, nor an environment x congruency interaction, *F*(1,93) = 0.40, *p* = .528, *BF* = 0.22, nor environment x load interaction *F*(1,93 = 1.88, *p* = .173, *BF* = 0.36. This means that running the experiment online in a less controlled environment, when a mouse was used to respond, had no effect on participants' overall performance in the task and did not significantly influence the number of action slips made.

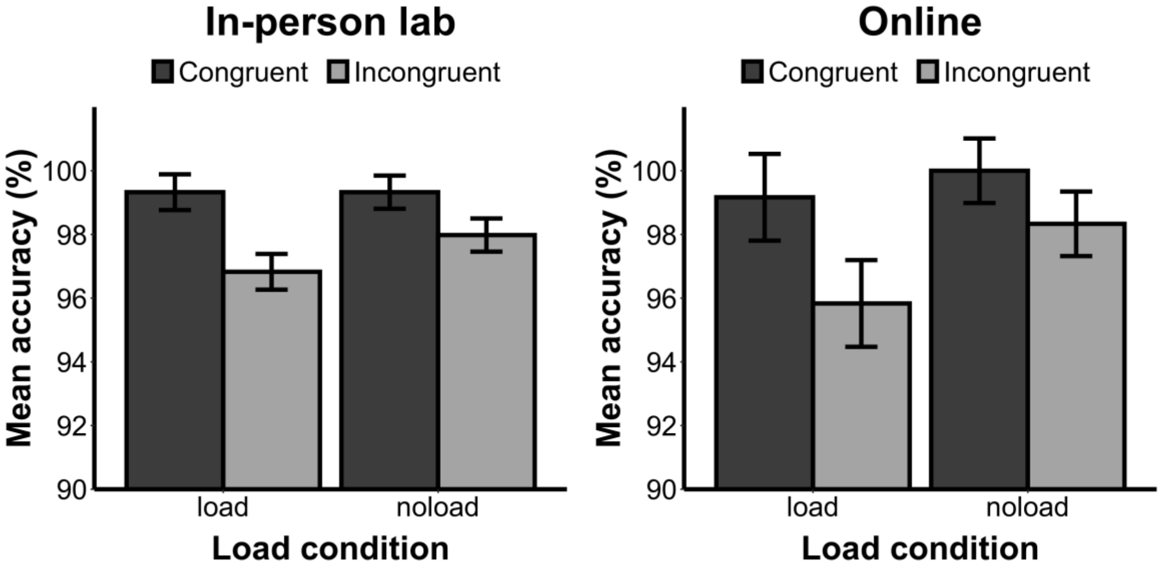


Figure 22: Bar graph of mean accuracy results from a post-hoc analysis of in-person verses online condition by congruency and load conditions in Experiments 5 and 6 (n=95). Higher accuracy is shown in congruent trials than incongruent trials for both in-person and online conditions. The graph on the left shows the results of participants who completed the experiment in-person in the lab (n=65), whilst the graph on the right shows the results of participants who completed the experiment online (n=30). Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Whilst online verses in-person environment had no effect on accuracy, it was considered important to also investigate the effects of online verses in-person environment on RTs. As shown in Figure 23, there was a main effect of online/in-person environment on RT *F*(1,93) =13.80, *p* < .001, BF > 100, with participants responding more quickly in the lab with a mouse (mean RT = 442 ms) compared to online with a mouse (mean RT = 560 ms). There was also a significant online/in-person x load interaction *F*(1,93) = 43.11, *p* < .001, *BF* > 100, where the online/in-person effect on RT was larger under load conditions than no load conditions. There was no online/in-person by congruency interaction *F*(1,93) = 0.49, *p* = .485, *BF* = 0.17. These results suggest that, although participants were generally slower at responding when conducting the experiment online, particularly under load conditions, this RT difference did not depend on the congruency of the trial.

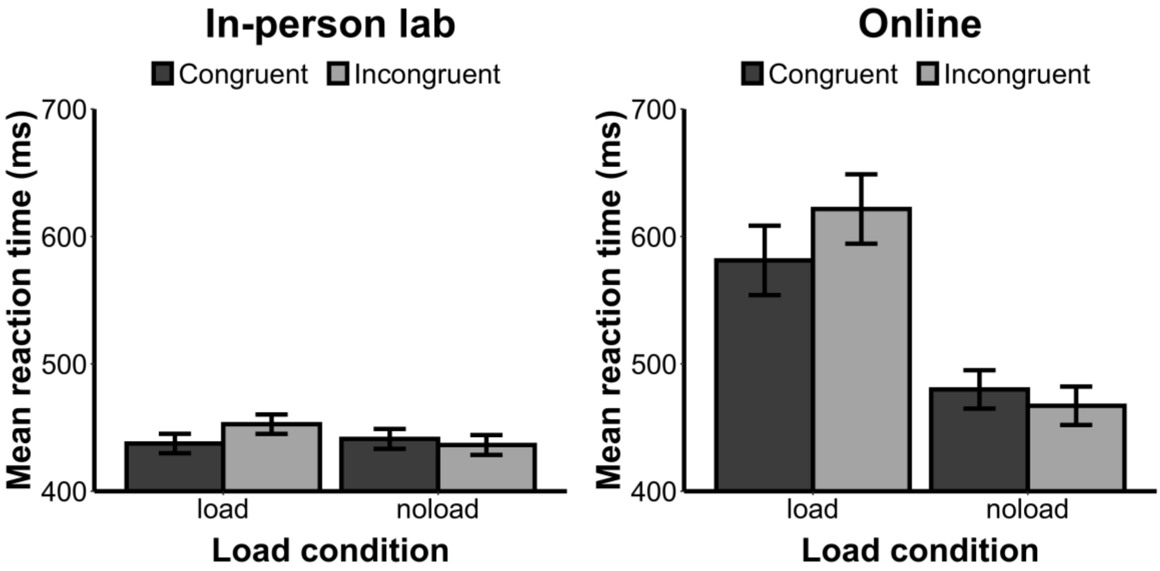


Figure 23: Bar graph of mean reaction times (ms) from a post-hoc analysis of in-person lab verses online condition *by congruency and load conditions in Experiments 5 and 6 (n=95). Faster responses are shown in the in-person lab condition than the online condition, where* the online/in-person effect on RT was larger under load conditions than no load conditions. The graph on the left shows the results of participants who completed the experiment in-person in the lab *(n=65)*, whilst the graph on the right shows the results of participants who completed the experiment online *(n=30)*. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Post-hoc analysis of combining Experiments 5 and 6. Figure 24 shows when the data from both the online and laboratory tasks are combined there is still a congruency effect on accuracy, $F(1,117) = 26.18$, $p < .001$, $BF > 100$, main load effect $F(1,117) = 6.44$, $p < .001$, $d =$, BF = 1.21, and congruency by load interaction $F(1,117) = 4.90$, $p = .029$, $d =$, BF = 0.60. However, the Bayes Factors for the load effect and the congruency by load interaction were inconclusive, so these results should be treated with caution. Participants were more accurate during congruent trials than incongruent trials. On congruent trials, accuracy was very similar on both load and no load trials. However, on incongruent trials, accuracy was lower on load trials than no load trials.

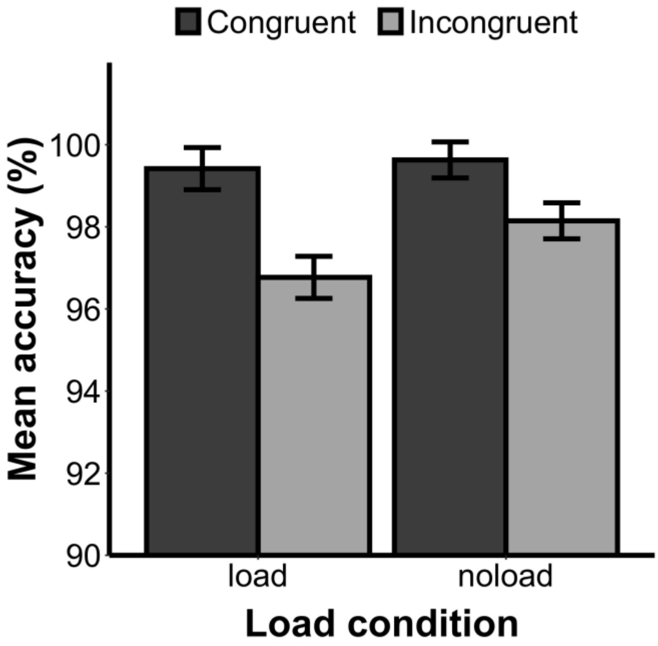


Figure 24: Bar graph of mean accuracy (%) for congruency and load conditions in Experiments 5 and 6 combined (n=118). Higher accuracy is shown in congruent trials than incongruent trials for both load conditions. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

In addition to investigating the effects of congruency and load on accuracy when combining the data from Experiments 5 and 6, we also explored the congruency and load effects on RTs. As seen in Figure 25, there was a main effect of load, $F(1,117) = 35.86$, $p < .001$, $BF > 100$, and no main effect of congruency on RT, F(1,117) = 2.48, p = .118, BF = 0.19. There was a significant congruency x load interaction this time, $F(1,117) = 6.48$, $p = .012$, $BF = 0.96$, where participants were slightly faster on incongruent trials than congruent trials under ‘no load conditions’ but were slower on incongruent trials than congruent trials under ‘load conditions’. This suggests that the complexity of the digit span task on load trials caused participants to slow down their responses on incongruent trials. According to dual-process theory, goal-directed actions are effortful, whereas habitual responses are fast and automatic (Heyes & Dickinson, 1990). Therefore, we would expect participants to respond quickly if they are demonstrating S-R behaviour. However, in Figure 25, our results appear to indicate an element of goal-directed effort being made during incongruent load trials, as participants slow down to compensate on the more difficult trials. Again, the Bayes Factor was inconclusive, so these results should be treated with caution.

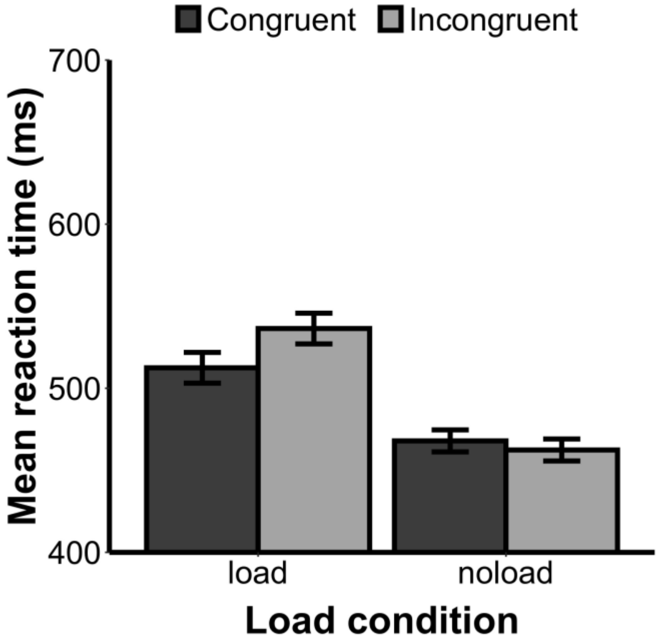


Figure 25: Bar graph of mean reaction times (ms) for congruency and load conditions in Experiments 5 and 6 combined (n=118). Faster responses are shown in the no load condition than the load condition, for both congruency conditions. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

It is worth noting the fast reaction times across all conditions in Experiment 6. These RTs were numerically faster than ‘online’ versions of the experiment. An obvious reason for this could be that participants face distractions in their home environments, which slow down their responses due to reduced working memory capacity. However, the RTs in the current experiment were also faster than previous ‘in-person’ versions of the experiment. A potential explanation may lie in the method of responding. In all other ‘in-person’ experiments in this thesis, participants used a joystick to make their responses. In our ‘online’ versions of the experiment, a large proportion of participants used a laptop trackpad. It is possible that use of a joystick or trackpad could have slowed down the overall mean RTs. Whereas, in the current experiment, participants were able to make their responses quickly, using a standard mouse on a desktop computer.

Regardless of the explanation for faster reaction times, it appears that participants may have prioritised responding quickly in this experiment. It is possible that the time pressure imposed across all trials also reduced participants’ working memory capacity on ‘no load’ trials, compromising the effect of our digit span load task. We expected it to increase the congruency effect on load trials, but the observed congruency x load effect was not supported by the Bayesian analysis. We have consistently found a congruency effect on accuracy in our experiments, but our distraction/load manipulations have not clearly increased this effect. Similar to Luque et al.'s (2020) findings, it seems that participants slow down their responses to avoid decreasing their accuracy in distraction/load trials. In this chapter, we attempted to manipulate action slips by restricting participants' working memory capacity. The next step was to explore the opposite approach: whether freeing-up working memory capacity allows participants to focus on the goal, and therefore removes the congruency effect. We felt a simple method to do this was to remove the time pressure in the experiment.

Experiment 7 – online between-subject time pressure

In order to explore the theory that imposing a time pressure has an effect on action slips, Experiment 7 applied our basic online experimental procedure (Method 2) with a congruent x incongruent trial manipulation, but no with additional within-subject manipulation (e.g., O-S delay or cognitive load task). Instead, we applied a between-subject manipulation of time pressure. In Experiments 1-6, participants were exposed to all conditions. This can lead to carryover effects where the experience of one condition influences responses in subsequent conditions (Poling et al., 1995). Therefore, it felt important to explore this procedure using a between-subject design in Experiment 7, rather than a within-subject design. Half of the participants were told to respond as quickly as possible, with points based on their response time, and the other half of participants were simply told to respond as accurately possible, with no mention of response time.

Method

Participants. 156 undergraduate psychology students (112 females, 41 males, 3 other) from the University of Plymouth participated in exchange for course credit (78 participants in each time pressure condition). The participants were aged between 18 and 51 years (*M* = 21.69 years, *SD* = 6.23 years).

Design, Stimuli and Procedure. The experiment followed a 2 (congruency condition: congruent, incongruent) by 2 (time pressure condition: time pressure, no time pressure) mixed design, where the congruency condition was manipulated within-subjects, and the time pressure condition was manipulated between-subjects. The ‘online’ (Method 2) experimental design was applied. The key difference between conditions was in the test phase instructions. Participants in the ‘time pressure’ condition were given the same instructions as previous experiments, where they were told to respond as quickly as possible. Participants in the ‘no time pressure’ condition were instructed, *“There is no time pressure in the main test trials, so you don’t need to rush. You will win 5 points for all correct responses and 0 points for all incorrect responses.”* This was to ensure participants in the ‘no time pressure’ condition focused on accuracy rather than speed.

Results

Exclusions. The exclusion criteria from previous experiments were applied, leaving 59 participants in each time pressure condition (time pressure: 43 females, 15 males, 1 other; no time pressure: 45 females, 12 males, 2 other).

Training. Figure 26 shows, as in previous experiments, participants learnt the contingencies quickly, with a long period during which participants were asymptote.

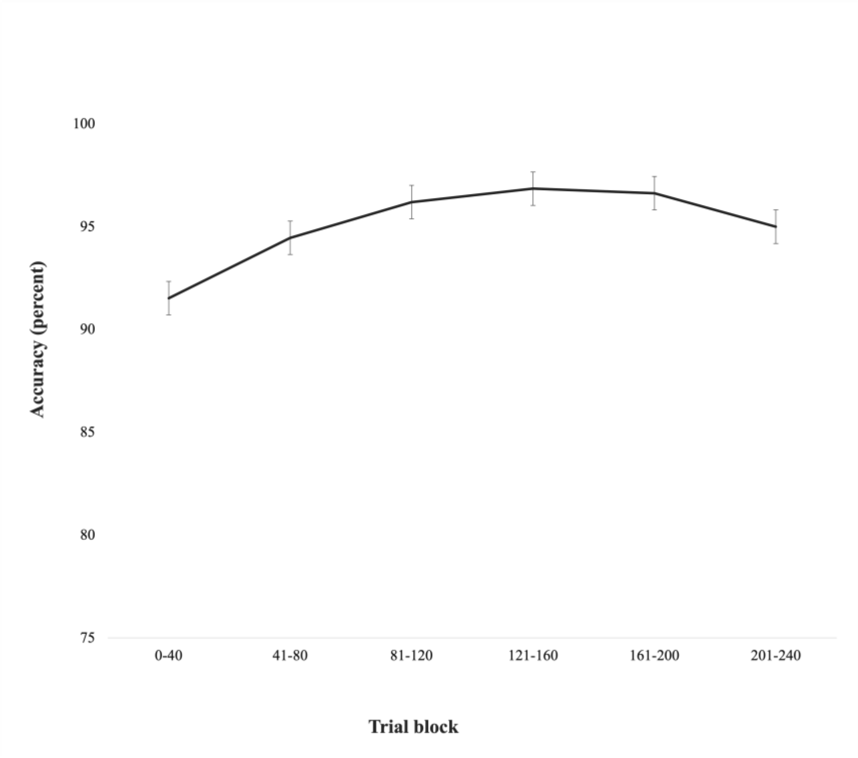


Figure 26: Line graph of mean accuracy (%) of training trials for S1 and S2 in Experiment 7. Over 90% correct responses are shown after the initials 40 trials. Error bars represent the standard error of the mean for each training trial interval.

Test response accuracy. Figure 27 shows the mean percentage of correct responses during the test, separated by congruency and time pressure conditions. The graph suggests that the participants were more accurate on congruent trials than incongruent trials in both time pressure groups. The graph also shows that the no time pressure group made more errors than the time pressure group. A 2 (congruency condition: congruent vs. incongruent) by 2 (time pressure condition: time pressure vs. no time pressure) mixed model ANOVA confirmed that there was a significant main effect of congruency on response accuracy, $F(1, 116) = 9.39$, $p = .003$, $d_z = 0.28$, $BF_{10} = 11.58$, with higher accuracy during congruent trials than during incongruent trials. Whilst there was a higher mean accuracy in the time pressure group than no time pressure group, this effect was not significant and the Bayes Factor was inconclusive, $F(1, 116) = 1.63$, $p = .204$, $d_z = 0.19$, $BF_{10} = 0.39$. There was also no congruency × time pressure interaction, $F(1, 116) = 0.24$, $p = .622$, $BF_{10} = 0.22$.

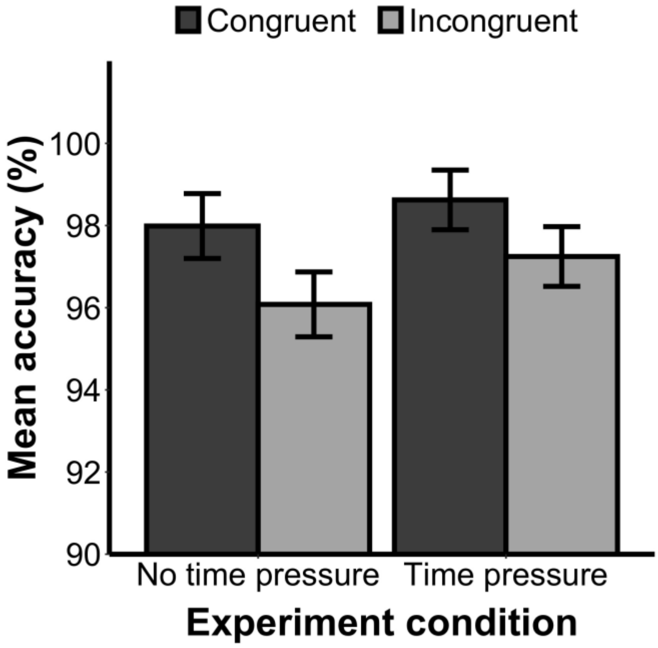


Figure 27: Bar graph of mean accuracy (%) for congruency and time pressure conditions in Experiment 7 (n=118). Higher accuracy is shown in congruent trials than incongruent trials in both time pressure groups. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Test reaction times. As in previous experiments, a comparable analysis was conducted on the RTs for correct trials only at test. Figure 28 shows there was a main effect of time pressure condition between subjects, $F(1, 116) = 4.34$, $p = .040$, $d_z = 0.37$, with participants responding more quickly on time pressure trials than no time pressure trials, but the Bayes Factor was inconclusive, BF_{10} = 1.69. No significant effect of congruency condition, $F(1, 116) = 0.67$, $p = .415$, $BF_{10} = 0.19$, nor congruency × time pressure interaction, $F(1, 116) = 1.53$, $p = .219$, $BF_{10} = 0.15$, was observed.

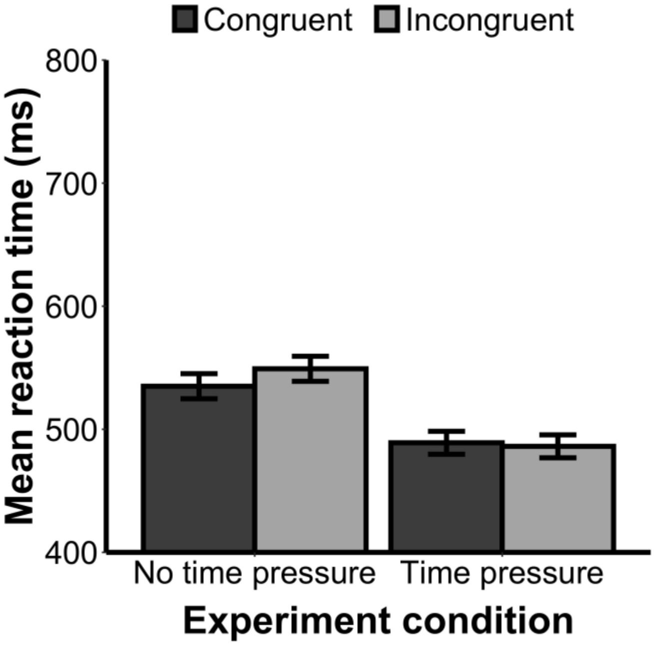


Figure 28: Bar graph of mean reaction times (ms) for congruency and time pressure conditions in Experiment 7 (n=118). Faster responses are shown in the time pressure condition than the no time pressure condition, for both congruency conditions. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Discussion

Experiment 7 used a simple congruent verses incongruent version of the experiment, with a manipulation to measure the effect of time pressure between subjects. We have consistently demonstrated a congruency effect in Experiments 1-6, where participants were instructed to respond as quickly as possible on test. Dual-process theories argue that goal-directed processes operate relatively slowly (Moors & De Houwer, 2006), compared to automatic habitual processes. Therefore, we predicted that removing any time pressure from the experiment would result in no congruency effect, as participants would have more capacity to keep the goal in mind and avoid action slips. We would, therefore, see a greater effect of congruency in the time pressure condition than in the no time pressure condition. Our results found a congruency effect on accuracy for both conditions, but no congruency effect on RT. As in previous experiments, this reveals no speed-accuracy trade-off.

Overall, accuracy was very high in both conditions, so it may be the case that without an additional within-subject manipulation, the task was easier than previous experiments for both groups and automatic processes were not triggered. As would be expected, there was a significant main effect of time pressure on RT. Participants who were told to respond as quickly as possible responded faster than participants under no time pressure, regardless of congruency.

General discussion

As in Chapter 2, this chapter provides evidence for habitual behaviour in all four experiments, where accuracy was higher on congruent trials than incongruent trials. There was no main effect of congruency on RT in any of the four experiments, suggesting that lower accuracy during incongruent trials was not due to a speed-accuracy trade-off. Rather, it seems that the more difficult incongruent trials produced more action slips, perhaps due to the previously associated S-R relationship during training. The main aim of Chapter 3 was to explore whether reducing participants' working memory capacity would lead to more action slips, indicating habitual behaviour. In addition to our standard congruency manipulation, we embedded a working memory task into half of the test trials in Experiments 4-6. Experiments 4 and 5 found no main effect of the working memory manipulation on accuracy, but there was a main effect of this manipulation on RTs for both experiments. This suggests that when trials became more difficult, participants compensated by taking longer to respond. Thus, accuracy was not affected when participants slowed their responses. Experiment 6 ran the same programme as Experiment 5, but in the lab, rather than online, and with twice as many test trials. This time, we found a main load effect on accuracy, but no effect of load on RT. Participants responded equally fast across all trials, regardless of load or congruency condition. Accuracy was higher during congruent trials than incongruent trials, and this difference for larger on incongruent trials than congruent trials. However, the observed congruency x load effect was not supported by the Bayesian analysis, so this result should be treated with caution. This indicates that the load manipulation increased action slips in Experiment 6, and this was not due to a speed-accuracy trade-off. It is not clear why the RT results of Experiment 5 and 6 differed, as the only difference in the procedure was the environment in which the experiment was conducted (Method 1 verses Method 2). Perhaps this was due to the added pressure of performing the experiment in a lab with an experimenter present in Experiment 6.

Our experiments consistently show a main congruency effect on accuracy, with more action slips on incongruent trials than congruent trials providing evidence of habitual behaviour. We have also shown that this congruency effect increases when working memory capacity is reduced. However, our congruency effect still appeared in Experiment 7 even when time pressure was removed. If our congruency effect is evidence of habitual behaviour, we would expect this effect to disappear when participants are told to focus on accuracy instead of speed. A possible explanation for this is that, rather than unconscious, S-R behaviour, the few errors made on incongruent trials were due to conscious goal-directed actions. Participants may have purposely chosen the incorrect response because they momentarily believed it was the correct response. It could be that our observed action slips are due to two competing goals, rather than two separate processes. This theory will be discussed further in Chapter 6.

It is not yet clear whether our congruency effect is due to two separate systems operating goal-directed action sand automatic habitual responses, or whether the action slips observed in our experiments are due something else that is primarily goal-directed. This raises the question of how the main effect of congruence – our observation of stimulus-driven action slips – relates to other similar results already in the literature. The standard manipulation to test for stimulus-driven behaviour is to devalue the outcome between training and test (Adams & Dickinson, 1981; de Wit et al., 2018; Luque et al., 2020; Tricomi et al., 2009). Therefore, it felt important to develop a version of our experiment that is more in line with existing devaluation procedures. If we are able to demonstrate a congruency effect when the outcome is devalued, this would provide evidence for habitual behaviour, and support dual process theories.

Chapter 4 – Devaluation and training

Chapter 4 aimed to explore the extent to which our congruency effects (shown in Chapters 2 and 3) connect to the existing habit literature. As discussed in Chapter 1, dual-process theories of instrumental learning propose that instrumental behaviours are controlled by two distinct systems: the goal-directed system and stimulus-triggered habitual system (e.g., Dickinson & Balleine, 1994; Verplanken, 2018). The goal-directed system is suggested to facilitate deliberate, reward-driven actions, while the habitual system is suggested to produce comparatively automatic responses based on stimulus-response (S-R) associations. When participants stop responding for outcomes that have been devalued in an outcome devaluation procedure, this is seen as evidence for goal-directed behaviour. The outcome is no longer valuable to the participant. Therefore, they no longer make the conscious effort to obtain it. On the other hand, if responding continues for a devalued outcome, this is considered evidence for automatic, habitual behaviour, developed through repeated S-R training. In line with this, extensive training is considered critical for developing habitual responses. Automatic habit behaviour is formed through repeating the behaviour to the point where conscious effort is no longer needed to make this response. For example, in Tricomi's et al.'s (2009) experiment, participants who had received moderate training performed significantly fewer instrumental responses for the devalued outcome than the valued outcome, showing evidence for goal-directed behaviour. Participants who received extensive training, by contrast, showed no significant difference in response rates for the still valued and devalued outcomes, providing evidence of habitual behaviour. The key difference between our procedure and devaluation procedures is that our experiments focus on outcome availability rather than outcome value. In the test phase of our procedure, each trial begins with the presentation of either outcome O1 or O2. The participants are instructed to perform the response that, during the training phase, produced the presented outcome (i.e., O1 → R1, O2 → R2). Hence, the presentation of the outcomes at the start of each test trial communicates to participants which outcome they could earn on that trial (which food points are available), rather than which outcome is worth more points. Therefore, Chapter 4 aims to adapt our procedure to follow traditional devaluation procedures more closely. Furthermore, we have not yet manipulated the amount of training participants receive to see whether this has an effect on action slips.

To summarise, Chapter 4 aims to apply two approaches to our experimental procedure that are commonly used in existing literature to measure habitual behaviour verses goal-directed action: manipulation of outcome value, and manipulation of amount of training (Adams & Dickinson, 1981; de Wit et al., 2018; Luque et al., 2020; Tricomi et al., 2009). These experiments will include the O-S delay manipulation from Chapter 2, to increase task complexity. As our procedure is simple, and participants have high rates of accuracy in both congruent and incongruent trials, we felt it would be too easy for participants to keep the goals in mind without imposing the delay manipulation. We ran an additional study without this variability in O-S delay and it did not produce the congruency effect. Therefore, the variability in the delay seems to enhance the congruence effect.

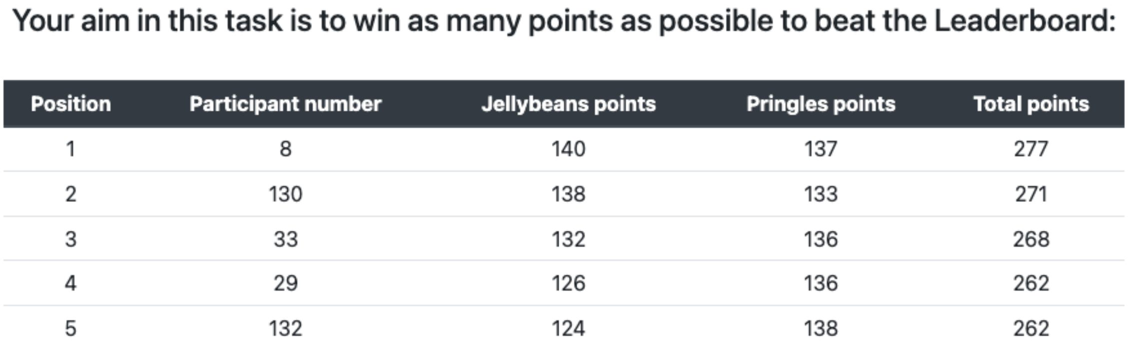
Experiment 8 – devaluation

As discussed in Chapter 1, outcome devaluation has become a standard procedure for distinguishing goal-directed behaviour from habitual behaviour by specifically manipulating the value of the outcomes at test and observing the resulting effects on participants' performance (Corbit, 2018). The aim of Experiment 8 was to investigate whether we could find evidence of habitual behaviour by adapting our procedure to closer reflect the existing devaluation procedures. As in previous experiments, participants were told that they would win jellybean/Pringles points for making the correct response during the test trials. However, in this experiment we devalued one of the outcomes by presenting both outcomes together at the start of each test trial, with a cross over the outcome that produced no points. This followed the same devaluation procedure of de Wit et al.'s (2007) Fabulous Fruit Task, where a cross superimposed on one of the fruits signalled that the instrumental response that produced that fruit previously no longer earned points, and participants were to perform the instrumental response that previously produced the still-valued outcome. We also emphasised the value of the outcome and the importance of the points by adding a Leaderboard to Experiment 8.

Method

Participants. Eighty-three undergraduate psychology students (73 females, 10 males) from the University of Plymouth participated in exchange for course credit. The participants were aged between 18 and 44 years (*M* = 21.86 years, *SD* = 5.88 years).

Design, Stimuli and Procedure. The online method (Method 2) was applied. The experiment was very similar to Experiment 1 in Chapter 2, with the following exceptions. At the beginning of the experiment, participants were told that the aim of the task was to win as many jellybeans and Pringles points as possible to beat the Leaderboard, as shown in Figure 29.



Your aim in this task is to win as many points as possible to beat the Leaderboard:

Position	Participant number	Jellybeans points	Pringles points	Total points
1	8	140	137	277
2	130	138	133	271
3	33	132	136	268
4	29	126	136	262
5	132	124	138	262

Figure 29: Example of Leaderboard shown to participants at the start of the Experiment 8. Participants are told the aim of the task is to win as many points as possible to beat the Leaderboard.

Whilst the short delay condition in the test phase was the same as Experiment 1 (3-4 seconds), the long delay condition was increased to 19-20 seconds. This was to increase the difference between the two delay conditions further, and to see what would happen when the long delay was increased to a point that may result in mind wandering (Schooler et al., 2011). In previous experiments, we observed more action slips in short delays than long delays. Our explanation for this is that participants did not always have enough time to prepare their response when there was a short O-S delay. Therefore, increasing the long delay could serve to give them even longer to prepare their response, or it could allow more time to lose focus on the goal. Participants were warned in the test trial instructions that there would sometimes be a long delay so as not to think the screen had frozen. During the test trials, instead of displaying one outcome picture before the screen changed colour, both the jellybeans and Pringles pictures were displayed together with a cross through one of them, as shown in Figure 30. Participants were instructed at the start of the test that the cross indicated which outcome had no points available on that trial, and therefore, no value. Presentation of the jellybean and Pringles pictures were randomised on each test trial.

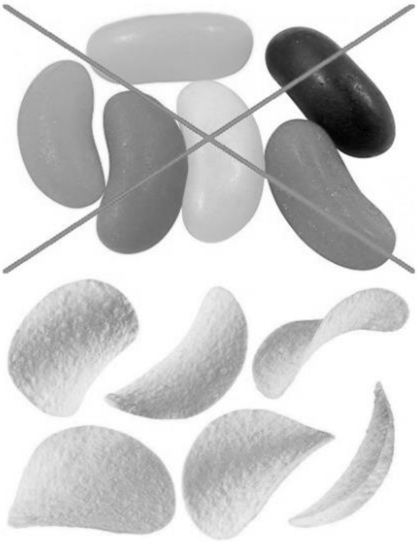


Figure 30: Example of how the devalued outcome is shown to participants at the start of each test trial in Experiment 8. The cross indicates which outcome is worth no points on that trial.

At the end of the experiment, participants were shown a Leaderboard of the top 5 scorers, based on the points earned in the test trials. If the participant's score was one of the five highest, their score was shown to them in red in the Leaderboard at the end of the experiment.

Results

Exclusions. The exclusion criteria from previous experiments were applied. Twenty-one participants were removed for achieving less than 80% accuracy in test trials, leaving 62 participants (54 females, 8 males).

Training. Figure 31 shows, as in previous experiments, the participants learnt the contingencies quickly, staying above 90% accuracy after 40 training trials.

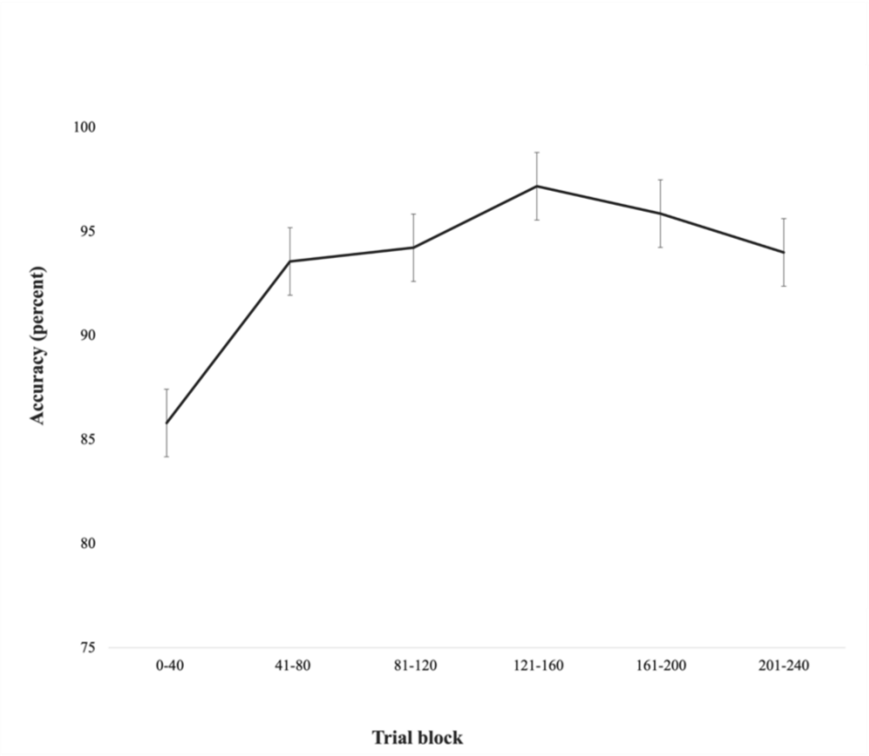


Figure 31: Line graph of mean accuracy (%) of training trials for S1 and S2 in Experiment 8. Over 90% correct responses are shown after the initials 40 trials. Error bars represent the standard error of the mean for each training trial interval.

Test response accuracy. Figure 32 shows the mean percentage of responses that earned points during the test, separated by congruency and delay conditions. The graph suggests that the participants were more accurate on congruent trials than incongruent trials, and that this pattern was not affected by the delay manipulation. A 2 (congruency condition: congruent vs. incongruent) by 2 (delay condition: short delay vs. long delay) repeated-measures ANOVA, however, revealed that the difference between response accuracy on congruent and incongruent trials was not significant, $F(1, 61) = 3.21$, $p = .078$, $d_z = 0.23$. However, the Bayes Factor was inconclusive, $BF_{10} = 0.47$. Therefore, this was not evidence for the null hypothesis. There was no main effect of the delay condition, $F(1, 61) = 0.14$, $p = .709$, $BF_{10} = 0.15$, nor a congruency × delay interaction, $F(1, 61) = 0.02$, $p = .885$, $BF_{10} = 0.21$.

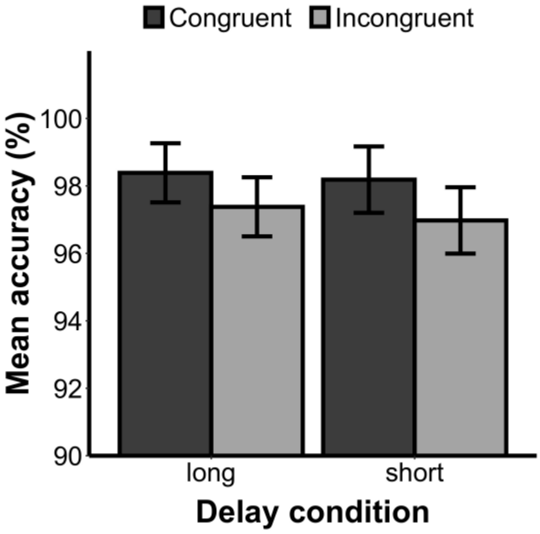


Figure 32: Bar graph of mean accuracy (%) for congruency and delay conditions in Experiment 8 (n=62). Higher accuracy is shown in congruent trials than incongruent trials in both delay conditions. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Test reaction times. As in previous experiments, a comparable analysis was conducted on the RTs for correct trials only at test. Figure 33 shows there was a main effect of delay condition on RT, $F(1, 61) = 11.68$, $p = .001$, $d_z = 0.37$, with participants responding more quickly on long delay trials than short delay trials. The Bayes Factor analysis provided extreme evidence for this delay effect, $BF_{10} > 100$. No significant effect of congruency condition, $F(1, 61) = 0.74$, $p = .394$, $d_z = 0.11$, $BF_{10} = 0.19$, nor congruency × delay interaction, $F(1, 61) = 0.33$, $p = .566$, $BF_{10} = 0.19$, was observed.

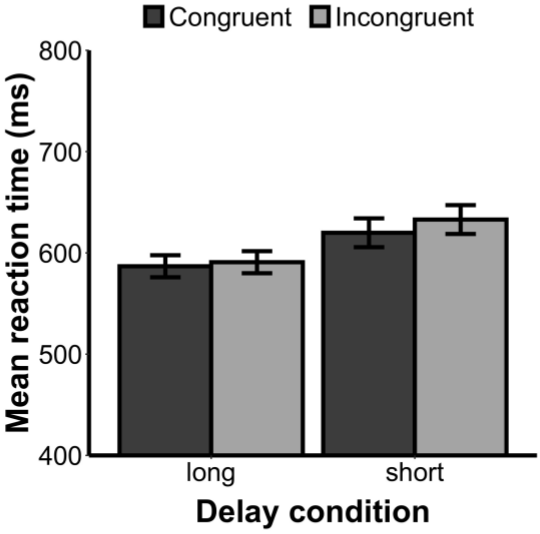


Figure 33: Bar graph of mean reaction times (ms) for congruency and delay conditions in Experiment 8 (n=62). Faster responses are shown in the long delay condition than the short delay condition, for

both congruency conditions. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Discussion

Experiment 8 applied a very similar procedure to Experiment 1 in Chapter 2, with a few changes in design to be more in line with existing devaluation procedures. A cross was shown over the outcome with no value to highlight that one of the outcomes had been devalued on the trial. A Leaderboard was also added to highlight the aim of winning points. In all previous experiments in Chapters 2 and 3, we found a significant congruency effect, indicating that participants were making action slips on incongruent trials. However, in this experiment, there was no significant congruency effect. This suggests that participants were able to keep the goal in mind, even for incongruent trials, and respond to the valued outcome. If participants were making automatic S-R responses, we would have expected to see more errors on incongruent trials, regardless of outcome value (Adams & Dickinson, 1981; de Wit et al., 2018; Luque et al., 2020; Tricomi et al., 2009). One possible explanation for this is that emphasising the value of the outcome strengthened goal-directed behaviour across all trials. Participants may have found it easier to keep the goal in mind and ignore the stimulus when the devalued outcome was also presented at the start of the trial. It is also possible that adding a Leaderboard for participants to beat other participants’ scores added an element of competition and therefore increased their motivation to earn points across all trials (Landers et al., 2015). It is important to note that this null result is not because the devaluation did not 'work' - participants were extremely accurate in this experiment. Therefore, the devaluation, if anything, worked too well. Participants responded for the outcome that would earn points and ignored the screen colour. Little, if no, action slips were made.

The results in Experiment 8 are in line with previous findings, that devaluation procedures with human participants result in evidence for goal-directed behaviour, rather than habitual behaviour. These results raise the question of whether the congruency effect seen in our Experiments presented in Chapters 2 and 3 is evidence to support a dual-process account of habitual behaviour. In order to investigate this further, Experiment 9 focuses on manipulating the amount of training presented to participants, as this is a procedure commonly used in the existing literature to measure habitual behaviour verses goal-directed action.

Experiment 9 – long verses short training

As existing research emphasises the importance of amount of training for increased habitual behaviour, Experiment 9 applied a between-subject manipulation of training. We ran our standard congruency x delay procedure applied in Chapter 2. However, whilst one group of participants were given the long training used in previous experiments (240 training trials), a second group of participants were given a shorter version of the training (24 training trials). As Experiments in Chapter 2 showed a congruency effect, we expected to see a congruency effect for participants conducting the same experiment (those in the ‘long training’ condition). However, in line with existing research on habitual behaviour (Adams & Dickinson, 1981; de Wit et al., 2018; Luque et al., 2020; Tricomi et al., 2009), we expected the congruency effect to significantly reduce for the participants in the ‘short training’ condition.

Method

Participants. 171 undergraduate psychology students (144 females, 27 males) from the University of Plymouth participated in exchange for course credit (90 participants in the short training condition, 81 participants in the long training condition). The participants were aged between 18 and 57 years (*M* = 19.68 years, *SD* = 3.79 years).

Design, Stimuli and Procedure. The in-person method (Method 1) was applied. The experiment used a mixed design, where the congruency and delay conditions were manipulated within-subjects, as shown in Experiment 3, and amount of training was manipulated between-subjects. As previously discussed, we have other data to suggest that a very simple task without the O-S delay manipulation does not produce the basic congruency effect. Therefore, we felt it was important to include it in this experiment. Participants in the ‘long training’ condition were given the same procedure as Experiment 3, with 240 training trials. The procedure was the same for participants in the ‘short training’ condition, however they only received 24 training trials.

Results

Exclusions. The exclusion criteria from previous experiments were applied, leaving 66 participants in each training condition (long training: 57 females, 9 males; short training: 56 females, 10 males).

Training. As in previous experiments, the participants in the long training condition learnt the contingencies quickly, with accuracy remaining above 90% after 42 training trials, as shown in Figure 34. The mean accuracy across the 24 training trials in the short training condition was 83.52% (this corresponds to halfway through the first bin in Figure 34).

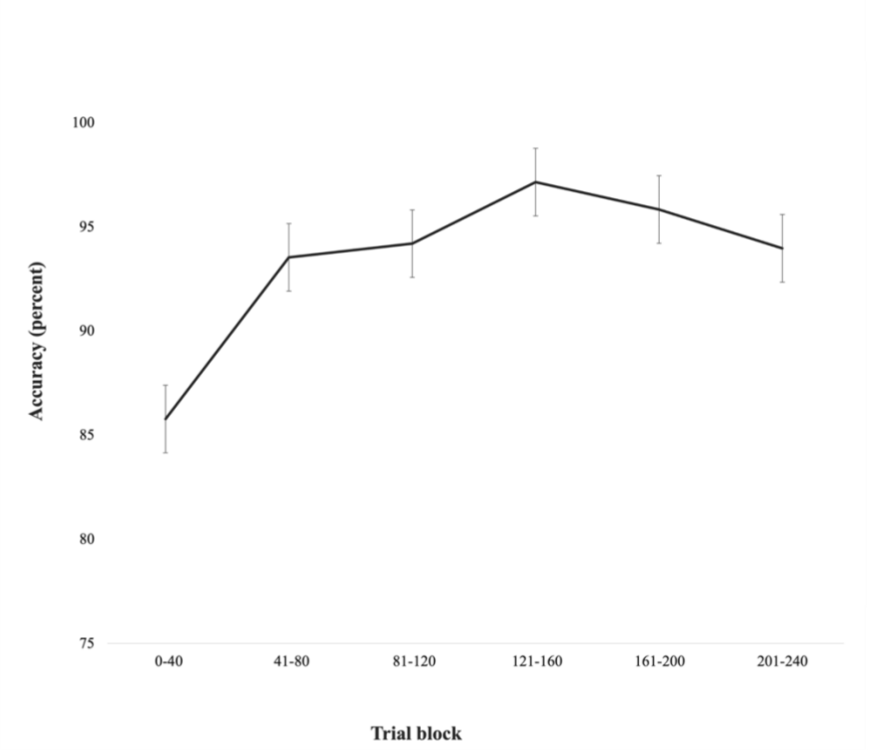


Figure 34: Line graph of mean accuracy (%) of training trials for S1 and S2 in Experiment 9. Over 90% correct responses are shown after the initials 40 trials. Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Test response accuracy. Figure 35 shows the mean percentage of correct responses during the test, separated by congruency, delay and training conditions. The graph suggests that the participants were more accurate on congruent trials than incongruent trials in both training conditions. However, for participants who received short training, this difference was only apparent in the short delay condition. The graph also shows that participants made more errors in the short delay trials than the long delay trials in both training groups, but this was particularly prominent in the incongruent trials. A 2 (congruency condition: congruent vs. incongruent) by 2 (delay condition: short delay vs. long delay) by 2 (training condition: short training vs. long training) mixed ANOVA confirmed that there was a significant main effect of congruency on response accuracy and extreme support from the Bayes Factor analysis, $F(1, 130) = 20.40$, $p < .001$, $d_z = 0.39$, $BF_{10} > 100$, with higher accuracy during congruent trials than during incongruent trials. There was also a significant main effect of delay with extreme evidence from the Bayes Factor analysis, $F(1, 130) = 16.90$, $p < .001$, $d_z = 0.36$, $BF_{10} > 100$, with higher accuracy in long delay trials than in short delay trials, and a significant congruency x delay interaction, $F(1, 130) = 11.60$, $p < .001$, $BF_{10} = 10.58$, where the congruency effect was more prominent in short delay trials than long delay trials. There was no main effect of training condition on response accuracy , $F(1, 130) = 0.08$, $p = .784$, $BF_{10} = 0.15$, congruency × training interaction, $F(1, 130) = 0.98$, $p = .323$, $BF_{10} = 0.21$, delay × training interaction, $F(1, 130) = 0.17$, $p = .682$, $BF_{10} = 0.14$, nor congruency × delay x training interaction, $F(1, 130) = 0.14$, $p = .706$, $BF_{10} = 0.06$.

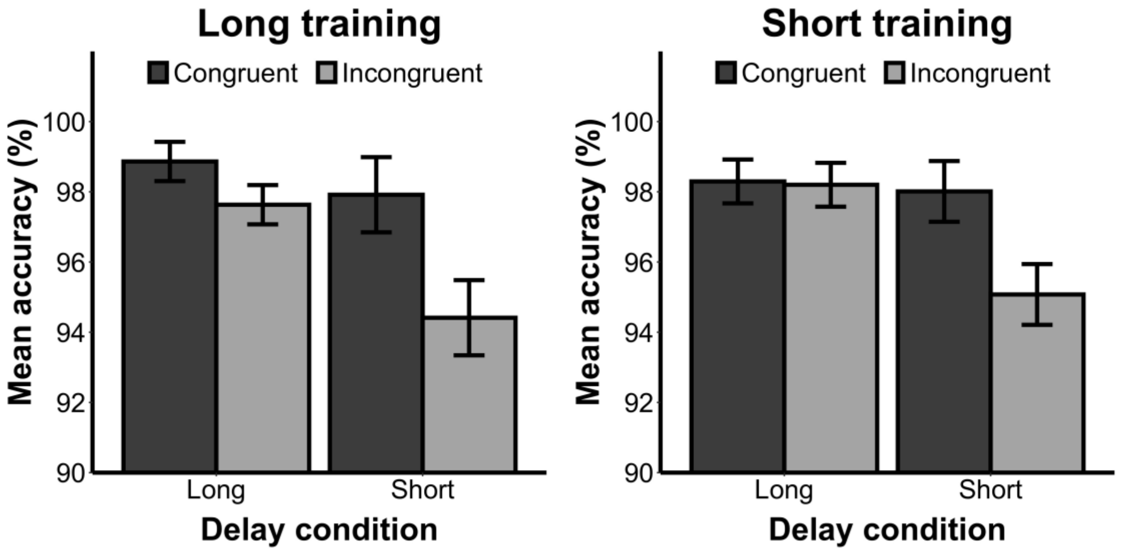


Figure 35: Bar graph of mean accuracy (%) for congruency and delay conditions for both training groups in Experiment 9 (n=132). Higher accuracy is shown in congruent trials than incongruent trials in both training groups. In the short training group, the congruency effect was only observed in the short delay condition. The graph on the left shows the results of participants in the long training condition (n=66), whilst the graph on the right shows the results of participants in the short training condition (n=66). Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Breaking our results down further, we ran repeated-measures ANOVAs on each training group individually. In the ‘long training’ condition, there was a significant main effect of congruency, on response accuracy, $F(1, 65) = 14.82$, $p < .001$, $d_z = 0.47$, $BF_{10} = 59.92$, with higher accuracy during congruent trials than during incongruent trials. There was also a significant main effect of delay, $F(1, 65) = 8.88$, $p = .004$, $d_z = 0.37$, $BF_{10} = 14.36$, with higher accuracy in long delay trials than in short delay trials. The congruency x delay interaction was not significant, $F(1, 65) = 3.65$, $p = .060$, however, the Bayes Factor analysis was inconclusive, and therefore did not support a null hypothesis, $BF_{10} = 0.76$.

In the 'short training' condition, there was a significant main effect of congruency, on response accuracy, $F(1, 65) = 6.36$, $p =.014$, $d_z = 0.31$, $BF_{10} = 3.45$, with higher accuracy during congruent trials than during incongruent trials. There was also a significant main effect of delay, $F(1, 65) = 8.06$, $p = .006$, $d_z = 0.40$, $BF_{10} = 10.44$, with higher accuracy in long delay trials than in short delay trials, and a significant congruency x delay interaction, $F(1, 65) = 9.60$, $p = .003$, $BF_{10} = 3.38$, where the congruency effect only occurred in the short delay trials, $t(65) = 3.38$, $p = .001$, $d_z = 0.42$. There was no congruency effect in the long delay trials, $t(65) = 0.15$, $p = .880$.

Test reaction times. As in previous experiments, a comparable analysis was conducted on the RTs for correct trials only at test. Figure 36 shows there was a main effect of congruency condition, $F(1, 130) = 13.48$, $p < .001$, $d_z = 0.32$, $BF_{10} = 15.15$, with participants responding more quickly on congruent trials than incongruent trials. There was also a main effect of delay on RT and extreme evidence for this effect, $F(1, 130) = 31.52$, $p < .001$, $d_z = 0.47$, $BF_{10} > 100$, with participants responding more quickly on long delay trials than short delay trials, and a significant delay x training interaction, $F(1, 130) = 9.03$, $p = .003$, $BF_{10} = 36.82$, where a short delay resulted in a slower reaction time for participants in the long training condition only. No significant effect of training condition, $F(1, 130) = 0.81$, $p = .371$, $BF_{10} = 0.45$, congruency × delay interaction, $F(1, 130) = 0.42$, $p = .520$, $BF_{10} = 0.15$, congruency × delay x training interaction, $F(1, 130) = 0.16$, $p = .687$, $BF_{10} = 0.34$, or congruency × training interaction, $F(1, 130) = 3.33$, $p = .070$, $BF_{10} = 0.44$, were observed.

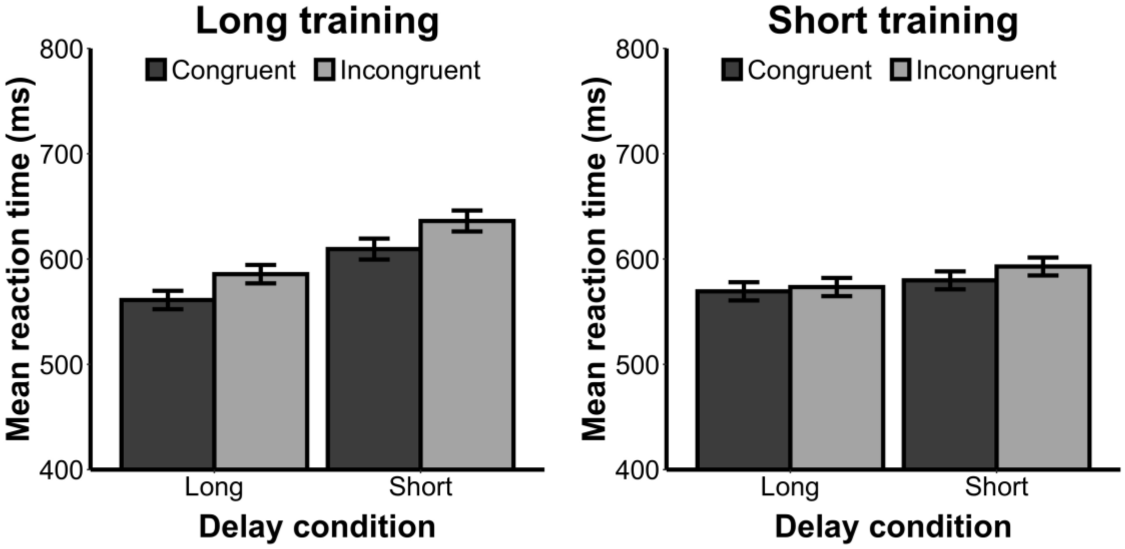


Figure 36: Bar graph of mean reaction times (ms) for congruency and delay conditions for both training groups in Experiment 9 (n=132). Faster responses are shown in congruent trials than incongruent trials, with faster responses in the long delay condition than the short delay condition in the long training group only. The graph on the left shows the results of participants in the long training condition (n=66), whilst the graph on the right shows the results of participants in the short training condition (n=66). Error bars represent difference-adjusted, 95% within-subjects confidence intervals.

Discussion

In Experiment 9, participants in the ‘long training’ condition conducted the same experiment as those in Experiment 3. As we found a significant congruency effect on accuracy in Experiment 3 (and in other experiments in Chapters 2 and 3), we expected to also find a significant congruency effect on accuracy for participants in the ‘long training’ condition. The ‘long training’ results were in line with our prediction, showing evidence for habitual behaviour with extensive training. In relation to the ‘short training’ condition, we expected the congruency effect on accuracy to disappear, as existing research specifies that extensive training is required to develop habitual behaviour (Adams & Dickinson, 1981; de Wit et al., 2018; Luque et al., 2020; Tricomi et al., 2009). However, we found a congruency effect in the ‘short training’ condition, and there was no difference in congruency effect between the ‘short training’ condition and the ‘long training’ condition.

As in previous experiments, we found a main congruency effect, providing evidence of action slips, which is in line with the dual process theory. However, there is evidence of action slips in the short training group as well as the long training group, which suggests our congruency effect may not be showing habits after all. Although, of course, it is possible that the task is so simple that even as few as 24 training trials in this experiment was enough to train a habit. This will be discussed future in the general discussion.

General discussion

In Chapters 2 and 3, we consistently saw a congruency effect in our experiments. However, when we tried to manipulate devaluation in Experiment 8, we did not observe a congruency effect. Existing literature uses the outcome devaluation procedure to test for evidence of habitual behaviour (de Wit et al., 2018; Luque et al., 2020; Tricomi et al., 2009). According to dual-process theory, continued responding for a devalued outcome can be considered evidence for habitual behaviour (Adams & Dickinson, 1981). On the contrary, high accuracy on incongruent trials suggests a strong effect of devaluation. Accuracy was very high across all conditions in Experiment 8. Therefore, it may be beneficial to use a more subtle devaluation procedure in future experiments to allow an effect of the Stimulus. We also found no effect of amount of training on congruency. According to dual process theories, we would expect to see goal-directed control after moderate instrumental training, but habitual control after extensive instrumental training. However, we found no significant difference between the two groups. The congruency effect was still there when we shortened training. It is possible that our congruency effect is measuring something other than habitual behaviour (or automatic habitual behaviour), or that our task is so simple that we would need to reduce the training even further to see an effect of training condition on congruency. Future research could explore this further by running a ‘no training’ condition, where the pre-test instructions simply state the S-R relationship rather than running training trials. The ’long training’ condition could also be increased to three days of training (e.g., Tricomi et al.).

An additional point to consider relates to the importance of the long-delay test trials for measuring habitual behaviour. There is a short-delay effect whatever the level of training, which appears to be related to preparation of the response on test. However, if there is a long delay between Outcome and Stimulus on test, participants have plenty of time to prepare their response. If a long-delay congruence effect is observed in the long training group, this could be considered evidence of habitual behaviour. In line with this, Experiment 9 found a long-delay congruency effect in the long training group, but no congruency effect in long delay trials after a short amount of training. Which could be evidence of habitual behaviour. Therefore, it may be that this is where one should look to find a habit effect. The implications of these findings will be discussed further in Chapter 6.

Chapter 5 – Individual differences

In Chapters 2-4, we attempted to manipulate action slips by a number of factors commonly associated with habitual behaviour: O-R delay, cognitive load, time pressure, outcome value and amount of training. We found a reliable congruency effect throughout all our experiments. This congruency effect appears to be a direct measure of habitual behaviour. The degree to which people show a tendency towards habitual behaviour varies significantly between individuals (Ersche et al., 2017). Several self-reported trait measures (which will be introduced below) have previously been used to capture these individual differences. However, there are limitations to self-report measures, as they rely on an individual's ability to recognise their own behaviour (Sniehotta & Presseau, 2012). Therefore, comparing performance on our tasks with these self-reported measures may provide a more objective measure of habitual tendencies. This is important, as it allows a greater understanding of habit formation and can inform interventions designed to enhance beneficial habitual behaviour or reduce detrimental habitual behaviour.

Creature of Habit

The Creature of Habit Scale (COHS; Ersche et al., 2017) is a self-reported measure of habits in everyday life. The questionnaire is made up of 27 items to quantify individual differences in habitual tendencies, as shown in Appendix 2. When developing this questionnaire, a factor analysis revealed two distinct aspects: automaticity, which captures some of the autonomous nature of habitual behaviour, e.g., "I often find myself running on 'autopilot', and then wonder why I ended up in a particular place or doing something that I did not intend to do"; and routine, which refers to consistent scheduling of the behaviour, e.g., "I tend to do things in the same order every morning (e.g., get up, go to the toilet, have a coffee…)". Answers are given on a 5-point Likert scale ranging from 1 = "strongly disagree" to 5 = "strongly agree". The final scores for the two distinct subscales of automaticity (11 items) and routine (16 items) are calculated by averaging the respective item scores. As correlation between these two subscales was found to be relatively low (r=.154; Ersche et al.), they are not combined to an overall score, but used as separate measures of habitual behaviour. Routines involve a time-based sequence of actions that are performed consciously to make daily life more orderly and efficient (Clark, 2000). Therefore, ‘routine’ habits may be more linked with goal-directed actions, as they are consciously made to achieve a desired outcome. In contrast to routines, automaticity does not involve any conscious effort, or dependency of outcome (Saling & Phillips, 2007).

As the COHS is still a relatively new self-reported measure of habitual behaviour, there is limited existing research on the link between this self-reported measure and performance on experiments measuring habitual behaviour in instrumental learning tasks. Therefore, the initial aim of this chapter is to see whether our experiments assess habits directly and then see if that correlates with self-reported habitual behaviour.

Goal-pursuit

As well as measuring self-reported habitual behaviour using the COHS, we measured goal-oriented personality traits using the Habitual Self-Control Questionnaire (HSCQ; Schroder et al., 2013). The aim is to see whether there is an association between performance on our experiments and self-reported tendencies towards goal-directed behaviour. The HSCQ is a 14-item questionnaire measuring variations in people's commitment to completing tasks and their self-perceptions around their drive for goal-pursuit, as shown in Appendix 3. Participants express their level of agreement to each item on a 5-point Likert scale, ranging from disagree strongly (1) to agree strongly (5). Goal-striving personalities are likely to run counter to regularity and repetition (Dunn, 2000). In line with this, Ersche et al. (2017) found a negative correlation between COHS automaticity and HSCQ goal-directed control. Therefore, we expected to find the same negative correlation between these two measures in our experiments. Furthermore, if performance on our tasks positively correlate with COHS automaticity, we expected to find a negative correlation between HSCQ goal-directed control and performance on our tasks.

Impulsivity

Another consideration in our experiments is that performance on our tasks is linked to participants’ levels of self-reported impulsivity. With respect to the influence of impulsive traits on habit formation, there is growing evidence that self-reported impulsivity is associated with deficits in goal-directed control (Hogarth et al., 2012). Hogarth and colleagues explored this theory by assessing smokers for self-reported impulsivity and capacity for goal-directed control over instrumental performance in an outcome devaluation procedure. They found that reduced goal-directed control was selectively associated with the motor impulsivity factor of Barrett's Impulsivity Scale (BIS; Patton et al., 1995), which reflects tendency for action without thought. Participants who scored high in motor impulsivity had a reduced devaluation effect, indicating habitual behaviour. The BIS-11 is a 30-item self-report questionnaire used to measure impulsive behaviour, as shown in Appendix 4. Answers are given on a 4-point Likert scale (1 = "rarely/never", 2 = "occasionally", 3 = "often", 4 = "almost always/always"). Items include statements about impulsive behaviour, for example "I act on impulse", "I say things without thinking", and "I plan tasks carefully". The items are grouped into three subscales – motor impulsivity (assessing tendency for action without thought), non-planning impulsivity (assessing future orientation), and attentional impulsivity (assessing capacity for sustained attention). As this measure is commonly used in research for assessing impulsivity, we used the BIS-11 in our experiments to assess whether performance on our task was associated with self-reported impulsivity.

As well as association between self-reported impulsivity and performance of outcome devaluation procedures, there is existing research associating self-reported impulsivity with self-reported habitual behaviour. Ersche et al. (2019) found a positive association between self-reported impulsivity on the BIS, and automaticity measured using the COHS. Impulsivity was positively correlated with automaticity. However, impulsivity was negatively associated with routine behaviours. Impulsivity seems to selectively enhance automatic stimulus-driven, goal-independent actions whilst diminishing the occurrence of routine behaviours. Impulsive actions are spontaneous and have been described as an 'inability to control automatic reactions to stimuli' (Kopetz et al., 2018). Therefore, the positive relationship between impulsivity and stimulus-driven behaviours captured by the COHS automaticity scale may seem intuitive (Ersche et al.). However, the negative correlation between impulsivity and routine is less intuitive. It seems habits have two quite different aspects – one is about control/impulsivity and the other is about carefully doing the same thing repeatedly. In other words, is someone susceptible to habitual behaviour when it is inappropriate (impulsive S-R-like behaviour), versus, does someone tend to arrange their life in such a way that their habits bring about a desired outcome (routine)? As existing research has found an association between self-reported habitual behaviour and impulsivity, we also ran correlational analyses between impulsivity and self-reported habitual behaviour and habitual control measured in the COHS and HSCQ, and included these in Appendix 8.

Stress, depression, and anxiety

Several other individual differences have previously been associated with habitual behaviour. Experiences of acute or chronic stress has been identified as a factor that can influence the switch from goal-directed action to habitual behaviour (Dias-Ferreira et al., 2009; Schwabe et al., 2011). Moreover, Heller et al. (2018) found that on a two-stage decision-making task, individuals high in depression demonstrated greater habitual and less goal-directed decision-making in the face of stress. Ample evidence supports rumination as a vulnerability marker for the development and maintenance of depressive symptoms and episodes (Nolen-Hoeksema et al., 2008; Watkins, 2008). Consistent with the habit-goal framework, rumination has often been described as habitual in the depression literature (e.g., Hertel, 2004). Trait anxiety and experiences of adversity during childhood —a well-known risk factor for depression (Nelson et al., 2017), are associated with increased automatic habitual responding in everyday life, when habitual behaviour is measured using the COHS (Ersche et al., 2017). Furthermore, a study by Ólafsson et al. (2020) found that on an outcome devaluation task, stronger habitual (relative to goal-directed) behaviour was associated with a greater number of previous depressive episodes in a group of formerly depressed individuals. To summarise, stress, depression and anxiety have all been correlated with habitual behaviour. Therefore, a further aim of this chapter was to see if there were associations between performance in our task and self-reported levels of stress, depression and anxiety. We expected stress, depression, and anxiety to be associated with greater self-reported daily habits (measured using COHS automaticity) and greater action slips in our experiments.

We aimed to measure self-perceived stress in participants using the Perceived Stress Scale (PSS; Cohen et al., 1983). The PSS is a 10-item self-report measure to assess the degree to which people perceive their lives as stressful, as shown in Appendix 5. Items include questions such as “in the last month, how often have you felt nervous and stressed?". Response choices are ranked from "0" (never) to "4" (very often), providing a 0-40 severity score. Scores ranging from 0-13 are considered low stress, scores between 14-26 are considered moderate stress, and scores above 27 are considered high perceived stress.

We assessed participants' level of depression using the 9-item depression module of the Patient Health Questionnaire depression scale (PHQ-9; Kroenke et al., 2001), as shown in Appendix 6. This questionnaire is used to diagnose depression and grade the severity of symptoms in general medical and mental health settings. Scores of each of the nine Diagnostic and Statistical Manual of Mental Disorders criteria of major depressive disorder are ranked from "0" (not at all) to "3" (nearly every day), providing a 0–27 severity score. A higher PHQ-9 score indicates greater depressive tendencies and depression severity. The cut off for moderate depression is a PHQ-9 score of 10.

Level of anxiety was assessed using the 7-item Generalised Anxiety Scale (GAD-7; Spitzer et al., 2006). The scale asks participants how often in the past two weeks they have felt certain states of anxiety, such "feeling nervous, anxious, or on edge," and "trouble relaxing", as shown in Appendix 7. Response choices are "0" (not at all) to "3" (nearly every day), providing a 0–21 severity score. The cut off for moderate anxiety is a GAD-7 score of 10.

To summarise, the current chapter aims to explore whether there is an association between performance in our experimental task and self-reported measures of habitual behaviour and goal-directed control (as potential individual traits). There is also the possibility that instead of measuring habitual behaviour, our experiments are measuring errors that are due to something else, such as impulsivity. Therefore, this chapter also aims to explore whether there is an association between performance in our experiments and self-reported impulsivity. Furthermore, we explored associations between performance in our experiments and self-reported measures of several other individual differences associated with habitual behaviour (levels of stress, depression, and anxiety).

Method

Each self-report measure was presented to participants on-screen and completed directly before one of the experiments outlined in Chapters 2-4. For experiments run online, the questionnaire was built into the experiment. If the experiment was run in the lab using E-Prime, the questionnaire was developed and presenting using Qualtrics on a web browser directly before running the E-Prime experiment.

Exclusions. The same exclusion criteria were applied to all self-report measures. As in other chapters, we excluded participants with a mean accuracy of less than 80% across all trials on the experiment. Participants were also excluded if they failed an attention check embedded within the questionnaire. In order to test whether participants were paying attention to the questions, an additional question was added to the regular set of questions within the measure, asking the participant to select a specific answer. Finally, participants were excluded if their score was considered implausible, based on the scoring guidelines of the measure (i.e., we followed Patton et al.'s (1995) guidance for removing scores of 52 or less on the BIS-11). The number of data points vary between each analysis, as not all experiments contained all the questionnaires. This means that the number of participants for each correlation vary depending on whether the participants completed both measures for each correlation.

Self-reported habitual behaviour (COHS). The Creature of Habit Scale (COHS; Ersche et al., 2017) was completed by a total of 459 participants. As above, the exclusion criteria were applied, leaving a sample of 313 participants with COHS scores. The mean COHS score of the remaining sample was 66.86 (SD=13.47). The structure of the COHS contains two subscales: routine and automaticity. The mean routine score was 41.87 (SD=8.97) and the mean automaticity score was 24.99 (SD=8.12).

Self-reported goal-directed behaviour (HSCQ). The Habitual Self Control Questionnaire (HSCQ, Schroder et al., 2013) was completed by a total of 414 participants. Following exclusions, this left a sample of 283 participants with HSCQ scores, with a mean HSCQ score of 42.99 (SD=7.84).

Self-reported impulsivity (BIS). The Barratt Impulsivity Scale Version 11 (BIS-11; Patton et al., 1995) was completed by a total of 793 participants. As detailed above, participants were excluded for scoring less than 80% accuracy in the experiment. Participants were also excluded for failing the attention check in the questionnaire. BIS-11 guidelines specify an additional reason for exclusion: scores below 52 are considered to represent inaccurate responding (Stanford et al., 2009). Scores between 52 and 71 represent normal levels of impulsivity, with scores above 71 considered high. Of the sample remaining after our own exclusion criteria were applied, 56 participants scored less than 52 and were therefore excluded based on implausible BIS scores. This left a sample of 547 participants. The mean BIS score of the remaining sample was 66.38 (SD=8.95), with 149 participants scoring over 71 (27.2%).

Self-reported stress (PSS). The Perceived Stress Scale (PSS-10; Cohen et al., 1983) was completed by a total of 576 participants. Once the usual exclusion criteria were applied, the remaining sample of 405 participants had a mean PSS score of 21.57 (SD=6.45). Forty-six participants scored between 0 and 13 (low stress), 265 scored between 14 and 26 (moderate stress), and 94 scored between 27 and 40 (high stress).

Self-reported depression (PHQ). The Patient Health Questionnaire 9-item depression scale (PHQ-9; Kroenke et al., 2001) was completed by a total of 352 participants. After exclusions, there were 229 participants with PHQ-9 scores, with a mean PHQ-9 score of 9.20 (SD=6.27).

Self-reported anxiety (GAD). The Generalised Anxiety Disorder 7-item scale (GAD-7; Spitzer et al., 2006) was completed by a total of 352 participants. After exclusions, the remaining number of participants with GAD-7 scores was 229, with a mean GAD-7 score of 7.59 (SD=5.49).

Measures of task performance. To obtain a measure of participants' responses that were controlled by the stimulus presented, an action slip score was first calculated for each participant. This action slip score was calculated by subtracting a participant's percentage of accuracy on incongruent trials from their percentage of accuracy on congruent trials. A positive action slip score thus revealed a decrease in response accuracy for incongruent trials – suggesting the stimulus controlled their responses. To represent the difference in mean reaction times for congruent and incongruent trials, an 'RT cost' score was also calculated for each participant by subtracting their mean RT on congruent trials from their mean RT on incongruent trials. Four scores from the experiments were used for the correlation analyses: action slip score, overall mean accuracy (including both congruent and incongruent trials), RT cost score, and overall mean RT (including both congruent and incongruent trials).

Results

It is worth noting, we only include scatter plots for significant correlations between performance in our experiment and individual differences measures throughout this chapter. Further significant correlations between the COHS, HSCQ, BIS, PHQ-9 and GAD-7 scores are shown in Appendix 8.

COHS and HSCQ correlations with performance in our experimental task

Table 4 shows the first analysis, which correlated the measures derived from our experiments (action slip score, mean accuracy, RT cost score, and mean RT) with self-reported habitual behaviour (based on total COHS score, and the two subscales, routine behaviour and automaticity) and self-reported goal-directed behaviour (based on total HSCQ score) using Pearson's correlation and Bayesian correlation tests.

Table 4: Results of COHS and HSCQ Pearson's correlation analyses with performance in Experiments 1-9

Pearson's Correlations

Variable			Action slip score	Mean Accuracy (%)	RT cost score	Mean RT (s)
COHS	Routine	N	313	313	313	313
		Pearson's r	-0.01	0.08	-0.01	0.07
		p-value	0.977	.167	.850	.207
		BF_{10}	0.07	0.18	0.07	0.16
	Automaticity	N	313	313	313	313
		Pearson's r	0.03	-0.03	0.01	0.12
		p-value	0.586	.651	.828	.034*
		BF_{10}	0.08	0.08	0.07	0.67
HSCQ	Total HSCQ score	N	283	283	283	283
		Pearson's r	-0.02	0.01	-0.05	0.01
		p-value	0.699	.885	.421	.872
		BF_{10}	0.08	0.08	0.10	0.08

* p < .05, ** p <.01, *** p <.001
† BF_{10} = 3-10, †† BF_{10} = 10-30, ††† BF_{10} = 30-100, †††† BF_{10} > 100

COHS and HSCQ scores and task performance. As shown in Table 4, there were no significant correlations between COHS scores and action slip scores, mean accuracy, or RT cost scores. As shown in Figure 37, there was a marginally significant positive correlation between mean RT and COHS automaticity, r=.12, p=.034, BF_{10} = 0.66, with participants responding more slowly in the experiment when they scored higher on the automaticity measure. However, this correlation was weak, and the Bayesian correlation was inconclusive. There were no significant correlations between COHS routine scores and task performance, nor HSCQ scores and performance in our experiments.

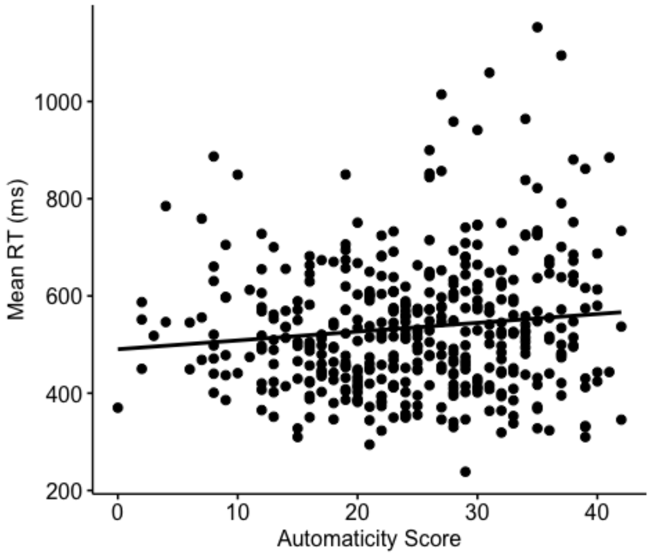


Figure 37: Scatter plot showing results of Pearson's correlation analysis between COHS automaticity score and overall mean RT (ms) in Experiments 1-9.

BIS correlations with performance in our experimental task

Table 5 shows the results of correlation analyses of BIS impulsivity scores and action slip score, mean accuracy, RT cost score and mean RT. As shown in Table 5, there were no significant correlations between impulsivity and action slip score or RT cost score.

Table 5: Results of BIS Pearson's correlation analyses with performance in Experiments 1-9

Pearson's Correlations

Variable			Action slip score	Mean Accuracy (%)	RT cost score	Mean RT (s)
BIS-11	Total BIS score	N	547	547	547	547
		Pearson's r	0.05	-0.13	-0.01	0.10
		p-value	.293	.002**	.962	.023*
		BF$_{10}$	0.09	6.33†	0.05	0.71
	Attentional score	N	547	547	547	547
		Pearson's r	-0.02	-0.02	-0.02	0.06
		p-value	.729	.603	.613	.141
		BF$_{10}$	0.06	0.06	0.06	0.16
	Motor score	N	547	547	547	547
		Pearson's r	0.05	-0.12	0.04	0.07
		p-value	.238	.005**	.344	.095
		BF$_{10}$	0.11	2.73	0.08	0.22

Non-planning score	N	547	547	547	547
	Pearson’s r	0.06	-0.15	-0.02	.09
	p-value	.133	<.001***	.596	.047*
	BF_{10}	0.17	33.75†††	0.06	0.38

* p < .05, ** p <.01, *** p <.001

† $BF_{10} = 3$-10, †† $BF_{10} = 10$-30, ††† $BF_{10} = 30$-100, †††† $BF_{10} > 100$

Total BIS scores and task accuracy. Table 5 shows that the correlation analysis between overall mean accuracy and total BIS impulsivity score revealed a weak negative correlation, r=-.13, p=.002, $BF_{10} = 6.33$, as illustrated in Figure 38, where impulsive participants made more overall errors in the tasks.

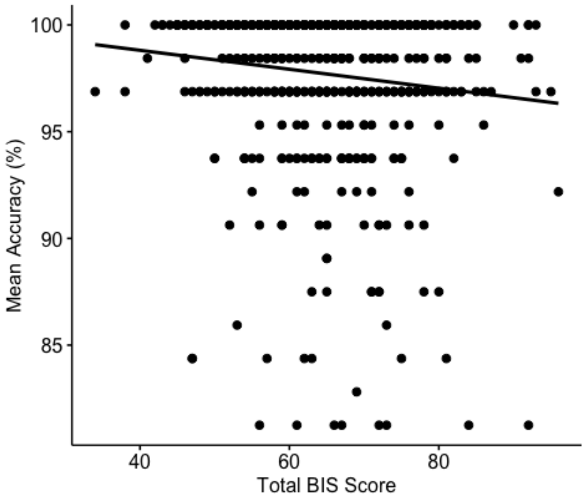


Figure 38: Scatter plot showing results of Pearson’s correlation analysis between mean accuracy (%) in Experiments 1-9 and total BIS score.

BIS subscale scores and task accuracy. Figure 39a shows overall accuracy was negatively correlated with motor impulsivity, r=-.12, p=.005, $BF_{10} = 2.73$. Overall accuracy was also negatively correlated with non-planning impulsivity, r=-.15, p<.001, $BF_{10} = 33.75$, as shown in Figure 39b. Those who scored higher on motor impulsivity or non-planning impulsivity made slightly more errors. However, these correlations were weak and the Bayesian correlation for motor impulsivity was inconclusive. There was no significant correlation between mean accuracy and attentional impulsivity.

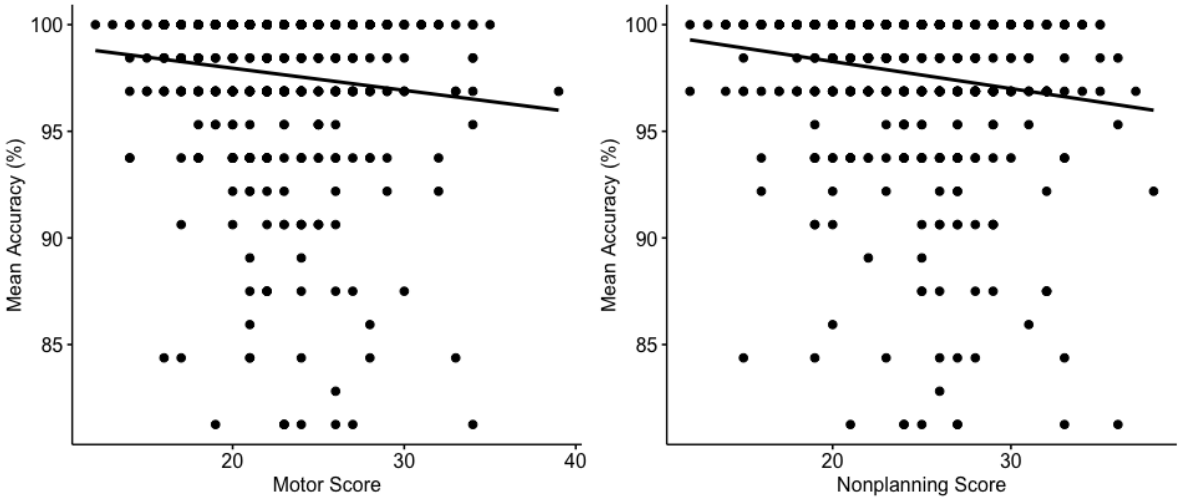


Figure 39: Scatter plot showing results of Pearson's correlation analysis between mean accuracy (%) in Experiments 1-9 and BIS subscale scores: a) motor impulsivity (bottom left); b) non-planning impulsivity (bottom right).

BIS scores and task RTs. The Pearson's correlations between mean RT and impulsivity showed a weak positive correlation between mean RT and total BIS score, r=.10, p=.023, as shown in Figure 40a, and a marginally significant positive correlation between mean RT and non-planning impulsivity, r=.09, p=.047, as shown in Figure 40b. Participants who scored higher on total BIS impulsivity and non-planning impulsivity responded marginally slower in our experiments. However, the Bayesian correlations did not support these findings for total BIS score, BF_{10} = 0.71, nor non-planning impulsivity, BF_{10} = 0.38. There were no significant correlations between mean RT and the other BIS subscales (attentional impulsivity and motor impulsivity).

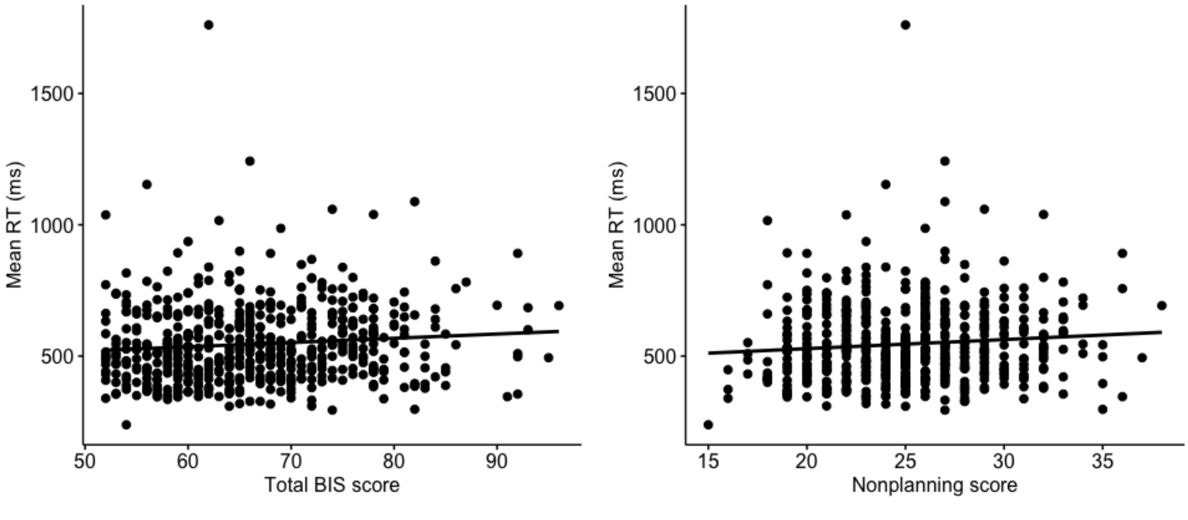


Figure 40: Scatter plot showing results of Pearson's correlation analysis between mean RT (ms) in Experiments 1-9 and the following BIS impulsivity scores: a) total BIS score (left); b) non-planning impulsivity score (right).

PSS, PHQ and GAD correlations with performance in our experimental tasks

As habitual behaviour has previously been associated with stress (PSS), depression (PHQ), and anxiety (GAD), we ran these correlations with performance on our own experiments to see if there were any associations between these self-reported individual differences and our measure of habitual behaviour. As shown in Table 6, there were no significant correlations between action slip score, overall mean accuracy, or RT cost score with any of these individual differences measures we explored (perceived stress, depression, and anxiety).

Table 6: Results of PSS, PHQ and GAD Pearson's correlation analyses with performance in Experiments 1-9

Pearson's Correlations

Variable			Action slip score	Mean Accuracy (%)	RT cost score	Mean RT (s)
PSS	Total PSS score	N	405	405	405	405
		Pearson's r	0.02	0.01	0.01	0.06
		p-value	0.728	.780	.933	.203
		BF_{10}	0.07	0.07	0.06	0.14
PHQ-9	Total PHQ score	N	229	229	229	229
		Pearson's r	-0.04	0.01	0.02	0.20
		p-value	0.509	.972	.787	.003**
		BF_{10}	0.10	0.08	0.09	$6.54^{\dagger}$
GAD-7	Total GAD score	N	229	229	229	229
		Pearson's r	0.05	-0.09	0.09	0.26
		p-value	0.413	.177	.184	<.001***
		BF_{10}	0.12	0.20	0.20	$> 100^{\dagger\dagger\dagger\dagger}$

* p < .05, ** p <.01, *** p <.001

† $BF_{10} = 3\text{-}10$, †† $BF_{10} = 10\text{-}30$, ††† $BF_{10} = 30\text{-}100$, †††† $BF_{10} > 100$

PSS, PHQ and GAD scores and task RTs. Both depression (PHQ score), r=.20, p=.003, $BF_{10} = 6.54$, as shown in Figure 41a, and anxiety (GAD score), r=.26, p<.001, $BF_{10} > 100$, as shown in Figure 41b, showed significant positive correlations with mean RT. Participants who scored higher on depression and anxiety tended to respond more slowly in our experiments. The Bayesian correlation showed particularly strong support for anxiety's correlation with mean RT. There was no significant correlation between perceived stress score (PSS) and mean RT.

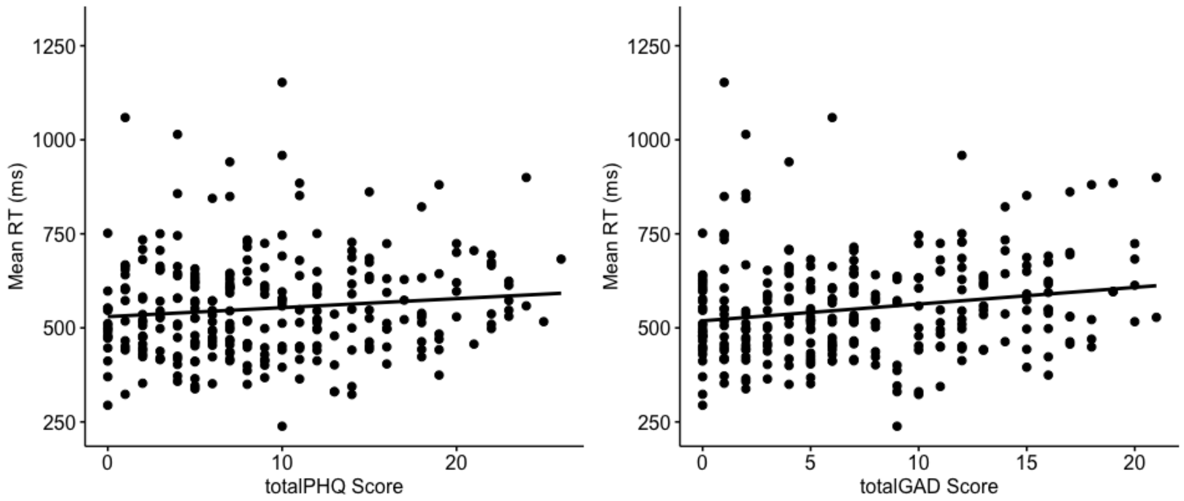


Figure 41: Scatter plot showing results of Pearson's correlation analysis between mean RT (ms) in Experiments 1-9 and the following individual differences measures: a) depression (PHQ-9) score (left); b) anxiety (GAD-7) score (right).

Discussion

The current chapter aimed to explore whether there are associations between performance in our experiments and self-reported measures of habitual behaviour and goal-directed control (as potential individual traits). We also explored whether performance on our experimental task was associated with self-reported impulsivity, stress, depression, and anxiety.

Our results found no association between self-reported habitual behaviour (COHS scores) and action slip score, nor RT cost score. As our action slip score was a representation of the congruency effect found in our experiments, these results suggest that the congruency effect is not associated with self-reported habitual personalities. We also found no association between self-reported goal-directed control (HSCQ score) and performance in our experiments. If anything, we found that higher scores in COHS automaticity led to slower responding across trials, regardless of trial congruency. This suggests that individuals with higher automatic habitual traits may have had to slow down their responding to ensure accuracy across trials (regardless of trial congruency). However, this correlation was very weak, and the Bayesian correlation was inconclusive, so these results should be treated with caution.

One significant finding was that there was a moderate negative correlation between COHS automaticity score and HSCQ score, with very strong support from the Bayesian correlation, as shown in Figure 44a in Appendix 8. This, therefore, replicates the findings from Ersche et al. (2017)'s correlations - that goal-striving personalities are likely to run counter to automatic, habitual personalities. A potential reason behind this is that goal-striving individuals manage their behaviour by considering the consequences of their actions. Individuals with automatic, habitual personalities, on the other hand, may be more inclined to allow environmental stimuli to take over their control.

It is important to note that using self-report measurements of behaviours that largely occur without awareness can be criticised (Sniehotta & Presseau, 2012). COHS Automaticity is the measure of an individual's tendency towards automatic, involuntary responding. However, as the measure is self-reported, the automaticity score is reliant on participants' awareness of their automatic behaviours. This poses a problem if automaticity is unconscious behaviour. In fact, you might expect the opposite correlation, where those people who are especially habitual give a low score on COHS automaticity because they are unaware of their automatic behaviour. Furthermore, given that automatic actions are dependent on the context that triggers them, automatic action patterns necessarily vary enormously with respect to individuals' lifestyles. This poses a particular challenge for self-report measures, which require questionnaire items that capture the specificity of the environment while applying to many people. As the consumption of food is largely automatic (Cohen & Farley, 2008), it is not surprising that many COHS automaticity items are related to eating. However, participants may have automatic habitual tendencies, but not in a food-related way. Therefore, they may score low on COHS automaticity simply because it does not capture the context of their automatic behaviour. Other non-food related behaviours are equally likely to be automatic but may be more difficult to assess by questionnaire (Ersche et al. 2017).

Another potential reason that we did not find an association between self-reported habitual traits and action slips scores is the possibility that, instead of measuring habitual behaviour, our experiments are measuring errors that are due to something else, such as impulsivity. We found a negative correlation between mean accuracy and non-planning impulsivity. Participants with more accurate overall performance in our experiments tended to score lower on the non-planning impulsivity measure. A possible explanation for this is that those with high non-planning impulsivity scores were likely to respond automatically to the stimuli, without planning their responses. On the other hand, those with low non-planning impulsivity scores were more likely to think about their responses in advance. Therefore, participants with low non-planning impulsivity scores performed more accurately across the experimental trials. This could also be interpreted as a form of goal-directed control.

To summarise, participants with high non-planning impulsivity were more likely to have low goal-pursuit traits, as shown in Figure 43d in Appendix 8 and high automatic habitual traits, as shown in Figure 42d in Appendix 8, and these participants tended to make more errors in the task in our experiments. This suggests that high accuracy in our experiments could be linked to goal-directed control, rather than automatic habitual behaviour. This also provides evidence for the association between impulsivity and performance in our experiments.

In addition to the correlations described above, is also worth discussing the attentional impulsivity correlations. This BIS subscale was negatively correlated with goal-directed control (HSCQ score), and positively correlated with automatic habitual behaviour (COHS automatic score). These results are shown in Appendix 8 and replicate Ersche et al.'s (2019) findings. However, attentional impulsivity did not correlate with performance in our experiments. Whilst we may expect there to be an association between attentional impulsivity and action slips, as self-reported automatic habitual behaviour is associated with high attentional impulsivity, our findings replicate those of Hogarth et al. (2012). In their devaluation experiment, they found no difference in devaluation effect between groups split by attentional impulsivity score. Individuals who struggle to pay attention to a task may also struggle with goal-directed control in real-life situations. However, the tasks in our experiments were relatively easy and short-lived compared to more complex real-life goals, such as losing weight or quitting smoking. Therefore, even participants with high attentional impulsivity may have found it easy to sustain attention in our experiments.

As well as investigating associations between performance in our task and several relevant behavioural traits (habitual behaviour, goal-pursuit and impulsivity), we explored a number of clinical states that have previously been associated with habitual behaviour. Experiences of acute or chronic stress has previously been identified as a factor that can influence the switch from goal-directed action to habitual behaviour (Dias-Ferreira et al., 2009; Schwabe et al., 2011). In line with this, our results found a moderate positive correlation between self-reported automatic habitual behaviour (COHS automaticity) and stress (PSS), where those who reported higher rates of stress also reported higher automatic habitual behaviour, as shown Figure 44b in Appendix 8. These results are in line with existing findings that stress can influence habitual behaviour (Dias-Ferreira et al.; Schwabe & Wolf). Therefore, we expected to find a positive correlation between self-reported perceived stress and action slip scores in our experiments. This was not the case.

As already discussed, it is possible that the action slips in our experiments were due to two competing goals, rather than S-R behaviour. This could explain why there was no association between perceived stress and action slip scores. It also possible that, whilst participants scoring high on the perceived stress scale (PSS) believe they have high levels of stress in general life, they may not have been feeling stressed at the precise time they completed the experiment. Therefore, their general levels of stress did not impact their performance in our experiments. Future research could apply a stress manipulation to our experimental procedure to measure the effect of stress applied at the time of the experiment. An example of existing research that has successfully induced stress in a laboratory setting is Le et al.'s (2021) ‘Simple Singing Stress Procedure’. In this procedure, participants in the stress condition were instructed to sing a song in front of the experimenter while being video- and audio-recorded. They were also told they would have to sing the song again at the end of the experiment whilst being assessed by a panel. Participants in the no-stress condition simply read the song lyrics in each test phase. They found that participants in the stress condition showed significantly higher blood pressure and salivary cortisol immediately after the initial singing session, compared to those in the no-stress condition. Therefore, applying the Simple Singing Stress Procedure during the training phases of our experiments, with the anticipation of singing again after the test phase to a panel of experimenters, could be a simple method of exploring the effects of induced stress on action slips.

A further investigation of this chapter was to explore the association between performance in our experiments and depression and anxiety. A study by Ólafsson et al. (2020) found that on an outcome devaluation task, stronger habit relative to goal-directed behaviour was associated with a greater number of previous depressive episodes in a group of formerly depressed individuals. Trait anxiety has also previously been associated with increased automatic habitual responding in everyday life, where habitual behaviour was measured using the COHS (Ersche et al., 2017). In line with this, our results found a moderate positive correlation between self-reported automatic habitual behaviour (COHS automaticity) and depression and anxiety, as shown in Appendix 8 Figures 44c and 44d, where those who reported higher automatic habitual behaviour also reported higher rates of depression or anxiety. Therefore, we expected to find an association between performance in our experiments and both self-reported depression scores (PHQ-9) and generalised anxiety scores (GAD-7). However, there were no correlations between action slip scores in our experiments and depression or anxiety.

One significant clinical finding in our experimental tasks was a positive correlation between overall RT in our experimental test trials and self-reported levels of depression and anxiety, with participants responding slower when they had also reported higher levels of depressive and anxious states. These findings are in line with existing evidence that depressed individuals perform more slowly on reaction time tasks than participants in a control group (Azorin et al., 1995). However, existing evidence for the effect of anxiety on RTs has been mixed and tends to depend on the motivation behind RTs (Farber & Spence, 1956). As performance was generally high across our experiments, these results indicate that participants with high levels of depressive or anxious states slow their responses to perform accurately in our experiments (regardless of trial congruency). This sounds like goal-directed behaviour, rather than habitual behaviour, which could explain why there was no association between action slip scores and depression, nor anxiety.

Overall, our results found no evidence for an association between performance in our experiments and self-reported measures of habitual behaviour, nor self-reported goal-directed control. However, we found evidence of an association between impulsivity and overall performance in our experiments, where low overall accuracy in our experiments is associated with participants who do not tend to plan their responses in advance (high non-planning BIS score). This could also be interpreted as a lack of goal-directed control - an interpretation that is supported by our correlation analysis, showing that low scores in goal-directed control (HSCQ score) were associated with high non-planning impulsivity (BIS score), as illustrated in Figure 43d in Appendix 8 These results suggest that performance in our experiments could be linked to goal-directed control, rather than automatic habitual behaviour, as participants' accuracy in our experimental tasks appear to depend on planning their responses in advance. Other individual differences (depression and anxiety) were also associated with self-reported habitual behaviour and goal-directed control but were only related to mean RT across all trials in our experiments, as shown in Table 6, indicating that these participants slowed their responses in order to respond accurately, which also sounds like goal-directed control. People who are anxious tend to slow their responses to ensure they respond correctly, which is in line with previous findings (Mkrtchian et al., 2017).

It is important to remember that correlations between these factors do not imply causation. Habitual behaviour and impulsivity may be related, but it is essential to interpret the findings in the context of the specific study design and research methods used. These results will be discussed further in Chapter 6.

Chapter 6 – General discussion

The experiments in this thesis aimed to investigate evidence of habitual behaviour and whether the amount of habitual behaviour observed varied with manipulations of response delay, working memory capacity, devaluation, and amount of training. We also aimed to explore whether individual differences were associated with performance in our experiments, including self-reported habitual behaviour, goal-directed control, impulsivity, perceived stress, depression, and anxiety.

Summary of results

A congruency effect was consistently found in our experiments (with one exception, which we will discuss below). Responding was more accurate when the presented stimulus was congruent with the available outcome (i.e., O1 followed by S1, or O2 followed by S2 on test) than when the stimulus was incongruent with the outcome (i.e., O1 followed by S2, or O2 followed by S1). These results indicate that participants' responses were influenced by the stimuli (S1 and S2) presented, even though participants were instructed explicitly to ignore those stimuli. This leads to the question of whether the stimuli presented on incongruent trials disrupted goal-directed behaviour.

In addition to the congruency manipulation, all three experiments in Chapter 2 implemented a manipulation in the delay between O and S on test. The short O-S delays had a tendency to increase the congruency effect. Although evidence for this effect of O-S delay was not strong, the numerical pattern was consistent across experiments, where shorter delays showed a bigger congruency effect on accuracy. In addition to the effect on accuracy, there was also an effect on RTs. In order to respond accurately, it took participants more time to respond on short delay trials than long delay trials. A possible interpretation of these findings is that, in short delay trials, the sudden presentation of the stimulus may not have allowed participants enough time to prepare their responses in advance, causing a longer response time. In long delay trials, on the other hand, participants had ample time to prepare their responses to obtain the presented outcome. They were then able to hold that response in mind until the stimulus was presented and make the prepared response quickly. These findings can be accommodated by dual-process theory by assuming that goal-directed control will be disrupted when there is less time to prepare responses in advance. This leads to a larger effect of the stimuli, as goal-directed behaviour is time consuming and effortful, in comparison with automatic habitual behaviour, which is fast and effortless (Heyes & Dickinson, 1990).

The main aim of Chapter 3 was to explore whether reducing participants' working memory capacity would increase the congruency effect. In addition to our standard congruency manipulation, we embedded a working memory task into half of the test trials in Experiments 4-6. This additional task was designed to create a distraction or increase cognitive load on half of the trials to see if this would increase the number of action slips made. The results suggest that a 'light' load task was not enough to interact with our congruency effect. When the digit-span load task was applied (which required a higher level of working memory), participants appear to slow down on incongruent trials during the more difficult load trials to ensure accuracy. When participants did not compensate in terms of speed (Experiment 6), a load x congruency effect on accuracy was observed. These results provided some evidence that the congruency effect can be increased by reducing working memory capacity, which again, supports the dual-process theory. It is not clear why the RT results differed when the load experiment was conducted online verses in-person. It is possible that the difference in speed of responding was due to the added pressure of performing the experiment in a lab with an experimenter present in Experiment 6.

Chapter 4 aimed to explore to what extent our congruency effects shown in Chapters 2 and 3 connect to the existing habit literature. We aimed to do this by exploring two approaches commonly used in existing literature to measure habitual behaviour verses goal-directed action: manipulation of outcome value, and manipulation of amount of training (Adams and Dickinson, 1981; de Wit et al., 2018; Luque et al., 2020; Tricomi et al., 2009). In Chapters 2 and 3, we consistently saw a congruency effect in our experiments. However, when we tried to manipulate devaluation (Experiment 8), this congruency effect was not observed. These results suggests that by putting more focus on the value of the outcome, participants were more able to focus on which outcome to respond for, resulted in high accuracy in both congruent and incongruent conditions. In other experiments, it is possible that the screen colour was more salient than the outcome, resulting in action slips. It is important to note that this lack of congruency effect only occurred in this one experiment. When we returned to the standard testing procedure of the previous chapters and manipulated the amount of training participants received (Experiment 9), we observed a main congruency effect, but we found no effect of increased training on accuracy. The congruency effect was still there, and undiminished, when we shortened training. The theoretical implications of these results will be discussed below.

In Chapter 5, as well as manipulating the congruency effect, we aimed to explore whether there is an association between performance in our experiments and self-reported measures of habitual behaviour and goal-directed control as potential individual traits. Additionally, we aimed to explore whether several self-reported individual differences (impulsivity, stress, depression, and anxiety) were associated with performance in our experiments. Our results found no association between self-reported habitual behaviour (COHS scores) and action slip score, nor RT cost score. This suggests that the congruency effect found in our experiments is not associated with self-reported habitual personalities. We also found no association between self-reported goal-directed control (HSCQ score) and action slips scores, nor RT cost score, overall accuracy, or overall RT. These are the results one might predict if our congruency effect is due to competing goals, rather than dual processes. The action slip score is formed from one goal (performance on congruent trials) subtracted from another goal (performance on incongruent trials), so it is possible that they cancel each other out.

There was strong evidence for a correlation between overall accuracy in our experiments (regardless of trial congruency) and non-planning impulsivity, with participants who do not tend to plan their responses in advance (high non-planning BIS scores) performing less accurately across the experimental trials. If they are right about their own behaviour when answering the BIS questions (they do not plan in advance), then these findings are consistent with Experiment 3's results. In Experiment 3, long delays led to higher accuracy. When participants had the opportunity to plan in advance, accuracy was higher. However, some people might not have taken that opportunity - because they have high non-planning impulsivity tendencies. This leads to the question of why the association between non-planning impulsivity and accuracy is seen on congruent trials no less than on incongruent trials (i.e., there is no correlation with the action slip score). It is possible that this comes down to the significance of the stimulus. In these experiments, S signifies nothing on test trials. So, on congruent trials, perhaps participants who have not planned their response are simply responding randomly when S appears, rather than responding for the outcome. Therefore, the association between performance in our experiments and non-planning impulsivity could be interpreted as a lack of goal-directed behaviour. In line with this, further correlation analyses shown in Appendix 8 found that non-planning impulsivity was positively associated with automatic habitual behaviour (COHS automaticity score) and inversely associated with goal-directed control (HSCQ score). These results indicate that people with high non-planning impulsivity may have lower goal-directed control, which may, in turn, have influenced performance in our experiments (regardless of congruency).

Theoretical implications

Dual-process theory proposes that instrumental behaviours are controlled by two distinct systems: the goal-directed system (facilitating deliberate, reward-driven actions) and the stimulus-triggered habitual system (producing automatic responses based on S-R associations). Therefore, if the dual-process account is correct, and S-R associations interfere with goal-directed control, we would expect to see a congruency effect in our experiments, where participants are less accurate on incongruent trials following interference from the trained S-R associations. As previously stated, we observed this congruency effect in all our experiments (with the one exception). A dual-process explanation of the action slips made on incongruent trials is that participants' goal-directed behaviour (e.g., for O1 by responding R1) is undermined by an incompatible response triggered by the presentation of an incompatible stimulus (e.g., S2 that is associated with R2). The proposed mechanism underlying the incompatible response is an S-R link formed during training. This link affects behaviour automatically – quickly and perhaps outside of the participant's awareness (Moors & De Houwer, 2006). Furthermore, the S-R mechanism is argued to be qualitatively different from the goal-directed system; it is evolutionarily older and has a distinct neural basis (e.g., Killcross & Coutureau, 2003). This S-R mechanism is the predominant explanation for action slips (Watson & de Wit, 2018). We also observed an increase in the congruency effect during short O-R delays and under high cognitive load. According to dual-process theory, goal-directed actions are effortful, in comparison with habitual S-R responses (Heyes & Dickinson, 1990). Therefore, on trials where responding for the goal (outcome) is incongruent with the S-R associations, participants would need more time to prepare their responses and more working memory capacity to avoid making an incorrect S-R response. Therefore, our O-S delay and cognitive load effects are consistent with the dual-process theory.

The main outcome of our results is that it appears habitual behaviour can be observed with humans using our basic instrumental experiments. We acknowledge that, by interpreting the congruency effect observed as evidence for habitual behaviour, we have somewhat “stretched” the definitions of habits. Insensitivity to outcome devaluation manipulations is typically regarded as the canonical measure of habitual control, at least in the animal learning literature (e.g., Adams & Dickinson, 1981). Still, there is a substantial but largely distinct literature that seeks to examine habits in humans from the ideomotor literature (e.g., Elsner & Hommel, 2001). In these studies, researchers often seek evidence of habits not through outcome devaluation manipulations, but rather through performance on reaction time and interference tasks, similar to that used here (e.g., Brass et al., 2001; see also Stroop, 1935). Gaining a clearer picture of how performance in our experimental procedure relates to devaluation manipulations was an important consideration we explored in Chapter 4. This link was required as interference tasks and devaluation tasks might measuring two different psychological processes.

It is important to consider why we did not see an effect of devaluation in Chapter 4. We used the same devaluation procedure as the Fabulous Fruit Task (de Wit et al., 2007; de Wit et al., 2013), which has previously been found to be an effective measure of habitual behaviour in lab-based experiments (although, see De Houwer et al.’s argument in Chapter 1). de Wit and colleagues found that participants will continue responding even when the outcome has been devalued, providing evidence for habitual behaviour. Therefore, if we are to believe the dual-process theory of habitual behaviour, we would expect to continue to see a congruency effect when our procedure is more in line with the Fabulous Fruit Task (Experiment 8). However, the devaluation procedure did not produce these results for us. Accuracy was high across all trials (around 98%), so it cannot be argued that the participants simply did not understand the instructions; participants clearly knew which outcome to respond to on each trial. Rather, our devaluation manipulation appeared to increase participants’ ability to hold the goal in mind or resist the influence of the nature of the stimulus S when it was presented. Rather than displaying both outcomes with a cross over the ‘incorrect’ outcome, our other experiments only displayed the ‘correct’ outcome at the beginning of each trial. Displaying the two outcomes (valued and devalued) simultaneously on screen before responding, rather than displaying the valued outcome alone, appeared to make participants almost immune to action slips. It is possible that indicating which outcome not to go for, as well as which outcome they should go for, made it easier for participants to make the correct response. In line with existing research, our devaluation results appear to provide evidence of goal-directed behaviour in humans, but not S-R responding.

A further common theme in previous literature on habitual behaviour is manipulating the amount of training participants receive (Adams & Dickinson, 1981; de Wit et al., 2018; Luque et al., 2020; Tricomi et al., 2009). As outlined in Chapter 4, a dual-process account would argue that more habitual behaviour will be observed when the amount of training is increased. However, we found no statistically significant difference between accuracy following short training and long training (Experiment 9). A congruency effect was observed even in the short training condition. As amount of training is considered fundamental in developing habitual behaviour, the fact that our congruency effect appears even after only a short amount of training suggests that we may be measuring something other than habitual behaviour in our experiments.

To summarise, in line with the dual process theory of habitual behaviour, our experiments almost consistently show a congruency effect, which can be manipulated through participants’ opportunity to prepare their goal-directed response (Experiments 1-3) and through decreasing their working memory capacity (Experiments 5-6). However, this congruency effect still appears even after a short amount of training (Experiment 9), which is inconsistent with the S-R account and dual process theory.

Methodological implications

Running experiments with participants face-to-face in a lab allows more control over the experimental conditions. Therefore, this would be the preferred method of conducting our experimental procedure. However, for the purposes of this thesis, we were forced to use the online method for some of the experiments due to social-distancing rules during the pandemic. Future experiments may benefit from following the face-to-face procedure, as we found a difference in RTs between participants in Experiment 5 (online) and Experiment 6 (in-person), possibly due to additional distractions in the online environment. Having said this, we still found our congruency effect in online experiments, indicating that it is possible to run this experiment online. Whilst face-to-face would be preferable, having an experiment that can run online has practicable benefits, such as reaching a wider range of the population and increasing sample size.

One small outstanding point relates to the training performance. One might expect accuracy to be close to 100% on such a simple task, rather than the 95% accuracy observed. Our only explanation for this poorer than expected performance is that participants were responding as quickly as possible to maximise points earned. This is not unusual for an instrumental training task. To some extent, participants can choose how accurate they want to be; faster responses earn more points, but incorrect responses earn zero points. Perhaps 5% errors were deemed acceptable in the pursuit of faster responses by our participants. The scoring system on test would have rewarded this adjustment; increasing reaction time by less than 100ms (the kind of increase we observed) would have led to a maximum reduction of a single point. In general, when accuracy is rewarded more than speed, participants are likely, if possible, to adapt to the kinds of interventions in place in our experiments to maintain high levels of accuracy. Perhaps a scoring system that heavily penalizes slightly slower responses would reveal stronger congruency effects under a O-S delay or cognitive load manipulation.

Behavioural and clinical research implications

As well as providing the existing literature with theoretical implications in the study of habitual behaviour, the research in this thesis has both behavioural and clinical implications. Understanding the mechanisms behind habitual behaviour relate heavily to behavioural change research, such as quitting smoking, increasing work productivity, and replacing unhealthy behaviour with more healthy alternatives (Ersche et al., 2016; Hogarth et al., 2003; Sebold et al., 2014; Voon et al., 2015). There are always competing response options in terms of food choices (unhealthy verses healthy), alcohol (alcoholic vs soft drinks), activity choices (watch tv vs go for run) etc., so our study could be used to measure habit formation in a variety of cases, where the outcomes in our experiment relate to these competing choices.

Our findings may also have implications for research into clinical conditions. As discussed in Chapter 5, chronic stress (Dias-Ferreira et al., 2009; Schwabe et al., 2011), depression (Heller et al., 2018) and trait anxiety (Ersche et al., 2017) have previously been associated with increased automatic habitual behaviour. Rumination is associated with stress (Nolen-Hoeksema et al., 1999), depression and anxiety (Watkins et al., 2009; Young & Dietrich, 2015), and rumination in terms of repetitive negative thoughts has been conceptualised as habitual behaviour (Hertel, 2004; Watkins & Nolen-Hoeksema, 2014). The analyses of our self-report measures, shown in Appendix 8, replicated these findings. We found positive correlations between self-reported automatic habitual behaviour (COHS scores) and self-reported stress (PSS scores), depression (PHQ-9 scores), and anxiety (GAD scores), as well as negative correlations between these three clinical traits and self-reported goal-directed control (HSCQ).

Although we found associations between self-reported habitual behaviour and the self-reported stress, depression and anxiety, we found no evidence for an association between action slip scores in our experiments and three clinical traits. A possible reason for this is that the self-report measures used in this thesis applied to levels of stress, depression, and anxiety in general life. For example, they may feel stressed in life generally but were not feeling high levels of stress at the time of the experiment. It is possible that a different study focusing on the current clinical state of the participants would find correlations between these clinical traits and performance in our experiments. For example, inducing stress at the time of the experiment by adding Le et al.'s (2021) 'Simple Singing Stress Procedure' to our task could be a simple method of exploring the effects of induced stress on action slips.

One significant clinical finding in our experimental tasks was that our correlational analyses found an association between overall RTs and levels of depression and anxiety. Participants with high levels of depression and anxiety tended to respond more slowly across all trials than participants with low levels of depression and anxiety. It would be beneficial to explore this further, as understanding the mechanisms behind habitual behaviour and goal-directed control may have implications for rumination research and could therefore play a role in developing interventions for treating depression and anxiety.

Future research

A key point in Chapter 6 is that we consistently saw a congruency effect in our experiments, except when we tried to manipulate devaluation (Experiment 8). One consideration for why participants' goal-directed control was not disrupted is that the devaluation manipulation made it too easy for them to keep the goal in mind. Therefore, they did not revert to habitual behaviour. To explore this further, and increase task complexity, future research could include our digit span load task (Experiments 5 and 6) in our devaluation procedure. This would allow us to see whether increasing task complexity and applying a manipulation that was previously found to effect congruency (Experiment 6) would result in a congruency effect when the outcome is devalued.

In general, future research could focus on exploring whether motivation plays a role in performance on our experimental tasks. Buabang et al., (2021) applied a variation of the Fabulous Fruit Task (de Wit et al., 2007) to explore the role of motivation in goal-directed control. Their primary finding was that performance after devaluation depended on participants' motivation to learn about the fruit outcomes. When motivation to learn about the fruit outcomes was not high enough, performance for the fruit outcomes was, consequently, reduced. Motivation was especially important for incongruent trials. They found that a lack of motivation led to the pattern of results that would usually be interpreted as evidence for habits, when in fact the behaviour was goal-directed. According to Buabang and colleagues, motivation to learn an association involves the individual deciding to engage in a set of mental actions (R), which depend on the value of achieving the outcome (O) and the expectancy that learning the association will lead to this outcome (R—O). In other words, if the outcome is not valuable enough to the individual, they may choose not to learn the association, which will result in lower R-O performance.

In order to explore the role of motivation further, a future study applying our experimental procedure could focus on increasing participants' motivation to win pringles/jellybean points. Motivation could be amplified by increasing the value of the points, e.g., by having a financial association to the points. The number of points won for each outcome could determine the value of a bonus payment awarded to the participant. In order to strengthen the motivation further, incorrect responses could result in a loss of points, and therefore a smaller bonus payment. Both outcomes could start with a high value of points. Then, on later trials, one of the outcomes would be devalued by associating it with negative points. If performance is goal-directed, we would expect participants to show greater accuracy and respond only for the outcome with value, resulting in no congruency effect. However, if participants make action slips, responding for the devalued outcome on incongruent trials, this could be considered evidence for habitual behaviour. Just like our existing devaluation experiment (Experiment 8), there is no personal incentive for making correct responses. Therefore, incorrect responses are not necessarily evidence for habitual behaviour, as participants could have chosen to respond to the stimulus (S) on some trials, as this was easier than holding R-O in mind during the O-S delay, resulting in action slips. However, unlike Experiment 8, action slips in the proposed future experiment result in a direct financial loss to the participant. Therefore, they are likely to be more motivated to consistently respond for the O and ignore the S. In Experiment 8, participants received no physical reward at the end of the Experiment 8. Therefore, the incentive for R-O accuracy may not have been high enough to motivate them.

Another key point in Chapter 6 was that we found no effect of reduced training on action slips (Experiment 9). Extensive training is considered critical for developing habitual responses (Tricomi et al., 2009). However, the congruency effect was still there in our experiment when we applied a shortened version of training for half the participants (although it was weaker). In line with existing research on habitual behaviour (Adams & Dickinson, 1981; de Wit et al., 2018; Luque et al., 2020), we expected the congruency effect to significantly reduce for the participants in the ‘short training’ condition. A potential future experiment could remove training altogether for half of the participants and simply state the S-R relationships at the beginning of the test phase as a verbal instruction. Conversely, the long training condition for the other half of participants could also be extended to three days of training, as this has been a method previously used in habit research (e.g., Tricomi et al.).

Conclusion

To conclude, the current thesis presents a new and very simple method to demonstrate habitual behaviour. The evidence for habits comes, not from the absence of an effect of outcome devaluation, but from the presence of stimuli that promote a different response from that which is consistent with the overall goal. Our hope is that this procedure might be helpful in future studies investigating the nature of habitual behaviour.

References

Adams, C. D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B, 34*(2), 77–98. https://doi.org/10.1080/14640748208400878

Adams, C. D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *The Quarterly Journal of Experimental Psychology Section B, 33*(2), 109–121. https://doi.org/10.1080/14640748108400816

Anderson, B. A. (2018). Controlled information processing, automaticity, and the burden of proof. *Psychonomic Bulletin & Review*, *25*(5), 1814-1823. https://doi.org/10.3758/s13423-017-1412-7

Augustinova, M., & Ferrand, L. (2014). Automaticity of word reading: Evidence from the semantic Stroop paradigm. *Current Directions in Psychological Science*, *23*(5), 343–348. https://doi.org/10.1177/0963721414540169

Azorin, J. M., Benhaïm, P., Hasbroucq, T., & Possamaï, C. A. (1995). Stimulus preprocessing and response selection in depression: A reaction time study. *Acta Psychologica*, *89*(2), 95–100. https://doi.org/10.1016/0001-6918(94)00024-B

Baguley, T. (2012). Calculating and graphing within-subject confidence intervals for ANOVA. *Behaviour Research Methods*, *44*, 158–175. https://doi.org/10.3758/s13428-011-0123-7

Bell, I. H, Marx, W., Nguyen, K., Grace, S., Gleeson, J., Alvarez-Jimenez, M. (2023). The effect of psychological treatment on repetitive negative thinking in youth depression and anxiety: a meta-analysis and meta-regression. *Psychological Medicine*, *53*(1), 6-16. https://doi.org/10.1017/S0033291722003373

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*(1–2), 3–22. https://doi.org/10.1016/S0001-6918(00)00024-X

Buabang, E. K., Boddez, Y., De Houwer, J., & Moors, A. (2021). Don’t make a habit out of it: Impaired learning conditions can make goal-directed behaviour seem habitual. *Motivation*

Science, *7*(3), 252–263. https://doi.org/10.54239/2319-022-001-001

Buabang, E. K., Köster, M., Boddez, Y., Van Dessel, P., De Houwer, J., & Moors, A. (2023). A goal-directed account of action slips: The reliance on old contingencies. *Journal of Experimental Psychology: General*, *152*(2), 496–508. https://psycnet.apa.org/doi/10.1037/xge0001280

Clark, F. A. (2000). The concepts of habit and routine: A preliminary theoretical synthesis. *The Occupational Therapy Journal of Research*, *20*(2000), 123S-137S. https://doi.org/10.1177/15394492000200S114

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behaviour*, *24*(4), 385-396. https://doi.org/10.2307/2136404

Cohen, D. A., & Farley, T. A. (2008). Eating as an automatic behaviour. *Preventing Chronic Disease*, *5*(1), 1–7. http://www.cdc.gov/pcd/issues/2008/jan/ 07_0046.htm.

Colwill, R. M., & Rescorla, R. A. (1988). Associations between the discriminative stimulus and the reinforcer in instrumental learning. *Journal of Experimental Psychology: Animal Behaviour Processes*, *14*(2), 155–164. http://doi.org/10.1037/0097-7403.14.2.155

Corbit, L. H. (2018). Understanding the balance between goal-directed and habitual behavioural control. *Current Opinion in Behavioural Sciences*, *20*(2), 161–168. https://doi.org/10.1016/j.cobeha.2018.01.010https://doi.org/10.1016/j.cobeha.2018.01.010

Corbit, L. H., & Balleine, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *Journal of Neuroscience*, *25*(4), 962–970. http://doi.org/10.1523/JNEUROSCI.4507-04.2005

Corbit, L. H., Janak, P. H., & Balleine, B. W. (2007). General and outcome-specific forms of Pavlovian-instrumental transfer: the effect of shifts in motivational state and inactivation of the ventral tegmental area. *European Journal of Neuroscience*, *26*(11), 3141–3149. http://doi.org/10.1111/j.1460-9568.2007.05934.x

De Houwer, J. Tanaka, A., Moors, A., & Tibboel, H. (2018). *Kicking the habit: Why evidence for habits in humans might be overestimated. Motivation Science, 4 (1), 50–59.* https://psycnet.apa.org/doi/10.1037/mot0000065

De Houwer, J. & Beckers, T. (2003). Secondary task difficulty modulates forward blocking in human contingency learning. *Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology*, *56*(4), 345–357. https://doi.org/10.1080/02724990244000296

de Wit, S., Niry, D., Wariyar, R., Aitken, M. R. F., & Dickinson, A. (2007). Stimulus-outcome interactions during instrumental discrimination learning by rats and humans. *Journal of Experimental Psychology: Animal Behaviour Processes*, *33*(1), 1–11. https://doi.org/10.1037/0097-7403.33.1.1

de Wit, S., Barker, R. A., Dickinson, A. D., & Cools, R. (2011). Habitual versus goal-directed action control in Parkinson disease. *Journal of Cognitive Neuroscience*, *23*(5), 1218–1229. https://doi.org/10.1162/jocn.2010.21514

de Wit, S. Corlett, P. R., Aitken, M. R., Dickinson, A., & Fletcher, P. C. (2009). Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behaviour toward food pictures in humans. *Journal of Neuroscience*, *29*(36), 11330–11338. https://doi.org/10.1523/JNEUROSCI.1639-09.2009

de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: A case for animal-human translational models. *Psychological Research*, *73*(4), 463–476. https://doi.org/10.1007/s00426-009-0230-6

de Wit, S., Kindt, M., Knot, S. L., Verhoeven, A. A. C., Robbins, T. W., Gasull-Camos, J., Evans, M., Mirza, H., & Gillan, C. M. (2018). Shifting the balance between goals and habits: Five failures in experimental habit induction. *Journal of Experimental Psychology: General*, *147*(7), 1043. https://doi.org/10.1037/XGE0000402

de Wit, S., Ridderinkhof, K. R., Fletcher, P. C., & Dickinson, A. (2013). Resolution of outcome-induced response conflict by humans after extended training. *Psychological Research*, *77*(6), 780–793. https://doi.org/10.1007/s00426-012-0467-3

de Wit, S., Watson, P., Harsay, H. A., Cohen, M. X., van de Vijver, I., & Ridderinkhof, K. R. (2012). Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control. *Journal of Neuroscience, 32*(35), 12066–12075. https://doi.org/10.1523/JNEUROSCI.1088-12.2012

Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., Costa, R. M.,

& Sousa, N. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. *Science*, *325*(5940), 621–625. https://doi.org/10.1126/science.1171203

Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *308*(1135), 67–78. https://doi.org/https://doi.org/10.1098/rstb.1985.0010

Dickinson, A., & Balleine, B. (1994). Motivational control of goal-directed action. *Animal Learning & Behaviour*, *22*(1), 1–18. https://doi.org/10.3758/BF03199951

Dickinson, A., Balleine, B., Watt, A., Gonzalez, F., & Boakes, R. A. (1995). Motivational control after extended instrumental training. *Animal Learning & Behaviour*, *23*, 197–206. https://doi.org/10.3758/BF03199935

Dunn, W. W. (2000). Habit: What's the Brain Got to Do with It? *The Occupational Therapy Journal of Research*, *20*, 6S-20S. https://doi.org/10.1177/15394492000200S102

Ehring, T., & Watkins, E. R. (2008). Repetitive negative thinking as a transdiagnostic process. *International Journal of Cognitive Therapy*, *1*(3), 192-205. https://doi.org/10.1521/ijct.2008.1.3.192

Elsner, B., & Hommel, B. (2001). Effect Anticipation and Action Control. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 229–240. https://psycnet.apa.org/doi/10.1037/0096-1523.27.1.229

Ersche, K. D., Gillan, C. M., Simon Jones, P., Williams, G. B., Ward, L. H. E., Luijten, M., De Wit, S., Sahakian, B. J., Bullmore, E. T., & Robbins, T. W. (2016). Carrots and sticks fail to change behavior in cocaine addiction. *Science*, *352*(6292), 1468–1471. https://doi.org/10.1126/science.aaf3700

Ersche, K. D., Lim, T. V., Ward, L. H. E., Robbins, T. W., & Stochl, J. (2017). Creature of Habit: A self-report measure of habitual routines and automatic tendencies in everyday life. *Personality and Individual Differences*, *116*, 73–85. https://doi.org/10.1016/j.paid.2017.04.024

Ersche, K. D., Ward, L. H. E., Lim, T. V., Lumsden, R. J., Sawiak, S. J., Robbins, T. W., & Stochl, J. (2019). Impulsivity and compulsivity are differentially associated with automaticity and routine on the Creature of Habit Scale. *Personality and Individual Differences*, *150*, 109493.

https://doi.org/10.1016/j.paid.2019.07.003

Farber, I. E., & Spence, K. W. (1956). Effects of anxiety, stress, and task variables on reaction time. *Journal of Personality*, *25*(1), 1–18. https://doi.org/10.1111/j.1467-6494.1956.tb01284.x

Hardwick, R. M., Forrence, A. D., Krakauer, J. W., & Haith, A. M. (2017). Skill acquisition and habit formation as distinct effects of practice. *BioRxiv*, 1–35. http://biorxiv.org/content/early/2017/10/14/201095.abstract

Hardwick, R. M., Forrence, A. D., Krakauer, J. W., & Haith, A. M. (2019). Time-dependent competition between goal-directed and habitual response preparation. *Nature Human Behaviour*, *3*(12), 1252–1262. https://doi.org/10.1038/s41562-019-0725-0

Heller, A. S., Ezie, C. E. C., Otto, A. R., & Timpano, K. R. (2018). Model-based learning and individual differences in depression: The moderating role of stress. *Behaviour Research and Therapy*, *111*, 19–26. https://doi.org/10.1016/j.brat.2018.09.007

Hertel, P. (2004). Memory for emotional and nonemotional events in depression. *Memory and Emotion*, 186–216. https://doi.org/10.1093/acprof:oso/9780195158564.003.0006

Heyes, C., & Dickinson, A. (1990). The intentionality of animal action. *Mind & Language*, *5*(1), 87–103. https://doi.org/10.1111/j.1468-0017.1990.tb00154.x

Hogarth, L. C., Mogg, K., Bradley, B. P., Duka, T., & Dickinson, A. (2003). Attentional orienting towards smoking-related stimuli. *Behavioural Pharmacology*, *14*(2), 153–160. https://doi.org/10.1097/00008877-200303000-00007

Hogarth, L., & Chase, H. W. (2011). Parallel goal-directed and habitual control of human drug-seeking: Implications for dependence vulnerability. *Journal of Experimental Psychology: Animal Behaviour Processes*, *37*(3), 261–276. https://doi.org/10.1037/a0022913

Hogarth, L., Chase, H. W., & Baess, K. (2012). Impaired goal-directed behavioural control in human impulsivity. *Quarterly Journal of Experimental Psychology*, *65*(2), 305–316. https://doi.org/10.1080/17470218.2010.518242

Holland, P. C. (2008). Cognitive versus stimulus-response theories of learning. *Learning and Behaviour*, *36*(3), 227–241. https://doi.org/10.3758/LB.36.3.227

Hull, C. L. (1933). Differential habituation to internal stimuli in the albino rat. *Journal of*

Comparative Psychology, *16*(2), 255. https://psycnet.apa.org/doi/10.1037/h0071710

Jefferys, W. H., & Berger, J. O. (1992). Ockham's razor and Bayesian analysis. *American Scientist, 80*(1), 64–72. https://www.jstor.org/stable/29774559

Ji, M. F., & Wood, W. (2007). Purchase and consumption habits: Not necessarily what you intend. *Journal of Consumer Psychology*, *17*(4), 261–276. https://doi.org/10.1016/S1057-7408(07)70037-2

Killcross, S., & Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex*, *13*(4), 400–408. https://doi.org/10.1093/cercor/13.4.400

Kopetz, C. E., Woerner, J. I., & Briskin, J. L. (2018). Another look at impulsivity: Could impulsive behaviour be strategic? *Social and Personality Psychology Compass*, *12*(5), e12385. https://doi.org/https://doi.org/10.1111/spc3.12385

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lally, P., van Jaarsveld, C. H. M., Potts, H. W. W., & Wardle, J. (2010). How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology*, *40*(6), 998-1009. https://doi.org/10.1002/ejsp.674

Landers, R. N., Bauer, K. N., & Callan, R. C. (2015). Gamification of task performance with leaderboards: A goal setting experiment. *Computers in Human Behaviour*, *71,* 508-515. http://dx.doi.org/10.1016/j.chb.2015.08.008

Le, J. T., Watson, P., Begg, D., Albertella, L., & Le Pelley, M. E. (2021). Physiological and subjective validation of a novel stress procedure: The Simple Singing Stress Procedure. *Behaviour Research Methods*, *53*(4), 1478–1487. https://doi.org/10.3758/s13428-020-01505-1

Le Pelley, M. E., Oakeshott, S. M., & McLaren, I. P. L. (2005). Blocking and unblocking in human causal learning. *Journal of Experimental Psychology: Animal Behaviour Processes*, *31*(1), 56. https://psycnet.apa.org/doi/10.1037/0097-7403.31.1.56

Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N., Wills, A. J., & Le Pelley, M. (2016). Attention and associative learning in humans: An integrative review. *Psychological*

Bulletin, 142(10), 1111. https://psycnet.apa.org/doi/10.1037/bul0000064

Liljeholm, M., Dunne, S., & O'Doherty, J. P. (2015). Differentiating neural systems mediating the acquisition vs. expression of goal-directed and habitual behavioural control. *European Journal of Neuroscience, 41*(10), 1358–1371. https://doi.org/10.1111/ejn.12897

Luque, D., Molinero, S., Watson, P., López, F. J., & Le Pelley, M. E. (2020). Measuring habit formation through goal-directed response switching. *Journal of Experimental Psychology: General, 149*(8), 1449. https://psycnet.apa.org/doi/10.1037/xge0000722

Marien, H., Aarts, H., & Custers, R. (2015). The interactive role of action-outcome learning and positive affective information in motivating human goal-directed behaviour. *Motivation Science*, *1*(3), 165. https://psycnet.apa.org/doi/10.1037/mot0000021

Miller, N. E. (1935). A reply to'Sign-Gestalt or conditioned reflex?'. *Psychological Review, 42*(3), 280–292. https://doi.org/https://doi.org/10.1037/h0054961

Mittner, M., Boekel, W., Tucker, A. M., Turner, B. M., Heathcote, A., & Forstmann, B. U. (2014). When the brain takes a break: A model-based analysis of mind wandering. *Journal of Neuroscience, 34*(49), 16286-16295. https://doi.org/10.1523/JNEUROSCI.2062-14.2014

Mkrtchian, A., Roiser, J. P., & Robinson, O. J. (2017). Threat of shock and aversive inhibition: Induced anxiety modulates Pavlovian-instrumental interactions. *Journal of Experimental Psychology: General*, *146*(12), 1694–1704. https://doi.org/10.1037/xge0000363

Moors, A., & De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychological Bulletin, 132*(2), 297. https://doi.org/https://doi.org/10.1037/0033-2909.132.2.297

Nelson, J., Klumparendt, A., Doebler, P., & Ehring, T. (2017). Childhood maltreatment and characteristics of adult depression: Meta-analysis. *British Journal of Psychiatry*, *210*(2), 96–104. https://doi.org/DOI: 10.1192/bjp.bp.115.180752

Neumann, O. (1984). Automatic processing: A review of recent findings and a plea for an old theory. *Cognition and Motor Processes*, 255-293. https://doi.org/10.1007/978-3-642-69382-3_17

Nolen-Hoeksema, S., Larson, J., & Grayson, C. (1999). Explaining the gender difference in depressive symptoms. *Journal of Personality and Social Psychology*, *77*(5), 1061.

https://psycnet.apa.org/doi/10.1037/0022-3514.77.5.1061

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking Rumination. *Perspectives on Psychological Science, 3*(5), 400–424. https://doi.org/10.1111/j.1745-6924.2008.00088.x

Ólafsson, R. P., Guðmundsdóttir, S. J., Björnsdóttir, T. D., & Snorrason, I. (2020). A test of the habit-goal framework of depressive rumination and its relevance to cognitive reactivity. *Behaviour Therapy*, *51*(3), 474–487. https://doi.org/https://doi.org/10.1016/j.beth.2019.08.005

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*(6), 768–774. https://doi.org/10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1

Piray, P., Toni, I., & Cools, R. (2016). Human choice strategy varies with anatomical projections from ventromedial prefrontal cortex to medial striatum. *Journal of Neuroscience*, *36*(10), 2857–2867. https://doi.org/10.1523/JNEUROSCI.2033-15.2016

Poling, A., Methot, L.L., LeSage, M.G. (1995). Between-subjects designs and non-experimental methods. In: *Fundamentals of Behaviour Analytic Research. Applied Clinical Psychology*. Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-1436-1_6

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. https://doi.org/10.1080/00335558008248231

Reason, J. T. (1979). Actions not as planned: The price of automatization. In G. Underwood & R. Stevens (Eds.), *Aspects of Consciousness* (pp. 67–89). Academic Press.

Roberts, H., Mostazir, M., Moberly, N. J., Watkins, E. R., & Adlam, A. L. (2021). Working memory updating training reduces state repetitive negative thinking: Proof-of-concept for a novel cognitive control training. *Behaviour Research and Therapy*, *142*, 103871.

Saling, L. L., & Phillips, J. G. (2007). Automatic behaviour: Efficient not mindless. *Brain Research Bulletin*, *73*(1), 1–20. https://doi.org/https://doi.org/10.1016/j.brainresbull.2007.02.009

Schneider, W., & Chein, J. M. (2003). Controlled & automatic processing: behaviour, theory, and biological mechanisms. *Cognitive Science*, *27*(3), 525-559. https://doi.org/10.1207/s15516709cog2703_8

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I.

Detection, search, and attention. Psychological Review, 84(1), 1–66. https://doi.org/10.1037/0033-295X.84.1.1

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, *15*(7), 319–326. https://doi.org/10.1016/j.tics.2011.05.006

Schroder, K. E. E., Ollis, C. L., & Davies, S. (2013). Habitual Self–Control: A Brief Measure of Persistent Goal Pursuit. *European Journal of Personality*, *27*(1), 82–95. https://doi.org/10.1002/per.1891

Schwabe, L., Dickinson, A., & Wolf, O. T. (2011). Stress, habits, and drug addiction: A psychoneuroendocrinological perspective. *Experimental and Clinical Psychopharmacology*, *19*(1), 53–63. https://doi.org/10.1037/a0022212

Seabrooke, T., Hogarth, L., & Mitchell, C. J. (2016). The propositional basis of cue-controlled reward seeking. *Quarterly Journal of Experimental Psychology*, *69*(12), 2452–2470. https://doi.org/10.1080/17470218.2015.1115885

Seabrooke, T., Le Pelley, M. E., Hogarth, L., & Mitchell, C. J. (2017). Evidence of a goal-directed process in human Pavlovian-instrumental transfer. *Journal of Experimental Psychology: Animal Learning and Cognition*, *43*(4), 377-387. https://doi.org/10.1037/xan0000147

Sebold, M., Deserno, L., Nebe, S., Schad, D. J., Garbusow, M., Hägele, C., Keller, J., Jünger, E., Kathmann, N., Smolka, M., Rapp, M. A., Schlagenhauf, F., Heinz, A., & Huys, Q. J. M. (2014). Model-based and model-free decisions in alcohol dependence. *Neuropsychobiology*, *70*(2), 122–131. https://doi.org/10.1159/000362840

Sjoerds, Z., De Wit, S., Van Den Brink, W., Robbins, T. W., Beekman, A. T. F., Penninx, B. W. J. H., & Veltman, D. J. (2013). Behavioural and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients. *Translational Psychiatry*, *3*(12), e337-8. https://doi.org/10.1038/tp.2013.107

Sniehotta, F. F., & Presseau, J. (2012). The habitual use of the self-report habit index. *Annals of Behavioural Medicine*, *43*(1), 139–140. https://doi.org/10.1007/s12160-011-9305-x

Spence, K. W. (1950). Cognitive versus stimulus-response theories of learning. In *Psychological*

Review, *57*(3), 159–172. https://doi.org/10.1037/h0058250

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. doi:10.1001/archinte.166.10.1092

Stanford, M. S., Mathias, C. W., Dougherty, D. M., Lake, S. L., Anderson, N. E., & Patton, J. H. (2009). Fifty years of the Barratt Impulsiveness Scale: An update and review. *Personality and Individual Differences*, *47*(5), 385–395. https://doi.org/https://doi.org/10.1016/j.paid.2009.04.008

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://psycnet.apa.org/doi/10.1037/h0054651

Thomson, D. R., Seli, P., Besner, D., & Smilek, D. (2014). On the link between mind wandering and task performance over time. *Consciousness and Cognition*, *27*, 14-26. https://doi.org/10.1016/j.concog.2014.04.001

Thrailkill, E. A., & Bouton, M. E. (2015). Contextual control of instrumental actions and habits. *Journal of Experimental Psychology: Animal Learning and Cognition*, *41*(1), 69–80. https://psycnet.apa.org/doi/10.1037/xan0000045

Tolman, E. C. (1932). *Purposive behaviour in animals and men*. University of California Press.

Treisman, M. (1960). Stimulus-response theory and expectancy. *British Journal of Psychology*, *51*(1), 49–60. https://doi.org/https://doi.org/10.1111/j.2044-8295.1960.tb00724.x

Tricomi, E., Balleine, B. W., & O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *European Journal of Neuroscience*, *29*(11), 2225–2232. https://doi.org/10.1111/j.1460-9568.2009.06796.x

Valentin, V. V, Dickinson, A., & O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *Journal of Neuroscience*, *27*(15), 4019–4026. https://doi.org/10.1523/JNEUROSCI.0564-07.2007

Verplanken, B. (2018). *The Psychology of Habit Theory, Mechanisms, Change, and Contexts.* Springer Cham. https://doi.org/10.1007/978-3-319-97529-0

Voon, V., Derbyshire, K., Rück, C., Irvine, M. A., Worbe, Y., Enander, J., Schreiber, L. R. N., Gillan, C., Fineberg, N. A., Sahakian, B. J., Robbins, T. W., Harrison, N. A., Wood, J., Daw, N. D., Dayan, P., Grant, J. E., & Bullmore, E. T. (2015). Disorders of compulsivity: A common bias towards learning habits. *Molecular Psychiatry, 20*(3), 345–352. https://doi.org/10.1038/mp.2014.44

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin, 134*(2), 163–206. https://doi.org/10.1037/0033-2909.134.2.163

Watkins, E. R., Baeyens, C. B., & Read, R. (2009). Concreteness training reduces dysphoria: proof-of-principle for repeated cognitive bias modification in depression. *Journal of Abnormal Psychology, 118*(1), 55–64. https://psycnet.apa.org/doi/10.1037/a0013642

Watkins, E. R., & Nolen-Hoeksema, S. (2014). A habit-goal framework of depressive rumination. *Journal of Abnormal Psychology*, *123*(1), 24–34. https://doi.org/10.1037/a0035540

Watkins, E. R., & Roberts, H. (2020). Reflecting on rumination: Consequences, causes, mechanisms and treatment of rumination. *Behaviour Research and Therapy, 127*, 103573. https://doi.org/10.1016/j.brat.2020.103573

Watson, J. B. (1924). The Unverbalized in Human Behaviour. *Psychological Review, 31*(4), 273–280. https://doi.org/10.1037/h0071569

Watson, P., & de Wit, S. (2018). Current limits of experimental research into habits and future directions. *Current Opinion in Behavioural Sciences*, *20*, 33–39. https://doi.org/10.1016/j.cobeha.2017.09.012

Wills, A. J., Graham, S., Koh, Z., McLaren, I. P. L., & Rolland, M. D. (2011). Effects of concurrent load on feature-and rule-based generalization in human contingency learning. *Journal of Experimental Psychology: Animal Behaviour Processes*, *37*(3), 308. https://psycnet.apa.org/doi/10.1037/a0023120

Wood, W., Mazar, A., & Neal, D. T. (2022). Habits and goals in human behaviour: Separate but interacting systems. *Perspectives on Psychological Science, 17*(2), 590–605. https://doi.org/10.1177/1745691621994226

Young, C. C., & Dietrich, M. S. (2015). Stressful life events, worry, and rumination predict depressive and anxiety symptoms in young adolescents. *Journal of Child and Adolescent*

Psychiatric Nursing, *28*(1), 35–42. https://doi.org/10.1111/jcap.12102

Appendices

Appendix 1: Photographs of famous faces


STYLE
RDS
H&M
ATRE
S 2016
ing
dard
HISCO
ATR
DS 2

Appendix 2: Creature of Habit Scale (COHS; Ersche et al., 2017)

Instructions:

Here are some statements relating to behaviours, feelings, or preferences that some people may have. Please indicate the extent to which you agree with each statement with regard to yourself. Please answer honestly, as there are no right or wrong answers.

1. I like to park my car or bike always in the same place.
2. I generally cook with the same spices / flavourings
3. When walking past a plate of sweets or biscuits, I can't resist taking one.
4. I tend to go to bed at roughly the same time every night.
5. I often take a snack while on the go (e.g., when driving, walking down the street, or surfing the web).
6. I quite happily work within my comfort zone rather than challenging myself, if I don't have to.
7. I tend to do things in the same order every morning (e.g., get up, go to the toilet, have a coffee...).
8. Eating crisps or biscuits straight out of the packet is typical of me.
9. Whenever I go into the kitchen, I typically look in the fridge.
10. I always try to get the same seat in places such as on the bus, in the cinema, or in church.
11. I often find myself finishing off a packet of biscuits just because it is lying there.
12. I normally buy the same foods from the same grocery store.
13. I rely on what is tried and tested rather than exploring something new.
14. I generally eat the same things for breakfast every day.
15. I tend to like routine.
16. I usually treat myself to a snack at the end of the workday.
17. In a restaurant, I tend to order dishes that I am familiar with.
18. I am one of those people who gets really annoyed by last minute cancellations.
19. I often find myself eating without being aware of it.
20. I usually sit at the same place at the dinner table.
21. I often find myself running on 'autopilot', and then wonder why I ended up in a particular place or doing something that I did not intend to do.
22. I always follow a certain order when preparing a meal.
23. Television makes me particularly prone to uncontrolled eating.
24. I tend to stick with the version of the software package that I am familiar with for as long as I can.
25. I often find myself opening up the cabinet to take a snack.

26. I am prone to eating more when I feel stressed.
27. I find comfort in regularity.

Answers:

0 = strongly disagree

1 = mildly disagree

2 = undecided

3 = mildly agree

4 = strongly agree

Appendix 3: Habitual Self-Control Questionnaire (HSCQ; Schroder et al., 2013)

Instructions:

Please indicate the extent to which you agree with each statement with regard to yourself. Please answer honestly, as there are no right or wrong answers.

1. I usually succeed in translating good intentions into action.
2. I always tackle important and difficult tasks without delay.
3. It would be easy for me to adopt a new habit such as doing exercise every day.
4. When a goal requires controlling my behaviour over a long period of time, I tend to give up after a while.
5. When I fail in a matter of some importance to me, I stick to my guns and try even harder than before.
6. If something important to me turns out to be quite difficult, I just persist in my efforts.
7. When I have made up my mind to complete an unpleasant task, nothing can stop me from doing so.
8. Even if I am really determined to do so, I often have difficulty rejecting a tempting offer.
9. When there is an opportunity to do something more enjoyable, my previous good intentions are usually lost.
10. I find it easy to motivate myself even if I do not enjoy a task at all.
11. Being rational and self-controlled does not seem to fit my way of life.
12. If I am convinced that completing a task is really important and worth the effort, I feel that I have a lot of willpower.
13. When there are more obstacles to overcome than I had expected, I tend to abandon the idea.
14. I was quite successful in controlling unwanted habits in the past.

Answers:

1 = disagree strongly
2 = disagree a little
3 = neither agree nor disagree
4 = agree a little
5 = agree strongly

Appendix 4: Barrett's Impulsiveness Scale (BIS-11; Patton et al., 1995)

Instructions:

People differ in the ways they act and think in different situations. This is a test to measure some of the ways in which you act and think. Do not spend too much time on any statement. Answer quickly and honestly.

1. I plan tasks carefully.
2. I do things without thinking.
3. I make-up my mind quickly.
4. I am happy-go-lucky.
5. I don't "pay attention."
6. I have "racing" thoughts.
7. I plan trips well ahead of time.
8. I am self-controlled.
9. I concentrate easily.
10. I save regularly.
11. I "squirm" at plays or lectures.
12. I am a careful thinker.
13. I plan for job security.
14. I say things without thinking.
15. I like to think about complex problems.
16. I change jobs.
17. I act "on impulse."
18. I get easily bored when solving thought problems.
19. I act on the spur of the moment.
20. I am a steady thinker.
21. I change residences.
22. I buy things on impulse.
23. I can only think about one thing at a time.
24. I change hobbies.
25. I spend or charge more than I earn.
26. I often have extraneous thoughts when thinking.
27. I am more interested in the present than the future.
28. I am restless at the theatre or lectures.
29. I like puzzles.
30. I am future oriented.

Answers:

1 = Rarely/Never
2 = Occasionally
3 = Often
4 = Almost Always/Always

Appendix 5: Perceived Stress Scale (PSS; Cohen et al., 1983)

Questions:

1. In the last month, how often have you been upset because of something that happened unexpectedly?
2. In the last month, how often have you felt that you were unable to control the important things in your life?
3. In the last month, how often have you felt nervous and “stressed”?
4. In the last month, how often have you felt confident about your ability to handle your personal problems?
5. In the last month, how often have you felt that things were going your way?
6. In the last month, how often have you found that you could not cope with all the things that you had to do?
7. In the last month, how often have you been able to control irritations in your life?
8. In the last month, how often have you felt that you were on top of things?
9. In the last month, how often have you been angered because of things that were outside of your control?
10. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?

Answers:

0 = Never
1 = Almost Never
2 = Sometimes
3 = Fairly Often
4 = Very Often

Appendix 6: Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001)

Instructions:

Over the last two weeks, how often have you been bothered by any of the following problems?

1. Little interest or pleasure in doing things?
2. Feeling down, depressed, or hopeless?
3. Trouble falling or staying asleep, or sleeping too much?
4. Feeling tired or having little energy?
5. Poor appetite or overeating?
6. Feeling bad about yourself - or that you are a failure or have let yourself or your family down?
7. Trouble concentrating on things, such as reading the newspaper or watching television?
8. Moving or speaking so slowly that other people could have noticed? Or the opposite - being so fidgety or restless that you have been moving around a lot more than usual?
9. Thoughts that you would be better off dead, or of hurting yourself in some way?

Answers:

0 = Not at all
1 = Several days
2 = More than half the days
3 = Nearly everyday

Appendix 7: Generalised Anxiety Scale (GAD-7; Spitzer et al., 2006)

Instructions:

Over the last 2 weeks, how often have you been bothered by any of the following problems?

1. Feeling nervous, anxious or on edge?
2. Not being able to stop or control worrying?
3. Worrying too much about different things?
4. Trouble relaxing?
5. Being so restless that it is hard to sit still?
6. Becoming easily annoyed or irritable?
7. Feeling afraid as if something awful might happen?

Answers:

0 = Not at all
1 = Several days
2 = More than half the days
3 = Nearly everyday

Appendix 8: Additional correlation results between self-reported measures

COHS and HSCQ correlations with BIS scores

As well as running correlational analyses between impulsivity and habitual behaviour measured in our instrumental experiments, we also ran correlations between impulsivity and self-reported habitual behaviour (measured using the COHS) and goal-pursuit traits (measured using the HSCQ). The results are shown in Table 7.

Table 7: Results of COHS and HSCQ Pearson's correlation analyses with BIS scores

Pearson's Correlations

Variable			Barrett Impulsivity Scale (BIS)			
			Total BIS score	Attentional score	Motor score	Non-planning score
COHS	Routine score	N	313	313	313	313
		Pearson's r	-0.09	-0.01	-0.10	-0.10
		p-value	.119	.985	.091	.070
		BF_{10}	0.24	0.07	0.29	0.37
	Automaticity score	N	313	313	313	313
		Pearson's r	0.33	0.33	0.17	0.25
		p-value	<.001***	<.001***	.002**	<.001***
		BF_{10}	$> 100^{\dagger\dagger\dagger\dagger}$	$> 100^{\dagger\dagger\dagger\dagger}$	$8.00^{\dagger}$	$> 100^{\dagger\dagger\dagger\dagger}$
HSCQ	Total HSCQ score	N	283	283	283	283
		Pearson's r	-0.49	-0.44	-0.17	-0.49
		p-value	<.001***	<.001***	.005**	<.001***
		BF_{10}	$> 100^{\dagger\dagger\dagger\dagger}$	$> 100^{\dagger\dagger\dagger\dagger}$	$3.75^{\dagger}$	$> 100^{\dagger\dagger\dagger\dagger}$

* p < .05, ** p <.01, *** p <.001

† BF_{10} = 3-10, †† BF_{10} = 10-30, ††† BF_{10} = 30-100, †††† $BF_{10} > 100$

COHS scores and BIS scores. Table 7 shows there was a positive correlation between COHS automaticity score and total BIS impulsivity, r=.33, p<.001, $BF_{10} > 100$, as illustrated in Figure 42a. Participants who scored higher on automaticity also scored higher on impulsivity. There was also a positive correlation between COHS automaticity score and the three BIS subscales: attentional impulsivity, r=.33, p<.001, $BF_{10} > 100$, as shown in Figure 42b; motor impulsivity, r=.17, p=.002, $BF_{10} = 8.00$, as shown in Figure 42c; and non-planning impulsivity, r=.25, p<.001, $BF_{10} > 100$, as shown in Figure 42d. There were no significant correlations between COHS routine and BIS scores.

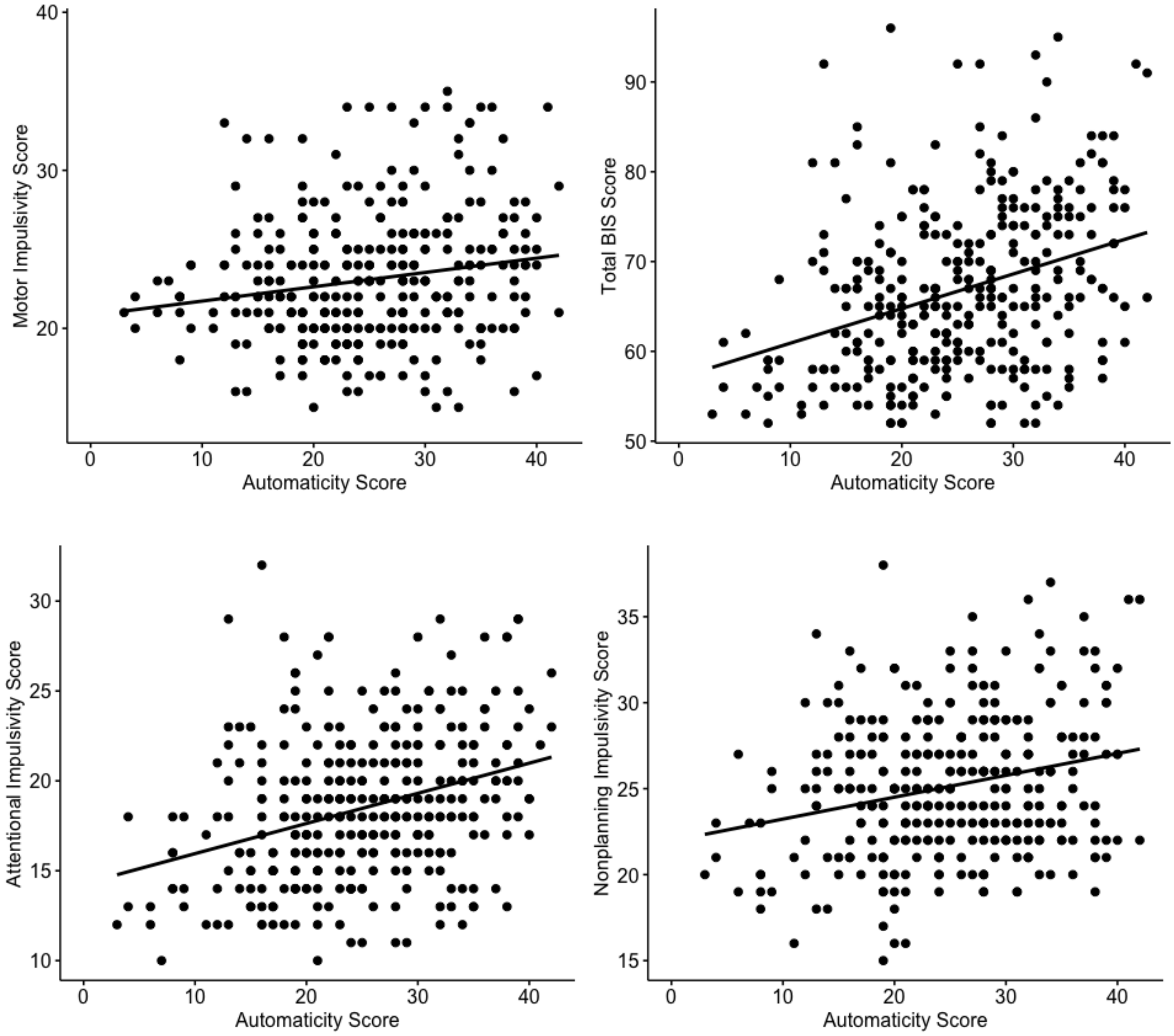


Figure 42: Scatter plot showing results of Pearson's correlation analysis between COHS automaticity score and BIS impulsivity scores: a) total BIS score (top left); b) attentional impulsivity score (top right); c) motor impulsivity score (bottom left); d) non-planning impulsivity score (bottom right).

HSCQ score and BIS scores. Figure 43a shows there was a negative correlation between HSCQ score and total BIS impulsivity, r=-.49, p<.001, BF_{10} > 100. Participants who scored higher on goal-pursuit tendencies scored lower on impulsivity. There was also a negative correlation between HSCQ score and the three BIS subscales: attentional impulsivity, r=-.44, p<.001, BF_{10} > 100, as shown in Figure 43b; motor impulsivity, r=-.17, p=.002, BF_{10} = 3.75, as shown in Figure 43c; and non-planning impulsivity, r=-.49, p<.001, BF_{10} > 100, as shown in Figure 43d.

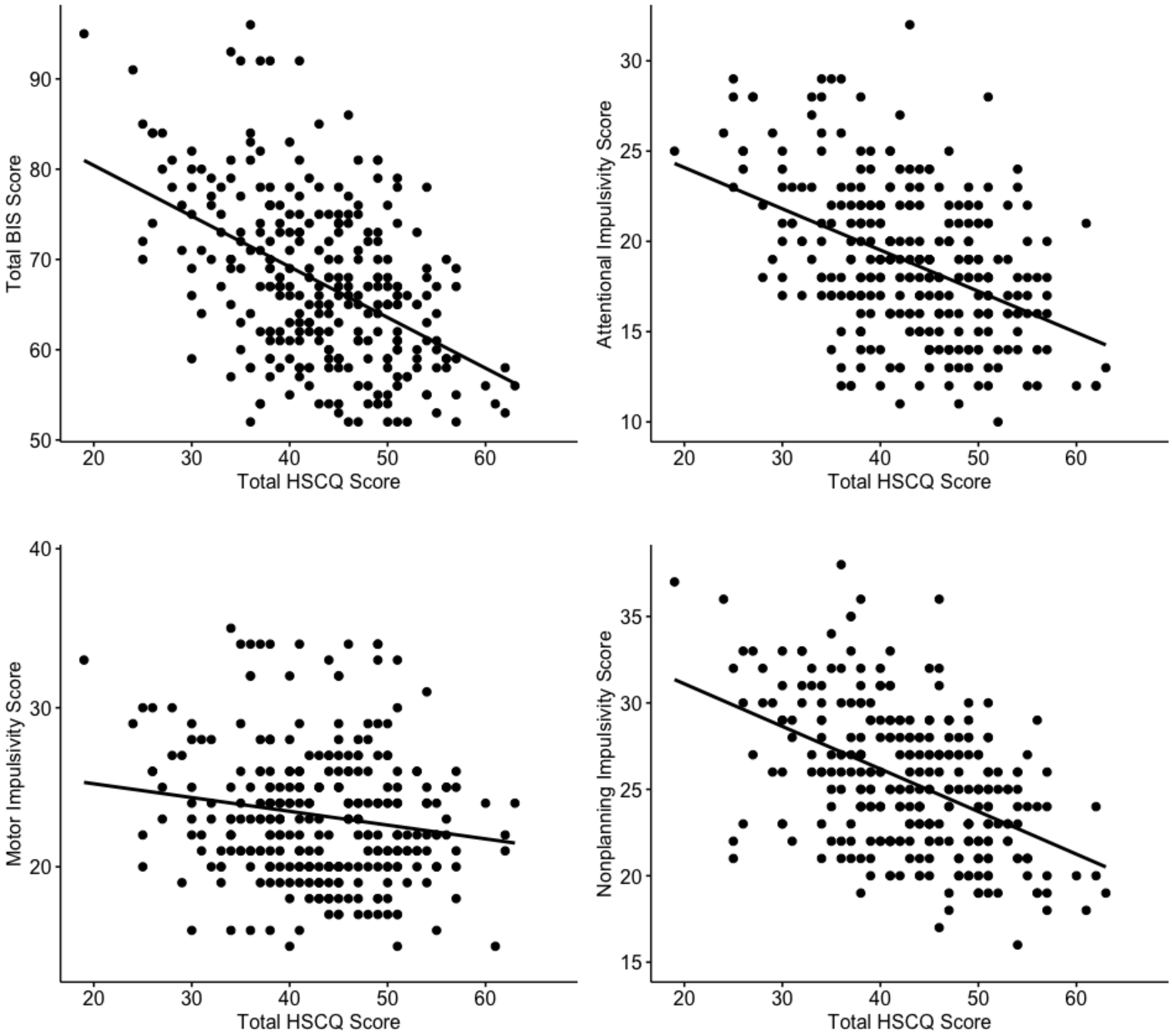


Figure 43: Scatter plot showing results of Pearson's correlation analysis between HSCQ goal pursuit score and BIS impulsivity scores: a) total BIS score (top left); b) attentional impulsivity score (top right); c) motor impulsivity score (bottom left); d) non-planning impulsivity score (bottom right).

PSS, PHQ and GAD correlations with COHS and HSCQ scores

As shown in Table 8, we also ran correlational analyses on self-reported habitual behaviour (COHS scores) and self-reported goal-directed behaviour (HSCQ score) against self-reported perceived stress (PSS), depression (PHQ), and anxiety (GAD) scores.

Table 8: Results of COHS and HSCQ Pearson's correlation analyses with PSS, PHQ and GAD

Pearson's Correlations

Variable			Total HSCQ score	Total PSS score	Total PHQ score	Total GAD score
COHS	Routine score	N	283	173	227	227
		Pearson's r	-0.08	0.13	0.19	0.22
		p-value	.183	.098	.005**	.001**
		BF_{10}	0.18	0.37	$4.36^{\dagger}$	$17.41^{\dagger\dagger}$
	Automaticity score	N	283	173	227	227
		Pearson's r	-0.40	0.44	0.31	0.32
		p-value	<.001***	<.001***	<.001***	<.001***
		BF_{10}	$> 100^{\dagger\dagger\dagger\dagger}$	$> 100^{\dagger\dagger\dagger\dagger}$	$> 100^{\dagger\dagger\dagger\dagger}$	$> 100^{\dagger\dagger\dagger\dagger}$
HSCQ	Total HSCQ score	N	-	143	217	217
		Pearson's r	-	-0.38	-0.20	-0.14
		p-value	-	<.001***	.003**	.039*
		BF_{10}	-	$> 100^{\dagger\dagger\dagger\dagger}$	$5.98^{\dagger}$	0.71

* p < .05, ** p <.01, *** p <.001

† BF_{10} = 3-10, †† BF_{10} = 10-30, ††† BF_{10} = 30-100, †††† BF_{10} > 100

COHS scores and HSCQ score. We firstly checked if COHS and HSCQ scores were correlated and found a significant negative correlation between COHS automaticity and HSCQ, r=-.40, p<.001, BF_{10} > 100, as shown in Figure 44a, where participants who scored higher on the automaticity measure tended to score lower on the goal-pursuit measure. There was no significant correlation between COHS routine and HSCQ score.

COHSS automaticity and PSS, PHQ and GAD scores. There were strong positive correlations between COHS automaticity and the three other individual difference measures: perceived stress (PSS), r=.44, p < .001, BF_{10} > 100, as shown in Figure 44b; depression (PHQ), r=.31, p < .001, BF_{10} > 100, as shown in Figure 44c; and anxiety (GAD), r=.32, p < .001, BF_{10} > 100, as shown in Figure 44d. Participants who scored higher on automaticity tended to report higher perceived stress, depression, and anxiety.

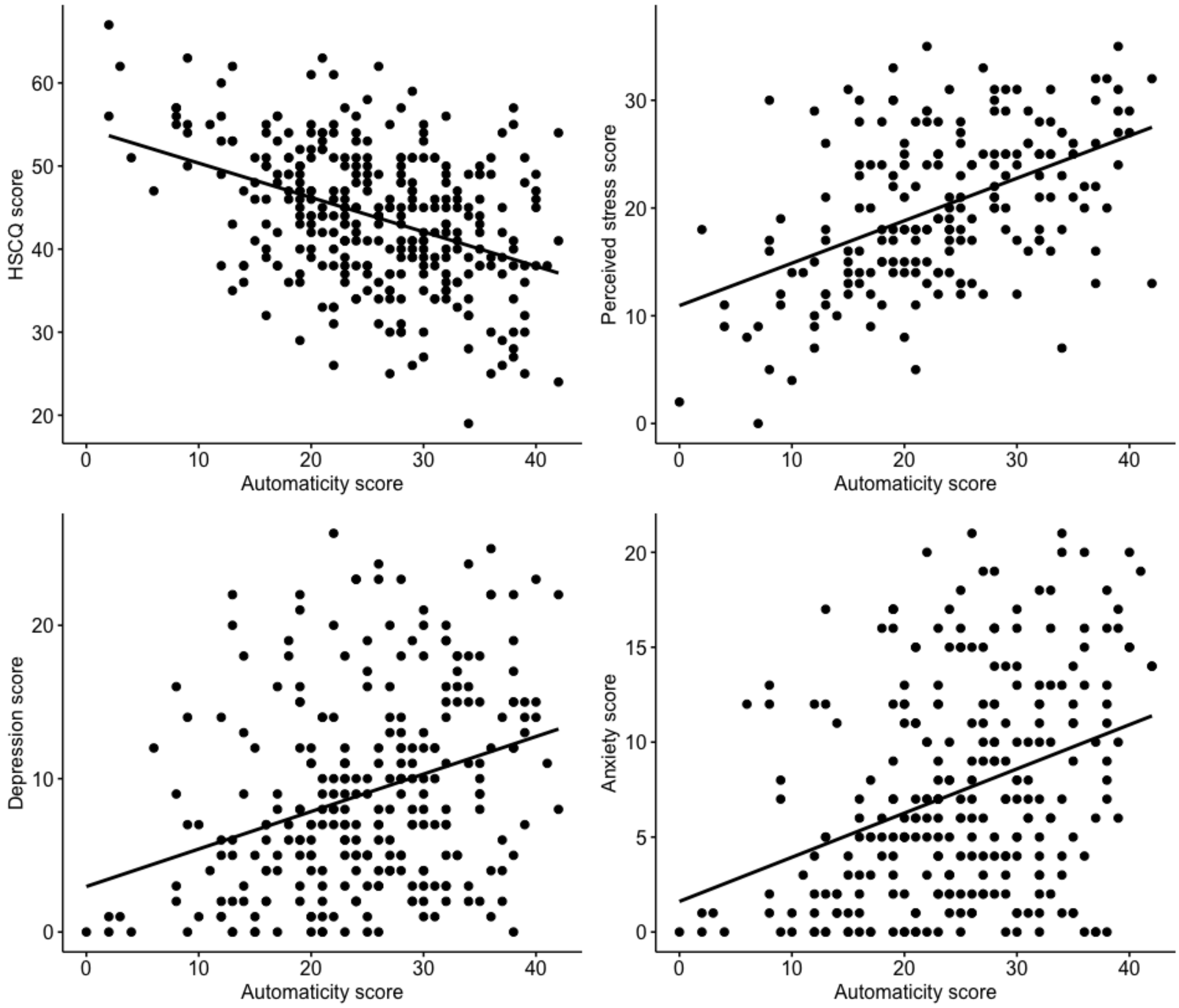


Figure 44: Scatter plot showing results of Pearson's correlation analysis between COHS automaticity score and the following individual differences measures: a) HSCQ goal pursuit score (top left); b) perceived stress (PSS) score (top right); c) depression (PHQ-9) score (bottom left); d) anxiety (GAD-7) score (bottom right).

COHSS routine and PSS, PHQ and GAD scores. COHS routine had positive correlations with depression, r=.19, p=.005, BF_{10} = 4.36, as shown in Figure 45a; and anxiety, r=.22, p=.001, BF_{10} = 17.41, as shown in Figure 45b. Participants who scored higher on routine behaviour tended to report higher depression and anxiety. There was no significant correlation between COHS routine and perceived stress.

HSCQ score and PSS, PHQ and GAD scores. HSCQ had significant negative correlations with perceived stress, r=-.38, p < .001, BF_{10} > 100, as shown in Figure 46a; depression, r=.20, p=.003, BF_{10} = 5.98, as shown in Figure 46b; and anxiety, r=-.14, p=.039, BF_{10} = 0.71, as shown in Figure 46c. Participants who scored higher on self-reported goal-directed traits scored lower on self-reported stress, depression, and anxiety. However, the Bayesian correlation did not support the anxiety and HSCQ correlation, $BF_{10} = 0.71$.

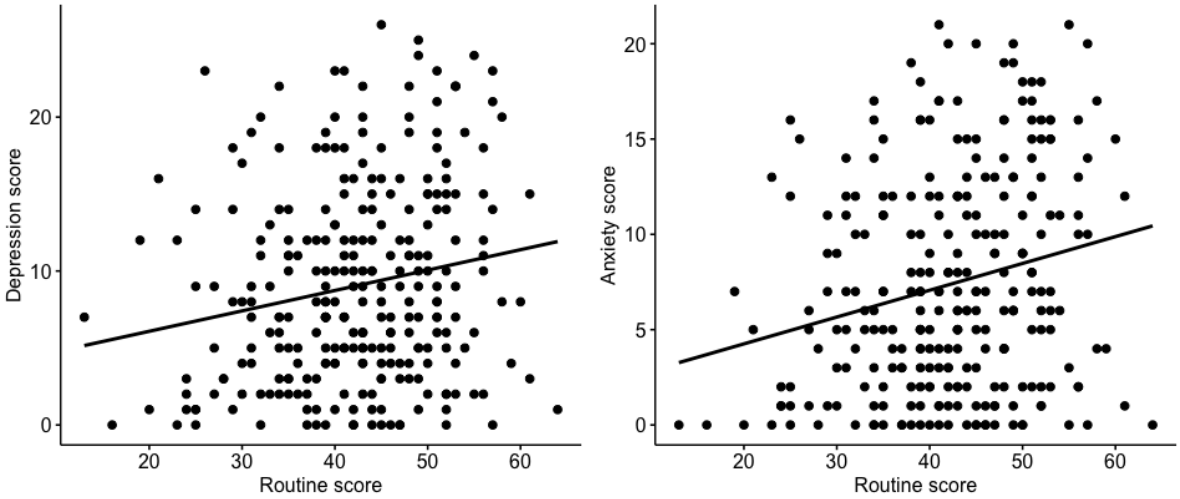


Figure 45: Scatter plot showing results of Pearson's correlation analysis between COHS routine score and the following individual differences measures: a) depression (PHQ-9) score (left); b) anxiety (GAD-7) score (right).

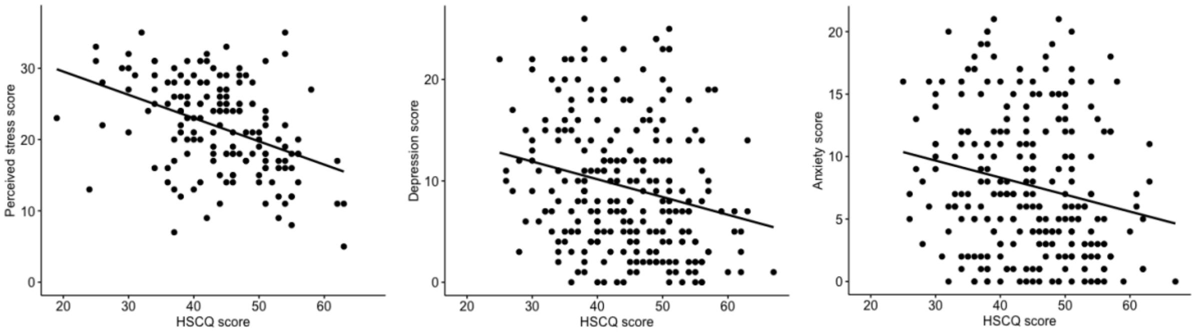


Figure 46: Scatter plot showing results of Pearson's correlation analysis between HSCQ goal pursuit score and the following individual differences measures: a) perceived stress (PSS) score (left); b) depression (PHQ-9) score (middle); c) anxiety (GAD-7) score (right).

PSS, PHQ and GAD correlations with BIS scores

As shown in Table 9, we also ran correlational analyses on self-reported impulsivity (BIS score) against self-reported perceived stress (PSS), depression (PHQ-9), and anxiety (GAD-7) scores.

Table 9: Results of BIS *Pearson's* correlation *analyse*s with PSS, PHQ and GAD

Pearson's Correlations

Variable			Barrett Impulsivity Scale (BIS)			
			Total BIS score	Attentional score	Motor score	Non-planning score
PSS	Total PSS score	N	405	405	405	405
		Pearson's r	0.30	0.48	0.10	0.10
		p-value	<.001***	<.001***	.054	.038*
		BF_{10}	> 100††††	> 100††††	0.40	0.53
PHQ-9	Total PHQ score	N	229	229	229	229
		Pearson's r	0.30	0.45	0.14	0.10
		p-value	<.001***	<.001***	.030*	.132
		BF_{10}	> 100††††	> 100††††	0.87	0.26
GAD-7	Total GAD score	N	229	229	229	229
		Pearson's r	0.25	0.39	0.13	0.05
		p-value	<.001***	<.001***	.045*	.497
		BF_{10}	> 100††††	> 100††††	0.61	0.10

* p < .05, ** p <.01, *** p <.001

† BF_{10} = 3-10, †† BF_{10} = 10-30, ††† BF_{10} = 30-100, †††† BF_{10} > 100

PSS and BIS correlations. Figure 47a shows stress (PSS) was positively correlated with total BIS impulsivity score, r=.30, p < .001, BF_{10} > 100, with more impulsive participants reporting higher levels of stress. This relationship was strongest for the attentional impulsivity subscale, r=.48, p < .001, BF_{10} > 100, as shown in Figure 47b. There was also a marginally significant positive correlation between stress and non-planning impulsivity, r=.10, p=.038, but this relationship was weak, and the Bayesian correlation did not support this, BF_{10} = 0.53.

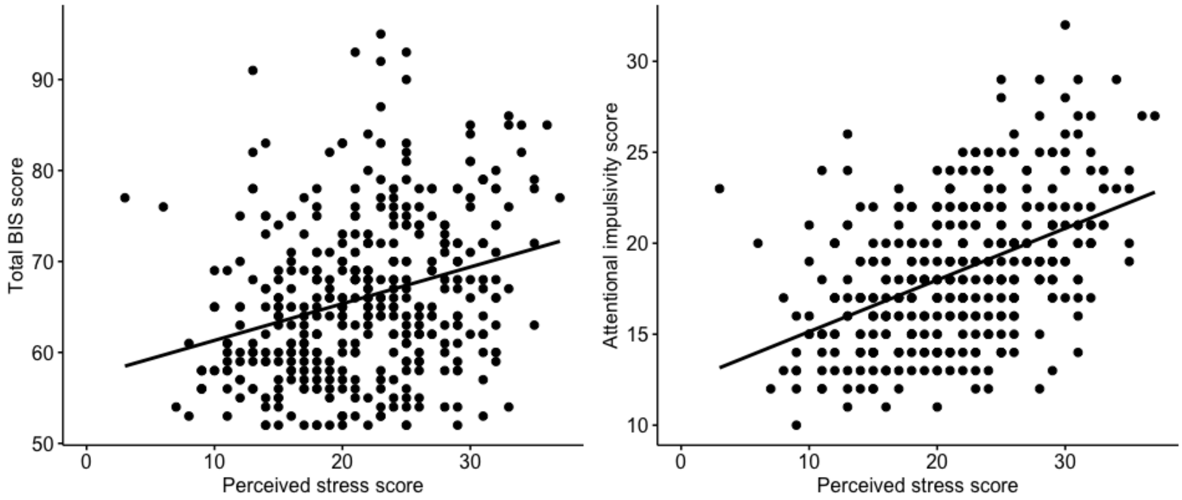


Figure 47: Scatter plot showing results of Pearson's correlation analysis between perceived stress (PSS) score and the following BIS impulsivity scores: a) total BIS impulsivity score (left); b) attentional impulsivity score (right).

PHQ and BIS correlations. Figure 48a shows depression was also positively correlated with total BIS score, r=.30, p < .001, BF_{10} > 100, with more impulsive participants reporting higher levels of depression. Again, this relationship was strongest for the attentional impulsivity subscale, r=.45, p < .001, BF_{10} > 100, as shown in Figure 48b. There was a marginally significant positive correlation between depression and motor impulsivity, r=.14, p=.039, but this relationship was weak, and the Bayesian correlation did not support this, BF_{10} = 0.87.

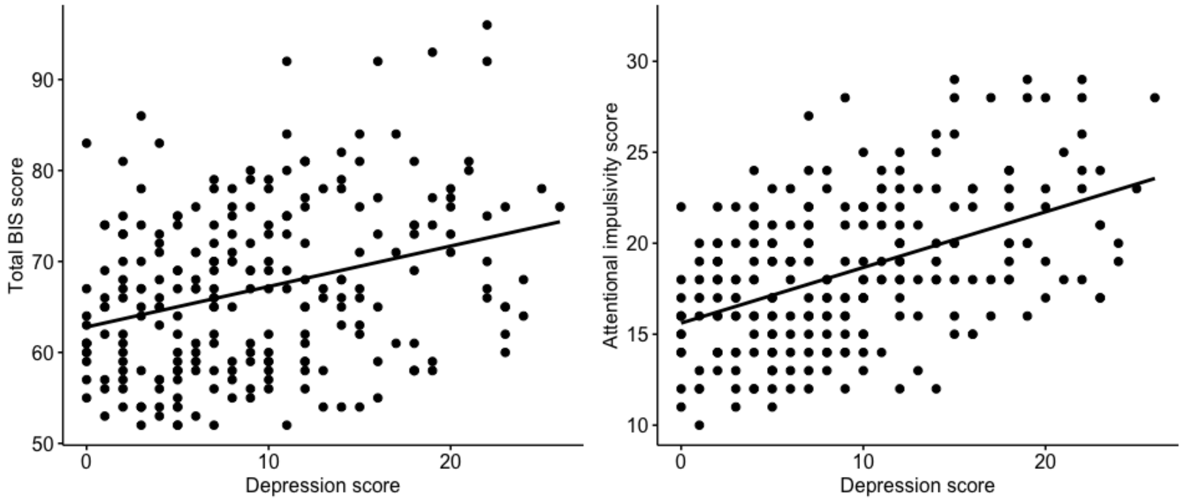


Figure 48: Scatter plot showing results of Pearson's correlation analysis between depression (PHQ-9) score and the following BIS impulsivity scores: a) total BIS impulsivity score (left); b) attentional impulsivity score (right).

GAD and BIS correlations. Figure 49a shows anxiety (GAD) was also positively correlated with total BIS impulsivity score, r=.25, p < .001, BF_{10} > 100, with more impulsive participants reporting higher levels of anxiety. As with stress and depression, the strongest relationship between impulsivity and anxiety was found in the attentional impulsivity subscale, r=.39, p < .001, BF_{10} > 100, as shown in Figure 49b. There was a marginally significant positive correlation between anxiety and motor impulsivity, r=.13, p=.045, but again this relationship was weak, and the Bayesian correlation did not support this, BF_{10} = 0.61.

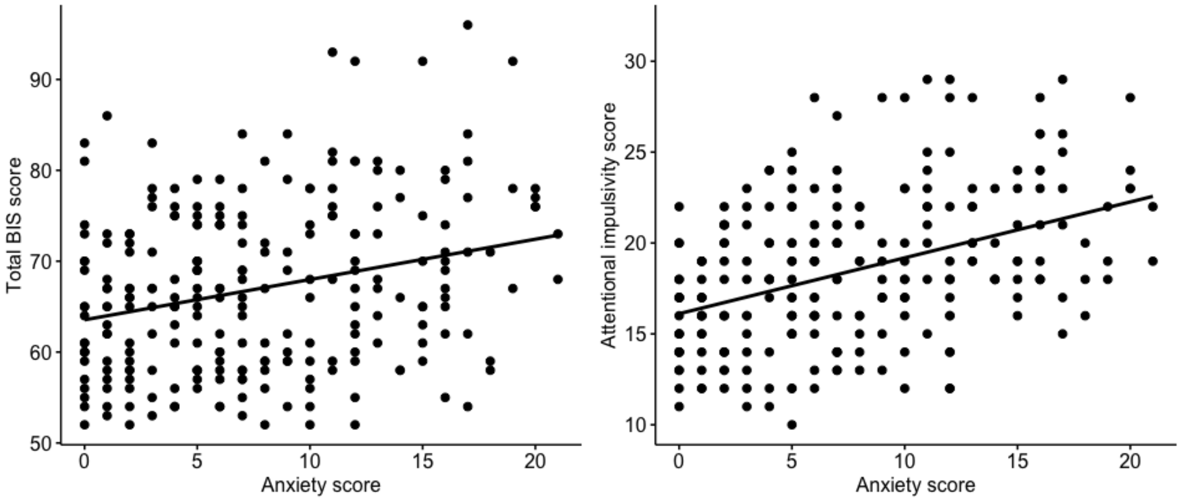


Figure 49: Scatter plot showing results of Pearson's correlation analysis between anxiety (GAD-7) score and the following BIS impulsivity scores: a) total BIS impulsivity score (left); b) attentional impulsivity score (right).