Youna Vandaele *Editor*

Habits

Their Definition, Neurobiology, and Role in Addiction

Springer

Habits

Youna Vandaele
Editor

Habits

Their Definition, Neurobiology, and Role in Addiction

Springer

Editor
Youna Vandaele
Laboratoire de Neurosciences
Expérimentales et Cliniques
Université de Poitiers, INSERM, U-1084
Poitiers, France

ISBN 978-3-031-55888-7 ISBN 978-3-031-55889-4 (eBook)
https://doi.org/10.1007/978-3-031-55889-4

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To Noa and Lionel

Preface

What is a habit? This question may appear trivial. Indeed, habit is intuitively easy to identify and describe; everyone can recognize his/her own habits and everyone knows that habits are carried out unconsciously and are notoriously hard to break. Very often, we realize that we have acted out of habit after making a mistake, for instance when taking the way back home from work before realizing that we were supposed to make a detour for grocery shopping. Habits are omnipresent in our daily life. This is because habits are so convenient; they allow us to act on a default mode, and free up cognitive space to perform other tasks requiring more attention. This is in contrast to purposeful, goal-directed behaviors, which involve planning, and thus, cognitive effort, and in which actions are explicitly performed with the goal of obtaining the desired outcome. Life would be unbearable without habits and if one must constantly think about all the consequences of all our actions! And yet, defining, operationalizing and investigating habits remain a real challenge.

Several issues have been receiving attention in the last few years; for instance, the operationalization of habit as an absence of goal-directed behavior, the difficulty in translating animal research to humans or the difficulty in measuring habits in more complex scenarios. Furthermore, different disciplines use different paradigms to investigate habits and it remains unclear whether we are studying the same concept. Thus, what is a habit? How are we forming and breaking habits? How do they relate to constructs such as perseveration, automaticity, skill learning, spatial navigation, or implicit associations? The first section of this book covers the multifaceted nature of habit. The first chapter presents the different definitions, methods and measures of habits in different disciplines. The second chapter focuses more specifically on the advantages and limitations of the outcome devaluation task, a powerful method to distinguish habitual from goal-directed behavior. The third chapter explores the behavioral mechanisms underlying formation and breaking of habits. In the fourth chapter, computational models of goal-directed and habitual learning systems are presented. Finally, the first section of this book ends with a thorough analysis of the commonalities and differences between the concepts of habit and perseveration.

Adaptive decision-making depends on the interaction between goal-directed and habitual processes, but how does our brain arbitrate between habitual and goal-directed processes? In the second section of this book, we explore the brain circuits underlying habits. The first chapter reviews computational theories on different decision-making models, including habitual and planned action sequences. In this chapter, the authors derive experimental predictions from these models and present neurophysiological data from hippocampus and striatum to shed light on the underlying neurobiological mechanisms. The second chapter proposes alternative options for defining habit and explores the function of the dorsolateral striatum in habitual learning from this broader perspective. The third chapter focuses instead on the role of different subregions of the prefrontal cortex in habitual and goal-directed control. The fourth chapter examines dopamine function in the striatum and its role in the transition from goal-directed behavior to automatic or habitual behavior. Finally, the last chapter explores neural networks involved in habitual and goal-directed processes, and provides evidence pointing towards a more integrated framework where the habitual and goal-directed systems work together to guide behavior.

An influential theory in the field of addiction suggests that compulsive drug use could arise from the progressive dominance of habits over goal-directed behaviors. However, it remains unclear whether and to what extent habits contribute to addiction. Chapters in the last section of this book explore the possible involvement of habits in alcohol and substance use disorders. This first chapter of this section reviews recent experimental work in humans supporting the hypothesis that maladaptive drug use in addiction arises from the progressive dominance of habits over goal-directed behaviors. The second chapter explores the relation between habits and implicit associations and delves into the critical role of implicit associations in addiction. The interaction and imbalance between the impulsive system, which drives automatic actions, and the reflective system, responsible for inhibitory control, is examined. The third chapter formalizes a recent elaboration on the "habit theory" of addiction using Bayesian reinforcement algorithms as models of habit and planning, to explore how the balance between goal-directed and habitual behavior could give rise to compulsivity and cognitive impulsivity. Finally, the last chapter explores the questions of whether individuals suffering from addiction are in control of their drug use behavior, and whether this control is habitual or goal-directed. In this chapter, it is suggested that the capacity of control fluctuates in addiction and that the interaction between habitual and goal-directed processes may help understanding this fluctuation.

This book explores the multiple facets of habit from diverse and complementary theoretical frameworks. It provides an exhaustive overview of the cognitive, computational and neural processes underlying distinct forms of habit and tries to address the delicate question of whether and how habit can contribute to psychiatric disorders such as alcohol or substance use disorder. This book is of interest to all researchers in behavioral and computational neuroscience, psychology and psychiatry, interested in associative learning and decision making, under normal and pathological conditions.

Poitiers, France Youna Vandaele

Acknowledgments

First, I would like to thank all the contributors of this book for their patience, perseverance and for their excellent contribution to this book. I am also grateful to William Lamsback who offered me this opportunity to edit a book on my favorite research topic. I thank him for his trust, encouragement and advice throughout the editing process. I would like to thank my two mentors for their continued support, Serge Ahmed for sharing his passion for Science and nurturing my critical thinking and Patricia Janak for guiding me through my first postdoc and helping me grow as an independent researcher. I thank Marcello Solinas and Nathalie Thiriet for the unfailing support they are giving me at a key step of my scientific career and for all the things I keep learning from them. Finally, I would like to give a special thanks to my husband, my forever ally in every personal and professional adventures.

Contents

Editor and Contributors

List of Contributors

Kenneth A. Amaya Tufts University School of Medicine, Boston, MA, USA

Priscilla Ambrosi Department of Neuroscience, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
Northwestern University Interdepartmental Neuroscience Program (NUIN), Evanston, IL, USA

Wesley C. Ameden Rutgers University – Newark, Newark, NJ, USA

Seiichiro Amemiya RIKEN Center for Brain Science, Saitama, Japan

Nidhi V. Banavar Department of Cognitive Sciences, University of California, Irvine, CA, USA

Suzanna Becker Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada

Aaron Bornstein Department of Cognitive Sciences, University of California, Irvine, CA, USA

Mark E. Bouton Department of Psychological Science, University of Vermont, Burlington, VT, USA

Armand Chatard Psychologie, Université de Poitiers, CNRS, Poitiers, Vienne, France

Laura H. Corbit Department of Psychology, The University of Toronto, Toronto, ON, Canada

Karen D. Ersche Department of Psychiatry, University of Cambridge, Cambridge, UK
Department of Addictive Behaviour and Addiction Medicine, Central Institute of Mental Health, University of Heidelberg, Mannheim, Germany

Isaac Kinley Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada

Talia N. Lerner Department of Neuroscience, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
Northwestern University Interdepartmental Neuroscience Program (NUIN), Evanston, IL, USA

Tsen Vei Lim Department of Psychiatry, University of Cambridge, Cambridge, UK

Andrew J. Miller-Hansen Department of Neuroscience, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

Ugurcan Mugan Department of Neuroscience, University of Minnesota, Minneapolis, MN, USA

Xavier Noël Université Libre de Bruxelles, ULB, Brussels, Belgium

Omar D. Perez Complex Engineering Systems Institute, Santiago, Chile
Department of Industrial Engineering, University of Chile, Santiago, Chile

Zachary Pierce-Messick Department of Psychology, The University of Toronto, Toronto, ON, Canada

A. David Redish Department of Neuroscience, University of Minnesota, Minneapolis, MN, USA

Paul S. Regier Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA

Kyle S. Smith Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA

Marcello Solinas Université de Poitiers, INSERM, Poitiers, Vienne, France

Elizabeth Tricomi Rutgers University – Newark, Newark, NJ, USA

Karly M. Turner School of Psychology, UNSW Sydney, Sydney, NSW, Australia

Youna Vandaele Laboratoire de Neurosciences Expérimentales et Cliniques, Université de Poitiers, INSERM, U-1084, Poitiers, France

Poppy Watson Psychology, University Technology of Sydney, Ultimo, NSW, Australia

Oulmann Zerhouni Université de Rouen Normandie, Rouen, Normandy, France

About the Editor

Youna Vandaele, PhD, is a neuroscientist working in the field of decision-making, behavioral control and addiction at the Inserm laboratory of experimental and clinical neurosciences, at the University of Poitiers, France.

The goal of her research is to understand the behavioral and neurobiological bases of maladaptive behaviors characterizing psychiatric disorders such as addiction, using animal models. She is particularly interested in improving the translational validity of animal models of addiction, by testing reward seeking in more complex scenarios that better capture the reality of alcohol or drug use in individuals suffering from this disorder.

Part I
Definitions, Operationalization and Investigation of Habits

Defining and Measuring Habits Across Different Fields of Research

Poppy Watson

Abstract Habits play an important role in our everyday behavioural repertoire, but they are conceptualised and measured differently across different fields of research. In this chapter, I compare the ways that habits are defined and measured by associative learning theorists, health and social psychologists, and those interested in the development of motor skills. The associative or knowledge structures that form once a habit has been acquired are also described and contrasted. Current theoretical debates about the representation of habits are discussed.

Keywords Habits · Goal-directed control · Outcome devaluation · Associative learning

1 Introduction

Habits are a frequently studied topic of human and animal behaviour. Broadly defined, habits are responses that are performed relatively automatically in stable contexts (Gardner, 2015; Robbins & Costa, 2017; Wood & Rünger, 2016). Habits are therefore computationally efficient, requiring minimal cognitive input and being mindlessly elicited by environmental cues. Habitual behaviour is thought to be a major component of our everyday behavioural repertoire and, in some models, is argued to be a critical component of pathological behaviour (i.e. behavioural inflexibility that may underlie compulsions and compulsivity; Everitt & Robbins, 2005, 2016; Verhoeven & de Wit, 2018).

Habits are defined differently across different fields of research in terms of the characteristics of the behaviour and how they should be measured. Furthermore, the representational structure that habits are argued to take differs across various fields. In this chapter, I will outline some of the different ways that habits are characterised and tested, considering the advantages and disadvantages of each approach. Finally,

P. Watson (✉)
Psychology, University Technology of Sydney, Ultimo, NSW, Australia
e-mail: poppy.watson@uts.edu.au

Y. Vandaele (ed.), *Habits*, https://doi.org/10.1007/978-3-031-55889-4_1

I will discuss the different representative structures that habits are proposed to take and highlight current discussions that are ongoing in this field.

2 Defining and Measuring Habits

2.1 Associative Learning Theory: Stimulus-Response (S-R) Habits

The narrowest and most empirically clear definition of habits is that offered by the associative dual-process framework (de Wit & Dickinson, 2009; Dickinson, 1985; Dickinson & Balleine, 1994). In a series of influential studies in the 1980s, Dickinson and colleagues demonstrated that responding could be controlled by the expected consequence of the action (response-outcome: R-O) or be triggered by stimuli (stimulus-response: S-R; Adams, 1982; Adams & Dickinson, 1981; Dickinson et al., 1983). Goal-directed behaviours (driven by anticipatory R-O associations) have two criteria—they are only carried out when the action reliably leads to the outcome (R-O contingency criterion) and the outcome is currently desired (desire criterion). On the flip side, habits are responses that are not sensitive to shifts in R-O contingency or outcome devaluation. That is, they persist even though the outcomes are no longer desired or despite the fact that the response no longer reliably produces the outcome.

The outcome devaluation test is the most commonly used method to assess whether observed behaviour is habitual or goal-directed (see Fig. 1a). Subjects learn associations between specific responses (e.g. a lever press or keyboard button press) and associated outcomes (e.g. food pellets or chocolate M&Ms). Following training, one of the outcomes is devalued (for within-subjects comparisons) or devalued for half the subjects (between-subjects comparisons). Outcome devaluation is typically achieved via feeding to satiation or via conditioned taste aversion (e.g. in rodents pairing the outcome with lithium chloride to produce gastric malaise). In human studies, ‘instructed devaluation’ is often used whereby participants are instructed that certain outcomes are no longer valuable and that responding to them will lead to loss of points/monetary bonuses. During the test phase, the subjects are given the opportunity to make instrumental responses (but now in extinction) and responding to instrumental responses associated with devalued versus valuable outcomes is compared. Behaviour that is immediately sensitive to the shift in outcome value is goal-directed (i.e. responding significantly more on responses associated with valued versus devalued outcomes), whereas behaviour insensitive to outcome devaluation is categorised as habitual—triggered directly by stimuli or contexts without consideration of the outcome that is produced.

One benefit of the associative dual-process model (and associative learning theory, more broadly) is that there are very clear definitions of the behaviours that can be classed as goal-directed or habitual. There is thus no ambiguity of what constitutes a ‘goal’ (something that is currently desired), and goal-directed

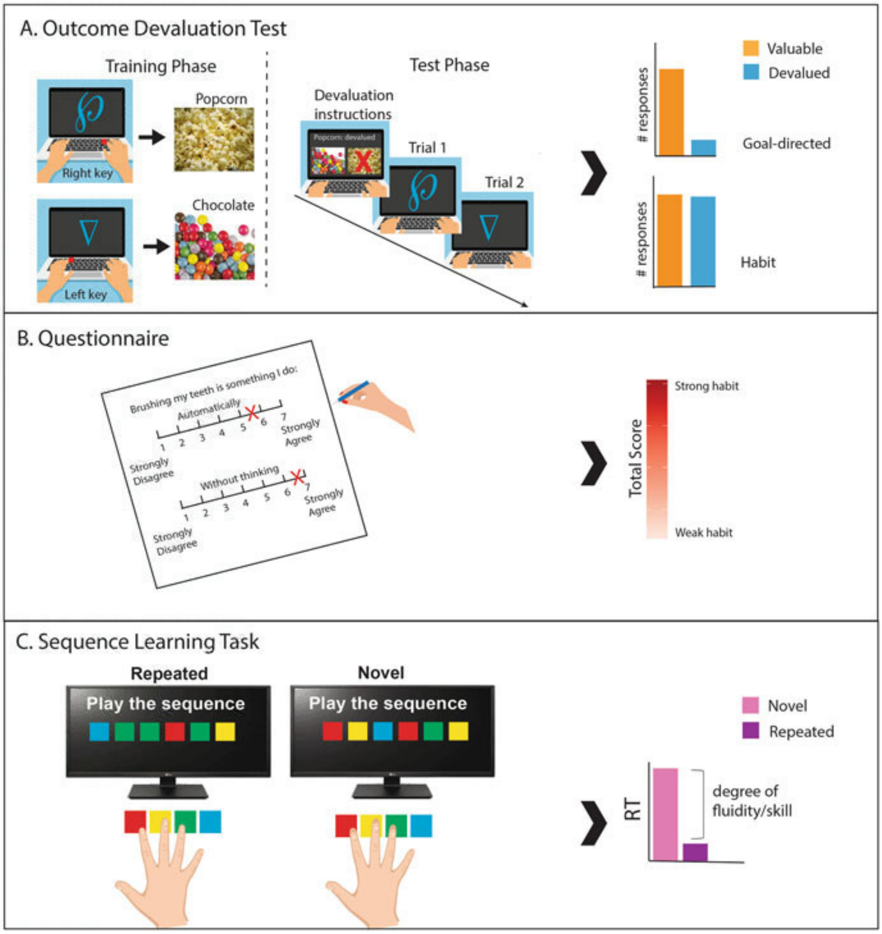

Fig. 1 Typical measurements of habitual behaviour. (**a**) The outcome devaluation test is commonly used in the field of associative learning to assess whether a behaviour is under goal-directed or habitual control. Participants learn that discriminative stimuli (e.g. symbols) signal whether a left or a right key press is required to earn outcomes (e.g. popcorn and chocolate). During the test phase, some of the outcomes are devalued (in this example by instruction). On each test trial, participants must then choose whether to make a response to earn the signalled outcome or not. Responding equally often for devalued as valuable outcomes is indicative of habit. (**b**) Questionnaire-based measures, commonly used in health and social psychology, ask participants to reflect on the degree to which certain behaviours are performed frequently and automatically. Higher scores represent stronger habits. (**c**) Motor skills are often studied using the Sequence Learning Task. Repeated sequences are interleaved within novel sequences and the RT advantage afforded towards repeated sequences at the end of training is indicative of fluid patterns of responding (i.e. skills or habits)

behaviour is a response that reduces in frequency when the R-O contingency is degraded or when the outcome is devalued. There are, however, some downsides to the near-exclusive use of the outcome devaluation paradigm in identifying habitual behaviour. Although the definition is clear (behaviour that persists despite outcome devaluation or contingency degradation), in practice, there are conceptual and operationalisation issues (see for review: Watson et al., 2022b). Perhaps most prominently, habits are typically inferred from a failure to find evidence for goal-directed control (Balleine & Dezfouli, 2019; Berridge, 2021; Robbins & Costa, 2017). That is, habits are inferred from a statistical null result. Furthermore, it is not always possible to break the contingency between actions and outcomes or devalue all types of outcomes. This is particularly vexing if we want to study behaviours in naturalistic contexts. Take, for example, toothbrushing behaviour. How can we establish whether such behaviour is habitual or goal-directed? According to the dual-process model, such assessment could not be made without devaluing the clean teeth outcome (or breaking the contingency between teeth cleaning and the sensation of clean teeth) and seeing whether the toothbrushing behaviour was immediately reduced. But of course, devaluing of the clean teeth goal (or breaking of the brushing-clean teeth contingency) is no easy feat.

This example illustrates that while the associative dual-process model of habits vs. goal-directed control has the benefit of being conceptually clear, it is perhaps best suited to study habits and goal-directed control in very controlled lab environments. This model has proved extremely valuable in the field of behavioural neuroscience where decades of work have established that different neural networks underly 'devaluation-insensitive' habits on the one hand and goal-directed control on the other (Balleine & O'Doherty, 2010; Turner & Parkes, 2020). In line with findings from non-human animals that damage to the dorsolateral striatum (DLS) reduces responding directed towards devalued outcomes (Yin et al., 2004), human neuroimaging studies have consistently found that an analogous region to the DLS in humans, the putamen, is implicated in studies of habit (see for reviews: Guida et al., 2022; Patterson & Knowlton, 2018). This conceptual framework, with clear experimental definitions of habits vs. goal-directed controls, has therefore proved extremely useful. However, for the study of human habit in naturalistic environments, the outcome devaluation and contingency-degradation paradigms become more difficult to implement.

Computational models of goal-directed control and habits that emerged from associative dual-process theories are more sophisticated, attempting to explain how goal-directed and habitual controllers interact to produce behaviour. For example, the hierarchical model proposed by Balleine and Dezfouli assumes that the habit controller is triggered in service to goal-directed behaviour (Balleine & Dezfouli, 2019; Dezfouli et al., 2014; Dezfouli & Balleine, 2013). Goal-directed decisions to act lead to initiation of well-rehearsed action sequences which may in turn be habitual (i.e. insensitive to the anticipated consequences). Such action sequences (or action chunks) are not sensitive to outcome devaluation as they are triggered as part of a sequence, distal from the outcomes that they eventually produce. Any devalued outcomes that are likely to be earned within the chunked action sequence are invisible to the goal-directed controller and thus not evaluated. This model retains the 'devaluation-insensitive' feature of habits, but in addition, characterises habits as fluid behavioural repertoires that are carried out fluently and quickly.

In contrast, model-free versus model-based theories argue that decision-making is driven by two distinct and non-interacting systems—the model-based system makes forward models of the likely outcomes and their desirability (akin to goal-directed control), whereas the model-free system uses cached outcome values to promote responses that previously worked in the past (Daw et al., 2005, 2011). Model-free decision-making is thus computationally efficient but not able to flexibly adjust when expected outcomes are no longer desired (similarly to S-R habits). This model assumes that the two systems are independent and that the strategy that will dominate decision-making is decided by an arbitrator. The controller that is eventually selected for action by the arbitrator is argued to be dependent on a variety of factors such as uncertainty about expected reward (Lee et al., 2014) or anticipated effort (Keramati et al., 2011).

These computational frameworks have their advantages, because more complex sequences of behaviours can be studied across the course of an experiment, rather than the crude measure offered by the outcome devaluation paradigm. These models also seek to explain the conditions under which behaviour is shifted back and forth between habitual and goal-directed controllers. However, this approach is not without its disadvantages. Because the computational model is so complex, mapping behaviour or neural activation uniquely to either controller can be difficult (Collins & Cockburn, 2020; da Silva & Hare, 2020). In regard to the model-free/model-based framework, it is not entirely clear what the model-free parameter represents, because it rarely correlates with other behavioural indices of habit (Watson & de Wit, 2018) and manipulations aimed at disrupting the model-free system have little impact (Collins & Cockburn, 2020). It has been disputed that model-free behaviour is akin to S-R habits (and therefore divorced from the outcomes that are produced) because both model-free and model-based learning are driven by a reward maximisation process (e.g. Miller et al., 2018, 2019).

2.2 Health/Social Psychology Habits

Habits have been intensely studied in the fields of health and social psychology as potentially amplifying the intention-behaviour gap (Kothe et al., 2015; Lally & Gardner, 2013). Take, for example, the frequent observation that an individual's intention to eat healthily and exercise regularly is often not sufficient to change behaviour. Instead, old patterns of sedentary behaviour and poor food choices dominate. Habits offer an explanation for why behaviour change can be difficult to implement—familiar contexts such as the living room can trigger the junk food eating response in a relatively automatic manner, regardless of any healthy food intentions (Gardner et al., 2011; Verhoeven et al., 2012; Webb et al., 2009; see for review: Verhoeven & de Wit, 2018). Furthermore, the implementation of 'healthy habits' is seen as a cornerstone in instigating successful and long-term behaviour change (Judah et al., 2013; Lally & Gardner, 2013; Mergelsberg et al., 2021). For new and desirable behaviours (such as regularly engaging in exercise) to become entrenched and durable, health psychologists promote the use of techniques that encourage habit formation. For example, in a randomised control trial of weight loss in obesity, participants in the intervention group were provided with simple diet and exercise advice in addition to being instructed to perform these behaviours repeatedly and in a stable context (Beeken et al., 2017). Weight loss was more rapid in participants who received these 'habit' instructions relative to the group who received standard weight loss advice from their primary care physician.

In the field of health psychology, habit strength is usually defined by self-report (Fig. 1b). Participants are asked to reflect on the frequency with which they perform certain behaviours as well as the perceived automaticity of the behaviour (e.g. 'I do it without thinking'). The most commonly used measurements are the 12-item self-report habit index (SRHI; Verplanken & Orbell, 2003) or the shorter self-report behavioural automaticity index (SRBAI; Gardner et al., 2012), which contains only four of the items from the SRHI. Many studies have demonstrated that habit strength, as measured with these scales, relates to patterns of behaviour in daily life (see for meta-analysis; Gardner et al., 2011). For example, Verhoeven et al. (2012) reported that the strongest predictor of caloric intake from unhealthy snacks across a 7-day period (as assessed with a snack diary) was self-reported habit strength of snacking on unhealthy junk foods over and above intentions to eat healthily, BMI, and scores on other eating questionnaires assessing hedonic responses to junk food.

Despite these interesting and important findings, there are a number of difficulties with relying on self-report as the main measure of habit strength. First, it is not clear how much insight participants are expected to have about a behaviour that is by definition largely outside of awareness (Hagger et al., 2015; Sniehotta & Presseau, 2012). Participants may also differ in the degree to which they are aware of their own behavioural automaticity. Second, as the dependent variable in many of these health behaviour studies is also a self-reported frequency measure (e.g. frequency of exercise as assessed from a daily diary), there is considerable conceptual and methodological overlap between the dependent variable and the habit measurement, potentially inflating the correlation between variables. It should be noted that these concerns are somewhat mitigated by studies that have related self-reported habit strength to more objective lab-based measures of behaviour (Watson et al., 2022a; Zhang et al., 2022) and/or made use of ecological momentary assessment (EMA) to measure daily health behaviours (Maher et al., 2021; Wouters et al., 2018). EMA involves delivering frequent short questionnaires to participants' smart phones asking about their eating or exercise behaviour over the past few hours rather than relying on participants to remember to regularly fill in a paper and pen diary.

The use of self-reported habit strength, while convenient to implement, also has implications for the way that habits are conceptualised. There is no clear cut-off for how much automaticity or behavioural frequency is required before a behaviour can be designated a habit. This blurs the distinction between 'habit' and 'intention' and means that there is no clear diagnostic test for identifying whether a behaviour is habitual or not. Instead, habit strength is always a relative measure, used to compare different behaviours and examine individual differences between participants (Fig. 1b).

2.3 Habits as Motor Skills

The broad definition of habits from associative learning and health psychology (as reflexive behaviours that are carried out relatively automatically in certain contexts) overlaps considerably with motor skills that are studied in the field of procedural learning (see for extensive review: Du et al., 2021). Skills are fluid behavioural repertoires that are acquired through repetition and carried out accurately and precisely with little cognitive effort. As such, this definition differs little from most other definitions of habits mentioned above. However, habits and skills are argued to be distinct (Du et al., 2021; Graybiel & Grafton, 2015). While a habit might be the relatively automatic initiation of a specific behavioural routine (e.g. the context prompting the initiation of the toothbrushing behaviour before bed each night, with little consideration of the 'clean teeth' outcome), the fluidity at which the sequence of motor behaviours is performed is a skill. Furthermore, an adept tennis player will be able to adjust his/her playing technique to different contexts (different types of courts and opponents), suggesting that behavioural inflexibility is less of a hallmark of motor skills than is usually attributed to habits.

Procedural motor skills are often studied in serial reaction time (SRT) paradigms where participants make repetitive sequences of behaviours (e.g. cued button presses) that are embedded within random sequences (Doyon et al., 2009; Hikosaka et al., 2002; Robertson, 2007). The speed at which these learned sequences are carried out relative to unlearned sequences is taken as evidence of a motor skill being acquired (see Fig. 1c). Thus, similar to the use of self-reported habit strength as evidence for habits in the fields of health and social psychology, there is no clear cut-off nor diagnostic test for assessing when a behaviour can be deemed a skill and whether this skill can transition to being habitual and inflexible.

Recently, Yang et al. investigated skill learning and the transition to behavioural inflexibility (Yang et al., 2022). Participants practised a difficult task where left and right hands had to move in different dimensions (up/down or left/right) to move a cursor on the screen in the opposite dimension (left/right or up/down, respectively). Participants had to move the cursor towards randomly presented targets and practise this for either 2, 5, or 10 days (between subjects). Skill was defined as the precision by which participants were able to move towards targets, which continued to improve in the 5-day group before plateauing. To assess habit strength, the mapping was flipped for the left hand after the last training session such that moving up and down would now move the cursor up and down rather than left to right. The authors reported that even after 2 days, participants were not able to adjust to the new mapping (making errors in line with the old mapping). This index of habit did not differ significantly across groups, regardless of the amount of training that they had.

The findings of Yang et al. (2022) are interesting because they suggest that behaviours can quickly become relatively inflexible (indicative of habit), but that beyond this point, they can still become more fluent and precise. There is thus the potential for a dissociation between the emergence of inflexible behaviour (which can emerge surprisingly quickly in lab-based tasks) and motor skills which may take longer to reach optimum efficiency of movement. In line with this, Vandaele et al. (2019) reported that in rats, insensitivity to outcome devaluation occurred prior to the development of highly optimised and stereotypical behaviour. In both humans and animals, there is considerable overlap in the regions of the basal ganglia implicated in both devaluation-insensitive habits and well-learned motor sequences in the SRT (Graybiel & Grafton, 2015; Janacsek et al., 2020; Patterson & Knowlton, 2018; Vandaele et al., 2019; Watson et al., 2018).

2.4 Summary

As the proceeding sections demonstrate, there is considerable overlap in the broad definition of habits across different fields of research. However, significant diversions in methodological approaches have implications for how habits are conceptualised. Associative dual-process models offer the clearest characterisation of habits, as distinct from goal-directed action. However, the outcome devaluation and contingency-degradation tests that are used to identify habits are often difficult to implement outside of the highly controlled lab environment. In health and social psychology, habits tend to be measured by self-report. This is convenient to implement, but sometimes, it is not entirely clear what is being measured beyond simply the frequency with which a behaviour is carried out. Finally, there is considerable overlap between motor skills and habit, with the former being measured as faster and more precise patterns of movement. Two meta-analyses reported that largely overlapping neural architecture is reported to underlie devaluation-insensitive habits and well-learned sequences of behaviour in the SRT (Patterson & Knowlton, 2018) as well as repetitive behaviours from daily life such as walking and driving (Guida et al., 2022). Some researchers have attempted to relate self-reported measures of automaticity to performance in outcome devaluation tests (Watson et al., 2022a) and model-free learning parameters to performance in outcome devaluation tests (Gillan et al., 2015; Sjoerds et al., 2016) and examined the different rates at which behavioural inflexibility and skills develop (Yang et al., 2022). Overall, while conceptually related, these different methodologies are likely tapping into different behavioural processes. More work is therefore needed to refine the concept and measurement of habit and elucidate the parameters of behavioural repetition, fluidity, and inflexibility following changes in outcome contingency or value that define habit.

3 Representative Structures

When one has acquired a habit, what exactly has been learned? Associative structures or knowledge structures are a way of parsimoniously describing how context or stimuli in the environment drive habitual responses. Unsurprisingly, there are considerable differences across fields about the representative structure that habits are argued to take. The main point of difference is whether stimuli can directly trigger responses, without reference to the associated goal or whether the outcome needs to feature in the representational structure.

3.1 S-R Associative Structures

According to the dual-process model, both S-R and bidirectional R-O associative links are formed during training (de Wit & Dickinson, 2009; Dickinson, 1985; Ostlund & Balleine, 2007). Initially, R-O associations will dominate responding (with behaviour being sensitive to outcome devaluation or changes in contingency), but eventually, behaviour will transition to being under the control of S-R associations (Fig. 2a). There are various ways to promote this transition from goal-directed to habitual behaviour, including schedules of reinforcement during instrumental training, pharmacological manipulations, and extended instrumental training (Adams, 1982; Nelson & Killcross, 2006; Yin et al., 2005; see for review: Balleine & O'Doherty, 2010). With limited instrumental training, behaviour should

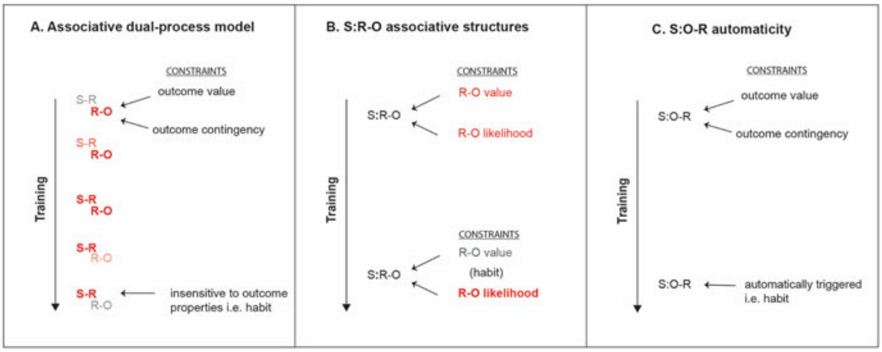

Fig. 2 Some of the representative structures that habits are argued to take. (**a**) According to the associative dual-process model, behaviour that is goal-directed is driven by R-O associations and sensitive to the value of the outcome and the contingency between responses and outcomes. Over time, however, the S-R association becomes stronger and eventually dominates responding (the outcome is not represented as part of the habit structure). (**b**) Hierarchical S:R-O models propose that the stimulus signals the R-O contingency that is most likely to be in place. Both likelihood and outcome value estimates are combined into a utility estimate and the response with the highest utility is selected for action. After extended training, likelihood estimates can override value estimates producing behaviour insensitive to devaluation. (**c**) S:O-R models propose that stimuli signal outcomes which in turn trigger the associated behavioural responses. Early in training, the desirability of goals and the potential responses are carefully considered, but after extended training, the O-R associative structure can be triggered automatically by stimuli, outside of awareness or desire for outcomes

be sensitive to outcome devaluation, whereas after extended training behaviour should be triggered by stimuli (S-R), regardless of the associated outcome value. Perhaps surprisingly, considering that the associative dual-process model is much concerned with representations of associative structures, direct evidence for S-R associative links that are proposed to underlie habits is largely lacking (Colwill & Rescorla, 1986; De Houwer, 2019; Watson et al., 2022b). Although overtraining of S-R associations is the most logical way to obtain direct evidence for S-R associative links, several human and animal studies have demonstrated that, surprisingly, S-R habits do not appear to be more pronounced following extended training (de Wit et al., 2018; Pool et al., 2022; Watson et al., 2022a).

Recently, Hardwick et al. (2019) appeared to show conclusive evidence for S-R associative links following overtraining, using a contingency reversal paradigm. Participants learned to make four specific keyboard responses in the presence of four discriminative stimuli, reinforced with correct/incorrect feedback. At test, two of the responses were switched. The study demonstrated that after 4 days, but not 1 day of training, participants were more likely to carry out the response that had been previously trained with that stimulus, when under considerable time pressure. This study has been widely interpreted as a demonstration of S-R associative links becoming stronger over the course of training and interfering with goal-directed control at test. However, a recent study from Buabang et al. (2022) calls this conclusion into question. Buabang et al. used the contingency reversal task from Hardwick et al. (2019) with some modifications. First, participants earned diamonds for correct responses and rocks for incorrect responses during training. In addition, participants completed occasional probe trials. On probe trials, a stimulus was presented and one of the potential responses was highlighted. Participants' task on probe trials was to indicate which outcome (rock or diamond) would be earned on the basis of the illustrated S-R mapping. Buabang et al. reported that participants who made more errors on these R-O catch trials (i.e. selecting 'diamond' when the actual outcome that would be earned was a rock) were also more likely to make the incorrect (previously learned) response to switched stimuli on the standard test trials. Although there may be different processes underlying responding to this task (on standard trials) and evaluation of outcomes (on probe trials), the results of this study raise the possibility that errors on the Hardwick task following extended training were not due to overtrained S-R associative links (without regard for whether or not that was correct) but incorrect S-R-O associations. That is, despite the fact that the outcome associated with a particular S-R contingency had changed (from a diamond to a rock), participants (under time pressure) still thought the previous S-R contingency was in place (Buabang et al., 2022).

In a typical outcome devaluation test, the relationships between stimuli, responses, and outcomes do not change, instead, the *value* of the outcomes changes and participants who are acting in a goal-directed manner should reduce responding to that outcome (de Wit & Dickinson, 2009; Sjoerds et al., 2016; Watson et al., 2022a). Thus, unlike the studies of Hardwick et al. (2019) and Buabang et al. (2022), R-O contingencies do not need updating in an outcome devaluation task and are thus unlikely to drive error responses directed towards devalued outcomes. In line with this, a number of studies have reported that participants have excellent knowledge of R-O contingencies at the end of outcome devaluation experiments (Watson et al., 2018, 2022a). However, it is theoretically possible that under time pressure participants are not confident of whether the outcome being signalled by the stimulus is still valuable or is now devalued. This is particularly pertinent in human studies where instructed devaluation is often employed and the valuable and devalued outcomes change on each block (e.g. Watson et al., 2018, 2022a). Assessing whether participants retain knowledge of the devalued/valuable outcomes after each block should be a standard practice in these types of tasks to establish whether errors are due to well-established S-R associations (as proponents of the association dual-process model argue) or due to miscalculations about expected outcome value. This is a relevant issue for understanding the role of habits in maladaptive reward-seeking behaviours (as seen in addiction, for example). Many studies have reported impaired performance on outcome devaluation tests in individuals with drug addiction, which is often interpreted as an overreliance on habits in this population relative to healthy controls (Ersche et al., 2016; Sjoerds et al., 2013; see for review: Verhoeven & de Wit, 2018). However, poor task contingency knowledge could also lead to performance impairments in outcome devaluation tasks and it is critical that any differences between clinical/healthy groups are carefully examined and controlled.

It is worth noting that the apparent R-O errors identified by Buabang et al. (2022) were stimulus-bound. That is, the incorrect R-O associations were elicited by specific stimuli (e.g. participants expected *left key press-diamond* and *right key press-rock* in the presence of a red door (S) when in fact the correct (updated) association was *left key press-rock and right key press-diamond* in the presence of that stimulus). Thus, the stimulus needs to be represented in the associative structure that produced the habitual behaviour in this task. There are at least two possible associative mechanisms that can account for interdependency between stimuli and R-O associations. One is the binary S-O-R account where the stimulus (e.g. red door) triggered anticipation of the diamond outcome (S-O) which then primed the associated response (due to the bidirectional nature of R-O/O-R associations). The other is a hierarchical explanation where the stimulus signals the most probable R-O contingency in place and this probability estimate also contributes to response selection (Colwill & Rescorla, 1990; Hogarth & Troisi, 2015). Assuming that behaviour in S-R reversal tasks (such as those used by Buabang et al., 2022 and Hardwick et al., 2019) does involve outcome anticipation, then the binary S-O-R account is insufficient to explain the findings. Because outcomes (diamonds/rocks or correct/incorrect feedback) were equally often associated with the various responses, it would be impossible for anticipation of an outcome (S-O) to trigger one response over another. A more complex hierarchal account is therefore required.

3.2 S:RO Associative Structures

Various hierarchical models of action control have been proposed by associative learning theorists over the years (Balleine & Ostlund, 2007; Colwill & Rescorla, 1990; Hogarth & Troisi, 2015; Rescorla, 1991). The most widely adopted of these is the S:R-O model, originally proposed by Rescorla, where the stimulus directly operates on the gestalt R-O representation (Bradfield & Balleine, 2013; Hogarth & Troisi, 2015). As can be seen in Fig. 2b, initially both the probability of the R-O being in place (given the current stimulus) and the expected value of the associated outcome are assigned to specific R-O contingencies and weighted for action. For example, in the contingency reversal paradigm described above, participants learn during training that a red door stimulus signals that the *left button-diamond* outcome contingency is more likely to be in place and more valuable than the *left button-rock* contingency and is as equivalently valuable but more likely to be in place than the *right button-diamond* contingency. These value and probability estimates are then combined into a utility estimate and the response with the highest utility is selected for action.

According to this hierarchical model (Hogarth & Troisi, 2015), devaluation-insensitive (or habitual) behaviour emerges because strong probability estimates may override value estimates. Thus, overtraining can lead to behaviour that is insensitive to outcome devaluation, but not because of stronger S-R associations, instead because of the high probability assigned to a specific S:R-O contingency. When this model was applied to the contingency reversal task (Buabang et al., 2022; Hardwick et al., 2019), then inflexible behaviour arose because the S:R-O contingencies had been extensively trained in the 4-day condition (but not 1-day condition) and the stimulus triggered the old (most probable) R-O contingency, despite the fact that it was no longer in place.

This is interesting because it suggests that hierarchical S:R-O associations may become more ingrained (and habit-like) following extended training, in the same way that S-R habits are proposed to. Although this seems logical, this is an area of research that has received little attention. Unlike the associative dual-process model, which clearly outlines the transition from goal-directed O-R to S-R associations as the mechanism underlying the emergence of behavioural inflexibility, hierarchical models are less clear on the 'tipping point' for behavioural inflexibility (see Fig. 2). More work is needed on assessing the hierarchical versus binary accounts of habit to see whether stimuli can directly trigger responses without recourse to outcome representations (as associative dual-process models propose) or whether the outcome also needs to be represented in the habit structure. It is of course theoretically possible that different models could explain behaviour depending on task requirements and experimental settings. In line with this, there is some evidence to suggest that both binary and hierarchical representations are acquired during instrumental training and that they compete for behavioural control (Bradfield & Balleine, 2013). Whether these different explanations for associative representations of habit can ultimately be teased apart and evaluated remains however to be seen.

3.3 S:O-R Automaticity

In terms of representative structures, most health psychologists are not concerned with how habits and goals are represented (focusing instead on how self-reported habit strength interacts with other variables to produce behaviour in the real world). In general, however, many of the assumptions of representational structures that were introduced from the field of social psychology have been adopted by health psychologists. Some of these align broadly with the associative dual-process model, proposing that contexts and stimuli can directly trigger responses in a relatively automatic manner, largely independent of cognitive processes (Wood et al., 2022; Wood & Rünger, 2016). However, a number of these theories are significantly different in that they assume that all behaviours are mediated by goal representations (Aarts & Dijksterhuis, 2000; Custers & Aarts, 2010; Kruglanski & Szumowska, 2020; Moors et al., 2017).

Aarts and colleagues, for example, proposed that by repeatedly acting in a goal-directed manner, the links between goals and responses made in service of achieving those goals get stronger (see Fig. 2c). In a hierarchical S:O-R manner, goal representations can be primed by situations and contexts which in turn can trigger the associated response in a reflexive and habitual manner (Aarts & Dijksterhuis, 2000; see for extensive discussion of these models: Balleine & Ostlund, 2007). It is generally assumed that over time, goal representations can be primed by stimuli outside of awareness and outside of desire (Aarts & Dijksterhuis, 2000; Custers & Aarts, 2010; Papies & Aarts, 2011). Take for example, a study of Aarts and colleagues where under working memory load (or not), Dutch students were asked to either recall or suppress the typical mode of travel they used to get to various destinations. Participants under working memory load struggled to *suppress* their typical transport mode when the destination (goal) was presented, whereas *recalling* their typical transport mode was not affected by memory load. This dissociation of suppressing vs. recalling while executive control functions were taxed is argued to be evidence that each goal (destination) was associated with a habitual travel response, which came automatically to mind when the goal was prompted (via O-R links). Under this model, goal representations are critical for driving habit responses, as opposed to the S-R links proposed by associative dual-process theory.

Other hierarchical models also feature outcome representations in habit structures (Balleine & Dezfouli, 2019; Cushman & Morris, 2015; Dezfouli & Balleine, 2013; Du et al., 2021). Some models propose that goals can be triggered in a habitual manner, but the behaviour used to reach that goal could be flexible (Cushman & Morris, 2015). Alternatively, others proposed that desired outcomes could be considered in a goal-directed manner, but triggered habitual patterns of behaviour as a means to achieve those goals (Balleine & Dezfouli, 2019; Dezfouli et al., 2014; Dezfouli & Balleine, 2013). In their flexible model, Du et al. (2021) argued that both accounts are plausible: 'If seeing a red light habitually triggers the goal of stopping, one could still select among actions to stop the car in a goal-directed way; conversely, if pressing the brake pedal to stop the car becomes habitual, one could still be flexible about what goal to pursue in the event of a red light' (Du et al., 2021, p. 379).

An extreme approach is taken by Kruglanski and Szumowska (2020), who argued that all behaviours are ultimately goal-directed. Under this framework, stimuli signal that certain rewards are available and behaviour will only be initiated *if and only if* the subject anticipates that this will lead to the reward. Thus, if a behaviour is observed, it must be assumed that the individual was anticipating the outcome (and currently desired it). If individuals are seen to be making responses that lead to outcomes that they realise they actually do not desire (leading to immediate self-correcting behaviour), it is proposed that rogue 'goals' must have momentarily dominated.

Kruglanski and Szumowska argued that goals can be elicited automatically by the environment, triggering unwanted behaviours that need to be immediately self-corrected. The authors give the example of turning left at an intersection and driving home rather than driving to the mall, a habitual response that is immediately corrected by turning around and heading to the mall as originally planned. While associative learning theorists would argue that the familiar intersection directly triggered the 'turning left' response without evaluation of the consequence, Kruglanski and Szumowska proposed that the 'goal' of driving home was elicited by the context. The fact that the goal was actually undesirable means that it is hard to see how the definition of goal provided here can be 'synonymous with reward' (Kruglanski & Szumowska, 2020, p. 1267). Presumably, the authors would argue that for a split second, the goal of driving home *was* desired, hence triggering the associated response which was immediately corrected. The actor thought that they wanted the goal because it had been desired many times in the past (in other words, a 'goal slip' occurred rather than an 'action slip').

Many of these theories (Aarts & Dijksterhuis, 2000; Du et al., 2021; Hogarth & Troisi, 2015; Kruglanski & Szumowska, 2020; Papies & Aarts, 2011) extend the concept of goals from 'outcomes that are currently desired' to 'outcomes that have been desired frequently in the past'. Some theories posit that goals can be elicited outside of cognition (e.g. automatically elicited by contexts) and that goal representations can be elicited outside of desire for that outcome. These types of accounts have been criticised as unfalsifiable (Luque & Molinero, 2020; Wood et al., 2022) and in conflict with associative dual-process models where outcomes need to be *evaluated* if they are to be labelled as goals with the potential to influence behaviour (Dickinson, 1985). Ultimately, however, the concept of habit as being behaviour which is fluidly elicited with minimal constraints from expected outcome value is common to all models.

If all behaviours are goal-directed or under S:R-O associations, with identical associative structures for habits as for goal-directed behaviour, we should expect largely overlapping brain networks to mediate both types of behaviours. There is some evidence against this notion. First, studies from the field of behavioural neuroscience have demonstrated in non-human animals that dissociable brain networks are seen to largely support devaluation-sensitive (goal-directed) vs. devaluation-insensitive (habitual) behaviour. For example, manipulations that lead to loss of function in the dorsomedial striatum shift behaviour to be immediately habitual, whereas damage to the dorsolateral striatum means that devaluation-insensitive habits do not develop (Yin et al., 2004, 2005; see for review: Balleine & O'Doherty, 2010). Similarly, this distinction has been largely replicated in human fMRI studies (Friedman & Robbins, 2022; Guida et al., 2022; Patterson & Knowlton, 2018; Watson et al., 2018). However, recent studies have suggested that the dissociation between habit and goal-directed regions may not be as clear-cut as previously assumed (Shipman et al., 2018; Vandaele et al., 2019; see for discussion: Turner & Parkes, 2020), highlighting that future developments to these neurobiological models may be forthcoming.

3.4 Summary

Various hierarchical models argue that the outcome needs to be represented in the habit structure. This contrasts with the associative dual-process model which proposes that stimuli can trigger responses directly, without representation of the associated outcome. Providing direct evidence for S-R links has proven to be difficult experimentally and the onus is on associative learning theorists to provide convincing evidence that stimuli can directly trigger behaviour in an S-R manner, without respect to the anticipated outcomes.

While this theoretical difference may seem significant, the habitual behaviour that is ultimately produced by these various models looks very similar. This is because the current desirability of outcomes may not be represented, or is represented very weakly, in all models. Outcomes are proposed to have many properties including, but not limited to, hedonic and motivational properties, sensory properties, and temporal characteristics (Delamater & Oakeshott, 2007). Many hierarchical models assume that following extended training, outcomes and their associated responses can be automatically primed and that current outcome value does not necessarily constrain this process. On this latter point, all the models examined here are broadly in agreement, suggesting that future experimental work may lead to concord across all fields on the representational structure of habits.

Acknowledgements Dr Poppy Watson was supported by an Australian Research Council Discovery Early Career Researcher Award (DE200100591) during the preparation of this manuscript.

References

Aarts, H., & Dijksterhuis, A. (2000). Habits as knowledge structures: Automaticity in goal-directed behavior. *Journal of Personality and Social Psychology, 78*(1), 53–63.

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