The Journal of Mind and Behavior
Summer and Autumn 2021, Volume 42, Numbers 3 and 4
Pages 317–346
ISSN 0271–0137

Mechanisms of Unconscious Thought: Capacities and Limits

Adrian P. Banks

University of Surrey

Unconscious thought has been linked with a wide range of mechanisms, capacities, and limits. These claims have changed over time and across different domains of thought. The aim of this review is to synthesise the research on unconscious thinking across the domains of reasoning, judgment, decision making, insight problem solving, and creativity and identify the commonalities between them. Three mechanisms underpin unconscious thought in all of these domains: automaticity, reward-based association, and spreading activation. The mechanisms are triggered by cues in the environment or internal states, and the output of the mechanisms are either specific outputs or affective responses. The mechanisms also define the limits of unconscious thought, expressed here as a "principle of integration": unconscious thought is not sufficient in tasks or problems that require concepts to be integrated in novel or unfamiliar ways. Where theories have made stronger claims for unconscious thought than this, analysis of the evidence supporting those theories proves equivocal. Nonetheless, unconscious thought based on these mechanisms is adaptive in frequently encountered situations and provides the capacity for highly effective thinking across a range of domains.

Keywords: unconscious thought, automaticity, reward-based association

The mechanisms and capacities of unconscious thought have long been debated with the pendulum often swinging from a view of the unconscious as capable of complex thought through to rather limited capacities and back again. Freud's influential theories have provided lasting support for the significance of the unconscious and contemporary psychodynamic theory still ascribes a wide range of thought to the unconscious with a particular emphasis on affect and motivation (e.g., Westen, 1998). Some experimental psychologists agree that the unconscious is capable of complex thought. For example, the seminal study of Nisbett and Wilson (1977) presented evidence that people have little or no

---

I would like to thank Adam Harris, Jonathan Fugelsang, Jane Ogden, and Peter Hegarty. Correspondence concerning this article should be addressed to Dr Adrian Banks, School of Psychology, University of Surrey, Guildford, Surrey, United Kingdom. Email: a.banks@surrey.ac.uk

introspective access to their thinking; it occurs without conscious awareness. Both Djiksterhuis and Nordgren (2006) and Hassin (2013) propose that the strict limitations on cognitive processing for conscious thought (e.g., Baddeley, 2007) do not seem sufficient for the number of complex decisions required daily, and therefore some of these must be carried out unconsciously.

In contrast, other experimental psychologists hold a much narrower view of the unconscious. For example, Simon (1987) argues that good judgement is simply analyses frozen into habit with the capacity for rapid response through recognition. Greenwald (1992) also proposes that unconscious cognition is severely limited in analytic capability, restricted to nothing more sophisticated than analysis of partial meanings of single words.

The diverging positions have persisted for both methodological and theoretical reasons. Methodologically, unconscious thought is difficult to test and different conclusions have been drawn based on different evaluations of the strength of the evidence available (e.g., Hesselmann and Moors, 2015; Moors and Hesselmann, 2017; Newell and Shanks, 2014). However, the aim of this review is theoretical rather than methodological. Rather than viewing unconscious thought as a black box into which all cognition that is not conscious is placed, I approach the unconscious as a set of several distinct mechanisms each of which has defined properties. The capacities and limits of unconscious thought are circumscribed by the mechanisms that generate it. By specifying what the mechanisms are, more detailed and testable ideas will become clear about how unconscious cognition operates, what tasks can be performed by unconscious cognition and which cannot.

The aim of this review is therefore to specify what these mechanisms are, focusing on cognitive rather than neural mechanisms. To do this I will review theories of unconscious thought from the domains of reasoning, judgment, decision making, insight problem solving, and creativity to identify the common mechanisms that are used in each of these areas and combine them to form a broad account of unconscious thought. I will present a model that synthesizes the theories from these areas and identifies the commonalities in the mechanisms of unconscious thought and the limits of these mechanisms.

Overview of the Review

Across the domains of reasoning, judgment, decision making, insight problem solving, and creativity, different terms and theories are used to describe similar mechanisms of unconscious thought. This review will bring these theories together and identify the underlying, fundamental mechanisms of unconscious thought that are common to all of these domains. I identify three mechanisms that are used to explain how prior experience drives unconscious thought: automaticity; reward-based association; and spreading activation.

First, the review outlines the theories and evidence for these three mechanisms. For each mechanism, the defining features of the mechanism are described that are common to all of the theories of that mechanism and differentiate it from the other mechanisms. The main theories used to explain the mechanism across different domains are described including how the relevant cognition is acquired and learnt from prior experience. Then the section looks at evidence demonstrating how the mechanism is deployed in the different areas of thinking that have been studied. Second, I describe how the mechanisms are used in a cue–mechanism–output process to generate an unconscious thought. Finally, the mechanisms have limits that determine what they can be applied to. That is, there are certain features that when present in a problem require conscious thought. I characterise the main limit of unconscious thought that recurs across the different domains as "the principle of integration." Unconscious thought is not sufficient in tasks or problems that require concepts to be integrated in novel or unfamiliar ways. I discuss the experimental evidence for these mechanisms and evaluate the strength of the evidence. It is possible that other mechanisms and limits could be identified in future; however, this review presents those that are currently dominant across the domains of thinking.

Consciousness, Attention, and Unconscious Thought — Defining Terms

Consciousness has been defined in multiple ways. Block (1995) differentiates between phenomenal consciousness and access consciousness. Phenomenal consciousness refers to feelings and sensations, such as pain, that a person might be aware of. This level of awareness is shared with many animals (e.g., Panskepp, 2005). Access consciousness refers to a conscious awareness that can be referred to and reflected upon. It is typically defined as a reportable inner state: conscious content that can be reported verbally or through an intended gesture such as a key press in an experiment (e.g., Baumeister, Masicampo, and Vohs, 2011; Dehaene and Changeux, 2011). It is access consciousness that is the main focus of studies of thinking. Studies of unconscious thinking therefore seek to examine thought processes whilst demonstrating an absence of this conscious awareness.

Theoretically, several roles for conscious awareness have been proposed. These include constructing meaningful sequences of thought, simulating future possible events, and facilitating social and cultural actions and communication (Baumeister and Masicampo, 2010). Within global workspace theory, conscious content is linked with working memory (Baars and Franklin, 2003).

Theories of attention are relevant to understanding unconscious thought; allocating attention to a task or stimuli is commonly associated with conscious awareness. Some theories view attention as a requirement for consciousness, for example allowing access to the global workspace which enables consciousness (Baars, 1997). Other theories propose that attention is required for access consciousness, if not other types (Block, 1995). Experimentally, loading attentional capacity with a secondary task has been shown to raise the threshold for conscious awareness of stimuli (Macdonald and Lavie, 2008). However, whilst conscious awareness may typically be associated with the allocation of attention, this does not mean that unconsciousness requires the absence of attention. There is not necessarily a one-to-one relation between attention and conscious awareness (Dijksterhuis and Aarts, 2010). Instead, it has been proposed that increasing attention moderates the quality of stimulus representations by enhancing or boosting processing of the stimuli (Dehaene and Naccache, 2001; Kiefer, 2012). Below a lower threshold stimuli are not processed. Above a higher threshold stimuli are processed consciously. Between these two thresholds is an intermediate level of stimulus representation and attention in which stimuli are processed unconsciously (Moors, 2016). That is, stimuli are allocated sufficient attention to be processed but not enough to reach conscious awareness. Cleeremans (2005) presents a similar idea of an underlying dimension of representation quality which, above a certain threshold, becomes conscious. A gradual increase in attention need not lead to a gradual increase in conscious awareness. It could lead to a qualitative shift from unconscious to conscious processing across a threshold much as a gradual increase in temperature results in a sudden qualitative change from ice to water. This metaphor aptly describes the neural response of "global ignition" described within the global neuronal workspace theory in which a large prefronto–parietal cortical network is activated that enables conscious access by making information globally available to multiple brain systems (Dehaene and Changeux, 2011; Dehaene, Kerszberg, and Changeux, 2001).

In this review I do not seek to address many aspects of the complex debate surrounding consciousness as the focus here is on studies that examine thinking without conscious awareness. Pragmatically, this involves assessing the effectiveness of methodological techniques that prevent conscious deliberation. Many methods have been used to do this, but they can be divided into two broad approaches. The first approach uses psychophysical techniques to present the task or stimuli without conscious awareness. The most common technique in studies of thinking is masked priming in which stimuli are presented for very brief periods of time, preceded and followed by irrelevant stimuli that render the experimental stimuli invisible (e.g., Breitmeyer and Öğmen, 2006). Other techniques are used such as continuous flash suppression in which the presentation of a rapidly changing stimulus such as repeated Mondrian-like patterns of multi-coloured squares to one eye act as a mask and dominate conscious awareness, rendering the presentation of a static experimental stimulus to the other eye invisible (e.g., Tsuchiya and Koch, 2005). The second approach uses typical experimental cognitive psychology techniques to prevent conscious thought. That is, the experimental stimuli are consciously perceived but the cognitive processes involved in thinking about the problem or making the decision are not.

Common techniques in studies of thinking are the use of a concurrent working memory load to block conscious cognitive work prior to the generation of a solution, and a fast response deadline to prevent conscious thought by ensuring there is insufficient time to engage in conscious cognition.

In summary, I have drawn on work connecting attention and consciousness to support the idea of a gradient of attention that can be allocated to a task and a threshold above which a cognitive process is conscious and below which it is unconscious. Thus unconscious processes may still be allocated some attention, but beneath the threshold for conscious awareness.

Mechanisms of Unconscious Thought

Automaticity

The first mechanism that is used to explain unconscious thought across a range of domains is automaticity. The concept of automaticity is complex and different theories propose different features as central to the concept. These include efficiency i.e., the process requires no attention (Shiffrin, 1988), and autonomy i.e., the process runs to completion without conscious guidance (Bargh, 1992). Some models propose more complex sets of features e.g., processes that are unintentional, goal independent, uncontrolled, autonomous, unconscious, efficient, and fast, and these must be examined separately (Moors and De Houwer, 2006). Whilst unconscious processing or lack of conscious awareness is rarely a defining feature of automaticity in these models, it is often an associated feature of automatic cognitive processing. Unconscious automatic processes are therefore likely to have other features, such as being efficient, autonomous, and fast, but current models do not typically require all of these features (Moors, 2016). Overall, the defining feature of an automatic cognitive process is efficiency. Through repetition, a cognitive process can be completed increasingly quickly and with low effort which is an adaptive response to frequently encountered sequences of cognition.

Three main theories have been proposed to explain how automaticity develops. Logan's (1988) instance theory explains automaticity rather narrowly in terms of direct memory retrievals. This theory proposes that initially problems are solved through nonautomatic application of a multi-step algorithm. The outputs of the algorithm are stored in memory until, after practice, the output can be directly retrieved when needed rather than calculated using the algorithm. Anderson's (1992) ACT* theory also proposes that problems are solved initially through nonautomatic application of an algorithm, but that with practice the algorithm is strengthened and becomes automatic. Hence, a difference between the theories is that instance theory explains the retrieval of specific facts whereas ACT* (and its subsequent development ACT-R; Anderson et al., 2004) explains the use of rules that can be applied to either the same or different stimuli. It is possible that both of these processes can occur, depending on the structure of the task (Taatgen and Wallach, 2002). Chein and Schneider's (2012) triachic theory of learning describes automaticity as a property of a representation system in which knowledge is slowly acquired through associative learning. Through interacting with the environment, key input–output relations from sensory, perceptual, and motor systems are slowly strengthened through Hebbian learning until they are sufficiently strong and stable that the cognitive control disengages from executing a task and task performance becomes automatic. These theories imply a gradual continuum of automaticity in which processes become increasingly automatic with practice rather than a dichotomy in which processes are either automatic or not. According to the model of conscious awareness used here, conscious awareness of these processes would depend on whether the level of attention required by the process was above the threshold of conscious awareness, or whether it was sufficiently automatic that the level of attention required falls below the threshold.

Automaticity in Unconscious Thought

Automaticity and reasoning. Automaticity in reasoning has recently returned as a focus of research because of the role that it has been proposed to play in parallel dual process theories. These are theories of reasoning that propose an initial heuristic or intuitive phase in which more than one response to a problem can be generated simultaneously (De Neys, 2012; Handley and Trippas, 2015; Pennycook, Fugelsang, and Koehler, 2012). It has been shown for some time that these initial responses could include heuristics such as choosing a stereotypical response (Kahneman and Tversky, 1973) or a believable response (Evans, Barston, and Pollard, 1983). On problems where these responses conflict with the logically or probabilistically correct solutions then the heuristics lead to an error such as belief bias. That is, fast and automatic reasoning is biased whereas logic requires deliberate, analytic thought. More surprising is recent empirical evidence showing that logically or probabilistically correct responses may be generated automatically and in parallel with the heuristic responses. That is, automatic responses may be logical and unbiased. These include task instruction manipulation studies in which a task of assessing the believability of statements (Handley, Newstead, and Trippas, 2011), the likeability of statements (Morsanyi and Handley, 2012), and the physical brightness of text on a computer screen (Trippas, Handley, Verde, and Morsanyi, 2016) were influenced by the logical structure of arguments embedded in the text. Participants appear to detect when problems have two potential responses that conflict, a logical and a non-logical response, by slower response times and increased skin conductance (De Neys, Moyens, and Vansteenwegen, 2010) even though they make no mention of this in concurrent verbal protocols (De Neys and Glumicic, 2008). The interaction of both processes can be very fast (Banks and Hope, 2014). Hence, logical reasoning can be automatic.

A theoretical approach that has been used to explain automatic reasoning is mental logic (e.g., Braine, Reiser, and Rumain, 1984 and Rips, 1994). Braine et al. proposed a theory of propositional logic reasoning in which people hold a set of inference schemata that spontaneously generate a conclusion when presented with premises that match their logical form. This happens directly for simpler logical forms such as modus ponens (if p then q, p, therefore q) and or elimination (p or q, not p, therefore q). More complex inferences occur indirectly through several steps as there is no inference schema that directly matches their form. The direct reasoning inference schemata are a possible process for automatic logical reasoning.

There is some evidence that these inferences are spontaneous, that is, displaying the unintentional feature of automaticity. Earlier work tested the spontaneity of inferences by embedding the logical forms within text and seeking evidence that inferences are drawn. For example, Lea (1995) presented participants with the following text:

> Tony was trying desperately to stick to his diet. “Well,” said his mother Maria, “You can have either bread or corn flakes with breakfast.” Tony seemed to take forever before giving an answer. “Alright Ma, since I have to skip something, I won’t have the corn flakes.” (p. 1471)

Participants may represent the text as "bread or corn flakes" and "not corn flakes." If they spontaneously apply the logical rule of inference “Or Elimination,” they will draw the conclusion “bread.” Participants were faster to name this word in a subsequent display, and faster to respond in a lexical decision task to a related word (butter). Participants also falsely recognised information as having been presented in the text when in fact it needed to be inferred (Lea, O'Brien, Fisch, Noveck, and Braine, 1990). Rader and Sloutsky (2002) embedded modus ponens and affirmation of the consequent (if p then q, q, therefore p) arguments in text and also showed that participants erroneously believed that they had read the concluding word of the argument when in fact it needed to be inferred. But the authors use this as evidence against inference schemata as affirmation of the consequent is not claimed to be a direct inference. The same question is raised by Handley et al.’s (2011) experiments on syllogistic reasoning. They show that the logical structure of syllogisms influences belief judgments even without instructions to reason logically, suggesting a spontaneous inference is drawn. But no direct reasoning schema for syllogisms has been proposed, leaving open the question of what the process is that explains how this inference is drawn.

Overall, these studies of embedded logical arguments show that the simple inferences (and possibly more complex ones) are relatively easy to draw, but are they automatic? It is possible that although there is no requirement to draw the inference, participants may well be aware of the logical structure and are intentionally drawing the inference as they reflect upon the text. There is no direct evidence about their lack of intention, conscious awareness, control or other features of automatic processes. Reverberi, Pischedda, Burigo, and Cherubini (2012) sought direct evidence by presenting a logical argument about a number, followed by subliminal masked primed presentation of a minor premise, then followed by a target number judgment task. When the target matched the conclusion that was primed by logical inference, the evaluation of the target was faster. But this priming effect was only found for modus ponens arguments; none of the more complex arguments demonstrated any priming. Therefore, so far, there is evidence that reasoning can have some features of automaticity — some inferences are fast and efficient — but there is less evidence that these inferences are unintentional, uncontrolled, and autonomous. Only modus ponens has direct evidence for being an unconscious, automatic logical inference.

Automaticity and judgment. By describing intuition as nothing more than recognition, Simon (1992) was making the claim that the only process of unconscious thought was an automatic association between a stimulus and a response. An example of this is Klein's (1993) theory of recognition-primed decision making. This theory describes how experts make fast decisions without evaluating decision options. The most frequent type of decision involves the expert decision maker matching the current situation to one in memory based on the similarity of goals and perceptual cues, i.e., recognising it, and the retrieving expectancies and a course of action for that situation. For example, in a study of fire officers, one fire officer saw a burning billboard on top of a house, recalled a time when a burning billboard had collapsed and fallen off the house, and ordered the crowd to move back (Klein, Calderwood, and Clinton–Cirrocco, 1986). However, as this theory is based on data gathered using retrospective interviews it is not certain that the decisions were made without conscious deliberation — participants may simply have forgotten the rapid decision process that occurred.

Experimental work conducted in several domains such as medicine (Crandall and Getchell–Reiter, 1993; Schmidt, Norman, and Boshuizen, 1990) and chess (Chase and Simon, 1973) have led to similar theories emphasising the recognition of patterns linked to outcomes. Studies of expertise are relevant as experts have engaged in sufficient practice for processes to become automatic (e.g., Ericsson, Krampe, and Tesch–Römer, 1993) unlike studies where judgments may occur with limited conscious awareness of their origin but without sufficient repetition to become automated (e.g., Betsch, Plessner, Schwieren, and Gütig, 2001). Chess expertise involves a combination of (slow) search and (fast) recognition, but it is the recognition of patterns of pieces and possible moves that primarily underpins expertise (Chase and Simon, 1973; De Groot, 1946, 1965). Chess experts can largely maintain their level of performance under considerable time pressure, supporting the role played by fast recognition (Gobet and Simon, 1996). Further evidence suggests that this might be an automatic process. Reingold, Charness, Schultetus, and Stampe (2001) used a modified chess task to assess expert and novice speed at detecting checks for a cued piece. There were additional non-cued distractor pieces that were to be ignored. Experts but not intermediates or novices were slowed by distractor pieces that cued an incongruent response (i.e., they were checking the king when the cued piece was not). They propose this Stroop-like (or flanker-like) interference indicates that the distractor piece automatically generated a response that was inhibited, slowing the overall response. Thus recognition and response to a checking pattern occurs in highly practiced participants quickly and without control, characteristic features of automaticity.

A comparison of automaticity in reasoning and judgment suggests that these domains may rely on different processes for developing automaticity. Some automatic judgments can be explained with Logan's instance theory. At first, an analysis might be applied consciously to a set of cues leading to a final judgment, and this pairing is stored in memory. Over repeated practice with that set of cues the judgment can be retrieved directly without the need for conscious analysis. For example, a chess player may initially analyse positions in practice and find the optimal move but over time recognise and retrieve from memory the relevant move. Alternatively, there may be no need for a conscious algorithm to generate the judgment initially. Cues in the environment may frequently be paired with outcomes and these associations are stored in a representation system and they guide task performance, as proposed by Chein and Schneider. In both of these cases, the judgment is context dependent as a specific input is tied to a specific output.

In contrast, theories of automaticity in reasoning rely on rules of inference, such as modus ponens. These rules differ from judgments because they can be applied to novel content, effectively containing variables that can be bound to information. Therefore they cannot be explained by learnt associations or direct memory retrievals. The rules are better explained by algorithm strengthening (Anderson, 1992). But where do these rules come from? One possibility is that rules are introduced through formal education, or knowledge that is shared socially or culturally. A second possibility is that general rules are learnt from specific examples, e.g., a child is told by her parent "If you finish your spinach then you can have pudding" — and over time the child learns modus ponens. Acquiring automated rules from examples was demonstrated by Anderson, Fincham, and Douglass (1997) who asked participants to memorise facts such as "Skydiving was practiced on Saturday at 5pm and Monday at 4pm." All of the examples followed the same underlying abstract rule that skydiving was two days later and one hour earlier (that is, forty-seven hours later). After several days of practice, participants became faster at applying the rules to new examples, showing that they could be applied more generally. Also, the rules were applied faster in the direction that the rule was learnt suggesting that the association was learnt as a rule rather than simply as the pairing of two items in memory.

Overall, there is evidence in the domains of reasoning and judgment that these thinking processes become automated, and some evidence that they can be applied unconsciously. However there are differences in the processes of deploying automaticity in these domains. Automatic judgments can be explained by context dependent associations such as instance theory or associative learning. Automatic reasoning is not context specific and so is better explained by the development of automatic rules.

Reward-Based Association

The second mechanism that is used to explain unconscious thought is the reward-based associations acquired through reward-based learning. Current theories of reward-based learning incorporate the main insight of Thorndike's law of effect: actions that are rewarded are more likely to be repeated in that situation whereas actions that are punished are less likely to be repeated (Thorndike, 1911). Behaviourist research on operant (instrumental) conditioning further explored this fundamental law and identified the effect of different schedules of reinforcement etc. (e.g., Skinner, 1963). Contemporary reinforcement learning algorithms provide a more precise account of how learning about rewards and punishments through trial and error leads to optimal decision making in the long term (e.g., Daw, Niv, and Dayan, 2005). The defining feature of reward-based associations acquired through reward-based learning is an association between a cue and a positive or negative outcome. Through repeated pairings a cue is associated with an outcome that leads to a reward or a punishment. Once these associations are learnt, the cues guide thinking as they indicate positive or negative outcomes and generate positive or negative emotions.

There are two broad classes of reinforcement learning methods, model-based and model-free algorithms (Dayan and Niv, 2008). These have been equated to goal-directed and habitual behaviour (de Wit, Corlett, Aitken, Dickinson, and Fletcher, 2009). Model-based algorithms build an internal model of all the states in an environment, the possible actions in each state, and the reward associated with each state. The model can then be used to achieve novel goals by predicting the likely outcome of different plans based on the values in the model and select the optimal sequence of actions required to attain the desired outcome. Model-based methods are also flexible. If the reward for any given state changes, then new plans can be developed immediately reflecting the change. However, the methods are also computationally costly as the problem space rapidly becomes large as the number of states and actions grows.

In contrast, model-free reinforcement learning algorithms are computationally undemanding. There is no internal model of the environment and no planning. Instead, much like the law of effect, the outcome of each sequence of actions is stored (or *cached*). In a given environment many different sequences of actions are explored through trial and error, slowly refining the cached value of each sequence until in the long run these become accurate estimates of the alternatives.

Making the optimal choice in a given situation is now computationally simple — select the sequence of actions that has the highest cached value. This simplicity comes at the cost of inflexibility. The cached values can only be changed slowly through trial and error. Therefore if the situation changes quickly, the values are no longer accurate guides to the optimal choice. Choices will be made based on the previous situation that are inappropriate for the current situation. That is, we may still pursue habitual behaviours even when they are no longer rewarding (Tricomi, Balleinie, and O'Doherty, 2009).

One of the main methods for refining the values of each option is the temporal difference learning algorithm (Sutton, 1988). These algorithms update the value of each action based on the discrepancy between the expected reward and the observed reward so that over time the expected reward becomes an accurate estimate of the value of that action. The algorithm also ensures that values of all of the actions in a sequence leading to an outcome are updated appropriately, not just the final step that brings the reward. This means that an optimal path can be selected simply by choosing the best action each step of the way and therefore no long range planning is required. The psychological relevance of these algorithms was affirmed by the striking finding that phasic dopamine responses during appetitive conditioning in the striatum have similar properties to the prediction error in temporal difference learning, suggesting that model-free reinforcement learning may provide an accurate account of habit formation (O'Doherty, Dayan, Friston, Critchley, and Dolan, 2003). Whilst temporal difference learning algorithms are neutral with regard to conscious or unconscious thinking in humans, model-free reinforcement learning has been proposed as a theory for learning how optimal decisions can be made without the need for conscious representations (Shea and Frith, 2016).

Reward-Based Associations in Unconscious Thought

Habits. Habits have been equated with unconscious thought (Lisman and Sternberg, 2013). Research on habits has been a rich source of naturalistic studies demonstrating how reward-based learning drives decision making by guiding towards choices with rewarding associations. Daily life provides ample opportunity for the development of habits as common scenarios are encountered and typical choices are made repeatedly. In a diary study in which participants provided hourly records of their thoughts and behaviours, around a third were performed every day in the same location, indicating they may be habitual or open to habit formation (Wood, Quinn, and Kashy, 2002). Furthermore, around 50% of the behaviours were performed without conscious guidance (although this was a self report measure and open to the limitation that conscious thought may have been present but forgotten). Repeated, rewarded actions lead to habit formation. A study of new gym members found that exercising four times per week for six weeks was necessary to form an exercise habit, and that positive feelings about exercise also predicted habit formation (Kaushal and Rhodes, 2015). Contextual cues aligned with the situation become associated with the reward and so become part of the habitual behaviour. In a cinema, habitual popcorn eaters ate as much unappealing stale popcorn as fresh popcorn, despite rating the stale popcorn as less attractive, whereas non-habitual popcorn eaters ate much less stale popcorn than fresh. But when participants were tested in a neutral meeting room both habitual and non-habitual popcorn eaters ate less stale popcorn. That is, in the habitual context of the cinema, participants with a popcorn habit were less sensitive to the poor reward and ate the stale popcorn based on learnt habits (the cached value of eating popcorn) rather than current taste, but this habit did not generalise to a novel context (Neal, Wood, Wu, and Kurlander, 2011).

Habitual responses can also be primed in naturalistic settings without conscious awareness. In a study to assess the priming effect of advertisements, adults watched a television programme interrupted by two commercial breaks. These breaks contained advertisements about either unhealthy snacks (e.g., candy), nutrition (e.g., orange juice), or non-food controls. In a second apparently unrelated experiment, participants were asked to taste and rate various foods. The amount of food eaten was recorded. Participants who had viewed snack ads ate more than those who had viewed the other ads. Debriefing interviews suggested that participants were unaware that the advertisements related to the aims of the experiment (if not, they were removed from the sample). Presenting attractive snack food images triggered habitual snack food eating without participants' conscious awareness (Harris, Bargh, and Brownell, 2009). Overall, naturalistic studies of habit formation show that repeated pairing of an action with a rewarding outcome increases the likelihood of choosing that action in future, without apparent conscious deliberation about that choice.

Model-free reinforcement learning. Whilst naturalistic studies provide valuable illustrations of habitual actions, they rely on self-report measures or interview techniques to assess the level of conscious awareness and thought. It is possible that the level of conscious awareness was greater in these studies than participants reported. More rigorous experimental techniques are required to control for this possibility. More controlled studies are also required to allow precise measurement in order to test the quantitative predictions of model-free reinforcement learning algorithms as a process for habitual or unconscious reward-based learning.

Pessiglione, Petrovic, Daunizeau, Palminteri, Dolan, and Frith (2008) tested how effectively participants were able to learn associations between subliminal cues and rewards and punishments. The abstract visual cues were masked and presented briefly for a time calculated in a pretest such that participants were not able to discriminate between cues. Three visual cues were used during the study, associated with a reward (winning £1), no reward (£0), or a punishment (losing £1). Over repeated trials, participants were able to learn to respond to the rewarding cues and avoid responding to the punishing cues. Their responses were modelled with a Q-learning algorithm, typical of model-free reinforcement algorithms, and this model provided a good fit to the slow learning of the values associated with the cues. This study provides good evidence for unconscious model-free reinforcement learning.

If model-free reinforcement learning is a habitual process, then a second prediction is that this learning will occur automatically, requiring limited cognitive resources. Otto, Gershman, Markman, and Daw (2013) differentiated between model-free and model-based reinforcement learning using a two-stage choice task. In this task a first stage choice is made between two fractal patterns which led to a second stage choice between two further fractals that led either to a reward or not. The probability of moving between these options varies and the task is to learn the choices that have the greatest reward. A model-free strategy would result in participants learning to choose the first stage choice that was ultimately rewarded in the second stage, regardless of changing probabilities of each stage because the model simply learns the sequence that is rewarded. The more sophisticated model-based strategy would result in participants learning the probabilities of the different transitions between stages and choosing their first choice in order to reach the most rewarding second stage options, even if that means choosing the first stage option that was not previously rewarded. Participants completed the task with or without a concurrent cognitively demanding secondary task, and those participants with the secondary task relied more on the model-free strategy. This suggests that model-free reinforcement learning requires fewer cognitive resources than model-based reinforcement learning.

Reward-based associations and reasoning. The majority of studies of reasoning follow the paradigm outlined above in which problems are described to participants who draw inferences from the premises. Participants are rarely provided feedback during the experiment, the process of learning to reason is rarely tested, and there is rarely a reward or payoff structure associated with different outcomes despite all of these being relevant factors in shaping the reasoning processes that people use naturalistically and bring with them to the lab. However, a largely unconnected literature on transitive inference does examine how the relations between facts are learnt through reward and new relations between them inferred. These studies typically use a narrow experimental paradigm in which the linear relation of five facts is learnt in pairs: A<B, B<C, C<D, D<E. For each pair, one item is associated with being the correct answer through repeated choices. This training is applied to all pairs until relations between all of them are learnt: A+B−, B+C−, C+D−, D+E− where "+" indicates a correct or rewarded choice and "−" indicates no reward. Thus the order is learnt by reward-based associations of one item within each pair. Participants are then tested on the relation between B and D, which was not explicitly trained. Successful performance on this task has been shown in humans (e.g., Bryant and

Trabasso, 1971) and also non-human animals such as rats (Davis, 1992) and pigeons (Weaver, Steirn, and Zentall, 1997). It is possible to complete this task by consciously learning the overall linear relationship between all five items. However, it has also been shown that participants are able to reason transitively without awareness of the overall relationship between items (Frank, Rudy, Levy, and O'Reilly, 2005) suggesting that the associative strengths are sufficient to support the inferences. Further research has replicated this effect and ruled out the possibility that this effect is a pseudoinference based on the differential strength of rewards associated with the test items (Leo and Greene, 2008). Thus reward-based learning of associations between facts can also be used as a process for reasoning without conscious awareness.

Affect may be conscious but the transition from a situation to a feeling may not be (e.g., Johnson–Laird, Mancini, and Gangemi, 2006). Nor are conscious reasons necessary for choosing an option. Zajonc (1980) has argued that "preferences need no inferences." People simply choose the option that they like and judgments of liking or disliking can occur independently and before the cognitive thought process. Justification of the choice may come later, but these reasons are not necessarily used to make the decision.

Damasio's (1996) somatic marker hypothesis connects theories of reward-based learning with affect to explain unconscious thought. Damasio suggests that as a person makes a series of decisions in an environment, they learn about different outcomes. These outcomes, positive and negative, are associated with different physiological effects which are emotional reactions to the outcome. These patterns of physiological responses are somatic markers. When the situation is encountered again, the somatic marker is activated and evokes a positive or negative feeling based on whether the previously learnt association was good or bad and whether this feeling guides the choice. A major line of evidence for this theory has been developed using the Iowa Gambling Task (Bechara, Damasio, Tranel, and Damasio, 1997). In this task participants make repeated choices from four decks of cards that each pay out different amounts of money. Some of the cards have gains and some have losses but in the long run two of the decks give a net positive outcome through a series of low risk small payouts but small penalties and two have a net negative outcome through high risk occasional large payouts but also large penalties. Participants chose cards from the good decks before they were aware that this was the best choice. Participants also showed anticipatory physiological responses (that is, increased skin conductance responses) when considering the bad, risky options — suggesting that the somatic marker was activated. These physiological cues provide information about the quality of the choice. However, this theory and the empirical work to support it is controversial and has been criticised on several grounds (Dunn, Dalgleish, and Lawrence, 2006).

A less controversial theory that describes the process through which affect influences thinking is the affect heuristic (Slovic, Finucane, Peters, and MacGregor, 2004). This theory also proposes that associations are learnt between events and positive or negative outcomes. As a result, the representations of those events are tagged with positive or negative affect. Judgments are made by simply pooling all the positive and negative tags that are associated with the representation and the overall judgment of liking influences the choice. This could be done consciously or unconsciously. Finucane, Alhakami, Slovic, and Johnson (2000) found that use of the affect heuristic varies according to the circumstances but was more common under time pressure as an affective response to a problem is generated quickly and automatically, unlike more complex analyses.

Overall, there is evidence that people learn associations between choices and rewarding or punishing outcomes. They are then able to draw on these associations in future choices. There is also evidence that they are able to do this without conscious awareness of the basis of their preferences. There are several processes for forming these associations. Specific behaviours or actions and their context may be directly associated with a reward so that the behaviour is reinforced and repeated in that context without conscious awareness. Alternatively, a more indirect process is that positive or negative outcomes may be represented through positive or negative affect. Future choices may feel good or bad as a result of these learnt associations and guide choices by selecting options that feel good. These feelings may arise without conscious awareness of their origin.

Spreading Activation

The third mechanism that is used to explain unconscious thought is spreading activation. Spreading activation assumes that semantic knowledge is stored in a network in which related concepts and facts are connected to each other through repeated associations that have been learnt over time. As one concept becomes the focus of attention, activation is spread to related concepts, raising their salience (Collins and Loftus, 1975). The level of activation of a fact determines its probability of retrieval from memory (Anderson, 1983). Spreading activation can be demonstrated in a semantic priming task. In this task the speed of identifying a target letter string as a word rather than as a non-word is faster if it is preceded by a word from the associative network than an unrelated word, e.g., Bird–Robin is faster than Bird–Arm (Neely, 1977). This effect can be found even when the words are masked so that they are not consciously perceived (Dehaene et al., 1998), demonstrating the potential for unconscious spreading activation in thinking. The defining feature of spreading activation is the connection of concepts into a network. Rather than a narrow set of cues and specific outcomes, the connections between concepts means that a wider range of concepts can influence thinking.

Spreading Activation in Unconscious Thought

Spreading activation and creativity. A long standing phenomenon of unconscious thought is incubation (Wallas, 1926). There are many anecdotal examples of incubation in which after a period of unsuccessful, hard deliberation the thinker is engaged in an unrelated task when suddenly a creative or insightful solution springs to mind apparently without conscious consideration. Incubation is the time seemingly not engaged in consciously thinking about the problem when a solution is found without deliberate thought. Sio and Ormerod's (2009) meta-analysis found evidence for a positive incubation effect, especially for divergent thinking tasks.

The alternative uses task is commonly used to study divergent thinking and creativity (Guilford, 1967). In this task participants are required to generate as many uses for a common object as possible. Participants use several different strategies such as directly retrieving alternative uses from memory or using properties of the object as a cue to retrieve alternatives (Gilhooly, Fioratou, Anthony, and Wynn, 2007). Retrieval from memory of alternative uses was explained with spreading activation from the common use to alternatives whereas the other strategies involve greater executive load. Gilhooly, Georgiou, Garrison, Reston, and Sirota (2012) investigated incubation in this task by requiring participants to complete a secondary task for a period of time before responding, simulating a period of incubation. Performance was best when incubation began immediately after the task was presented, before conscious thought had begun, suggesting that some beneficial unconscious thought occurred during this incubation period. Gilhooly proposed that the mechanism for this thought was spreading activation. Hélie and Sun (2010) similarly propose that incubation relies on a low attention, unconscious or implicit associative process that spreads activation to novel concepts through implicit networks.

Spreading activation and insight. Insight is the sudden realisation of the solution to a problem after a period of time stuck at an impasse, that is, when no path to the solution is apparent (Kaplan and Simon, 1990). A commonly used task to study insight is the Remote Associate Task (RAT; Mednick, 1962). In this, participants are presented with a word triad of distantly associated words and are asked to find a fourth word that connects them. For example, if presented with "rat," "blue," "cottage," the connecting word is "cheese." Bowers, Regehr, Balthazard, and Parker (1990) extended this task by contrasting pairs of triads in which one triad was coherent (had an associated connecting word) and one was incoherent (no connecting word). Participants were asked to choose the coherent triad, and they were able to do this even when they could not explicitly name the associate. Bowers et al. propose that the clues unconsciously and automatically activate relevant associates through spreading activation in a semantic network, raising the activation of the solution word that fits the triad. The process of activating the common associate in a Remote Associate Task triad occurs without intention. Topolinski and Strack (2008) found that simply reading the triad without attempting to solve it reduced the length of time to recognise the coherent solution word. Spreading activation also increases the speed of responding. Participants were faster to name solution words to triads, even when they had not solved the triad. This suggests that there was some semantic activation of the solution (Bowden and Beeman, 1998) and participants were able to discriminate coherent from incoherent triads under time pressure that prevented explicitly solving the problem (Bolte and Goschke, 2005). This suggests that the activation of the semantic associate is fast and automatic. Together, the RAT studies show that insight can occur through spreading activation in a semantic network without conscious awareness, without intention, and rapidly.

The slower unconscious process of incubation also benefits insight. Sio and Ormerod's (2009) meta-analysis found that incubation benefited linguistic and visual insight tasks. Zhong, Dijksterhuis, and Galinsky (2008) compared performance on the RAT with either a five minute incubation period in which participants completed the cognitively demanding 2-back task (unconscious condition) or thought consciously about the problems for five minutes. After this, participants in the unconscious condition responded more quickly to the solution words in a lexical decision task than those in the conscious condition, suggesting that the solution words were more active following unconscious incubation than was conscious thought. This supports the predictions of Dijksterhuis and Nordgren's claim that unconscious thought is more effective at associative processes than conscious thought (Dijksterhuis and Nordgren, 2006). This final claim may not be proven, but the meta-analytic evidence more strongly indicates that incubation supports insight and that spreading activation is one mechanism that has been used to explain this unconscious thought process.

Spreading activation and judgments. Spreading activation can also influence judgments. Numeric anchoring effects arise because the anchor activates knowledge that it is semantically related to, so this related knowledge becomes more accessible. When the target is evaluated, this more accessible knowledge is salient and more easily retrieved in the context of the target, biasing judgments of the target (Mussweiler and Strack, 2001). Hence the bias that people are overly influenced in their judgments of the target by the first piece of information they receive, the anchor. This has been demonstrated in diverse domains such as estimating the weight of Julius Caesar (Wegener, Petty, Detweiler–Bedell, and Jarvis, 2001) or the likelihood of nuclear war (Plous, 1989). Mussweiler and Englich (2005) also showed that this can occur unconsciously. Participants were asked to estimate the annual mean temperature in Germany, and were presented with an anchor of either 20 degrees or 5 degrees before responding. The anchor was presented for 15ms as a masked prime. Those who were presented with the high anchor estimated a mean temperature that was higher than those who were presented a low anchor, despite not being aware of the content of the subliminal prime. Mussweiler and Englich also found that participants were faster to recognise words that were semantically associated with the subliminal anchor, supporting the idea that spreading activation from the anchor caused the anchoring effect.

Overall, there is evidence that spreading activation is a mechanism that is involved in unconscious thought in several domains. Spreading activation can be fast and operate without intentional control to raise the salience of related concepts as potential solutions. It can be influenced subliminally, raising the salience without conscious awareness. It can also operate slowly, with incubation periods of five minutes and more, resulting in sudden insights or greater salience of solutions compared to conscious thought.

Deploying Mechanisms of Intuitive Thought

The three mechanisms of unconscious thought are triggered by cues and these are processed by the mechanisms which generate outputs. The outputs are either a specific response that is associated with the cues or an affective response.

Automaticity. Cue recognition is the trigger for an automatic process. Cues could be a complex or simple pattern perceived in the environment or internal states, such as the critical information assessed by a fire officer (Klein et al., 1986) or chess expert (Chase and Simon, 1973). These trigger the unconscious thought and have features typical of automaticity such as a lack of intention or control. However, there is no specific reward as a motivating factor for instigating the process. Simply, as a result of repetition, the cognitive process has become highly efficient and will autonomously operate in the situations in which it has been implemented before. The output is a specific response that has been acquired through practice such as the diagnosis of a medical condition in response to medical signs and symptoms or the application of a rule to set of cues such as a logical inference during conversation.

Reward-based associations. Reward-based processes are also triggered by cues in the environment or internally, but in addition they are motivated by the expectation of a desirable reward. For example, the habitual eating of popcorn is triggered by specific environmental cues such as being in a cinema, but it is also driven by the association of these cues with previous enjoyment of popcorn leading to the expectation that this popcorn will be enjoyable (Neal et al., 2011). This contrasts with an automatic process that occurs in response to cues simply because it is efficient to repeat a highly learnt process as it is likely to be useful again in the same context. A reward could take many different forms such as pleasant food, social affiliation, self esteem, financial gain etc. The output is an affective response that is associated with the outcome and used in future decisions, for example when applying the affect heuristic to judge an object according to the balance of positive affect associated with it (Slovic et al., 2004).

Spreading activation. Spreading activation is also triggered by cues that raise the activation of semantically related neighbours which in turn raise the activation of more distant neighbours. The output is not a specific response though. Rather than refining a solution, the unguided process would expand ever wider. To prevent this, the mechanism is guided by the goal of finding a coherent set of activated concepts (Öllinger and von Müller, 2017). That is, concepts mutually spread activation to other concepts that form a consistent network, satisfying the constraints between concepts in the network. The output is a set of coherent concepts and an affective response that indicates coherence (Topolinski and Strack, 2009). Coherence is also indicated by positive affect (Gamblin, Banks, and Dean, 2020).

The goal of forming coherent representations has a long history within psychology. Gestalt psychologists proposed laws that explain how we perceive complex stimuli as coherent groups rather than as individual objects. Heider (1958) proposed a theory of attitude change in which we seek balance between our values and beliefs and will change them in order to achieve a coherent state. Festinger (1962) similarly proposed that we are motivated to avoid cognitive dissonance, an uncomfortable state when our actions, values, and beliefs are not coherent. More recent theories have also explained thinking as a search for coherent representations. Thagard's (2002) computational theory of explanatory coherence uses a constraint satisfaction process to simulate thinking in which choices are selected and inferences made based on how well they fit within the constraints of the network. Bowers et al. (1990) proposed that a perception of coherence could be built up unconsciously that guides thought towards a solution by spreading activation to relevant semantic networks. When the activation of coherent information is sufficient, it crosses a threshold into conscious awareness. This can be fast, creating the impression of a sudden insight. Topolinski and Strack (2009) also argue that partial activation of coherent representations guide thought without conscious awareness. In this theory, the partial activation of coherent representations means that they are processed more fluently than incoherent representation. This leads to a conscious awareness of ease of processing that is felt as positive affect. That is, people feel a problem is coherent before they are consciously aware of the cause of the coherence. This feeling can be used to make a conscious judgment guided by the unconscious processing of the coherence of the problem.

Overall, the different mechanisms of unconscious thought are triggered by simple or complex patterns of cues perceived in the environment or internal states. Automatic processes efficiently generate outputs in response to specific cues. Reward-based processes also occur in response to specific cues but here the mechanism is different: the cues create an expectation of a reward which motivates the associated response. The output is an affective response that is associated with the outcome. Finally, spreading activation is guided by coherence, with coherent representations becoming more strongly activated leading to insights and judgments that fit within a coherent network. The output is a feeling of coherence indicated with an affective response.

Limits of Unconscious Thought — The Principle of Integration

It has been stated that the purpose of conscious awareness is to enable us to engage in complex internal processing such as constructing meaningful sequences of thought, reasoning logically, and simulating events such as the perspective of a partner in social interaction, whereas simpler animals and unconscious humans are limited to only learning associations between concepts and patterns (Baumeister and Masicampo, 2010). Completing these tasks requires the novel integration of high level semantic concepts. Whilst simpler integration may occur without conscious awareness, it may be necessary for more complex integration (Mudrik, Faivre, and Koch, 2014). Theories of consciousness such as global neuronal workspace describe the function of conscious awareness as allowing the long range connections required for this integration to occur (Dehaene and Changeux, 2011). Greenwald (1992) phrased this as a "two-word challenge," suggesting that it had not been possible at that time to demonstrate that unconscious awareness was able to extract the meaning of a two-word sequence. If so, this places a strict limit on the capacity for unconscious thought. Complex thinking across the domains of reasoning, judgment, decision making, insight problem solving, and creativity would seem to require the integration of novel, high level concepts (e.g., comparing two novel options) and therefore cannot occur without conscious thought. However, recent research has challenged this position with studies showing complex thinking without conscious awareness. This section will review some of these studies and highlight methodological problems that raise doubts about the conclusions drawn about the capacity of unconscious thinking. A limit of unconscious thought appears to be the integration of information. This leads to a "principle of integration" that specifies the limit of the capacity of unconscious thought: unconscious thought is not sufficient in tasks or problems that require concepts to be integrated in novel or unfamiliar ways.

Several theories directly challenge this principle. For example, unconscious thought theory makes strong claims for the capacity of unconscious thinking (Dijksterhuis and Nordgren, 2006). There are several principles within the theory. The bottom–up–versus–top–down principle suggests that information from several criteria are slowly integrated and organised by the unconscious into a summary evaluation. The weighting principle suggests that these criteria are then weighted in importance relative to each other. These principles require the integration of novel information to form a judgment. For example, Dijksterhuis (2004) presented participants with information about different apartments and found that decisions made after a three minute period of distraction (allowing unconscious thought) were better than decisions made after a three minute period of conscious deliberation. This suggests that participants were more effective decision makers when unconsciously weighting and integrating the attributes of the apartments than when consciously doing so. This effect has been demonstrated in complex decisions in other domains such as buying cars and other consumer products such as furniture (Dijksterhuis, Bos, Nordgren, and Baaren, 2004) as well as medical decisions (de Vries, Witteman, Holland, and Dijksterhuis, 2010). However, further research has failed to replicate these effects (Calvillo and Penaloza, 2009) and meta-analyses have failed to support them (Nieuwenstein et al., 2015). A likely alternative explanation for the effect, aside from the problems of weak effects presented in some studies (e.g., Dijksterhuis, 2004), is that some conscious analysis of the decision takes place during the actual presentation of the problem — prior to any unconscious distraction period — and this is used as a basis for the decision (Lassiter, Lindberg, González–Vallejo, Bellezza, and Phillips 2009; Newell, Wong, Cheung, Rakow, 2009). The evidence for the capacity of the unconscious to make complex decisions within this paradigm is equivocal.

Claims have also been made about the capacity of unconscious thought to solve complex math problems. Sklar, Levy, Goldstein, Mandel, Maril, and Hassin (2012) used continuous flash suppression whilst presenting an addition or subtraction problem subliminally, e.g., 9 – 4 – 3 =. Immediately after this prime, a number was presented supraliminally and participants were asked to pronounce it. Participants were faster to respond when the number was the solution to the subliminal math problem (i.e., 2 in the above example) than when it was not. This effect was found for subtractions but not additions. Whilst the arithmetic task is simple, it does require the integration of different numbers into a simple function for a solution to an equation that is sufficiently uncommon that it is unlikely to be directly retrieved from memory, suggesting that abstract symbols can in fact be integrated without conscious awareness. However, there are reasons to doubt this effect. Theoretically, it is surprising that the effect would be found for subtraction and not addition as the cognitive processes in these tasks seem similar. It is more surprising given that Ric and Muller (2012) examined unconscious math using masked priming and found an effect for addition and not subtraction. One possibility is that Sklar et al.'s effect is an artefact of their method. They did not analyse all participants, but selected a subgroup of participants after testing who had not shown conscious awareness of the prime. This method of post hoc selection of an extreme group on one variable (conscious awareness) has been shown to artefactually create an effect on the second variable of interest (response time to respond to the target number) because of regression to the mean (Shanks, 2017). Further Bayesian reanalysis of the data also suggests that the effect is not as strong as claimed (Moors and Hesselmann, 2017) and it has not been successfully replicated (Karpinski, Yale, and Briggs, 2016).

Van Opstal, de Lange, and Dehaene (2011) presented four digits subliminally using a masked prime followed by a further four digits which were presented supraliminally. Participants had to assess whether the sum or mean of the target set of digits were above or below a given figure. Participants were faster to make this calculation when the mean or the sum of the prime digits was congruent with the answer than when it was incongruent. This suggests that participants were able to extract an approximate average of the prime unconsciously, another apparent example of applying a function to integrate abstract symbols. But again, it is not certain if participants were entirely unaware of the primes. The effect of congruency was shown to occur without conscious awareness using a method of regressing the performance in the task onto the level of prime conscious awareness (Greenwald, Draine, and Abrams, 1996). Some concerns have been raised about the validity of this method (Miller, 2000) and it is possible that participants did have some conscious awareness of the primes.

A simple example of integrating concepts is processing negations. To process a negation such as "this drink is not hot" requires understanding the word "hot" and retrieval of the semantic opposite "cold." The words "not" and "hot" are processed in combination. Reasoning with negations is more difficult (Wason, 1959) and the negations are not always processed accurately. For example the innuendo effect arises when a negative attribute is negated, e.g., "the politician was not bribed." This leads to more negative attitudes to the target (the politician) as the negative attribute is remembered but the negation is not (Wegner, Wenzlaff, Kerker, and Beattie, 1981). Similarly the illusion of truth arises when claims are repeatedly explained to be false — the claim is recalled, but not the fact it was false. Skurnik, Yoon, Park, and Schwarz (2005) found that three days after explaining to elderly participants that a set of health claims were false, the claims were more likely to be remembered as true because they were familiar but their falsehood was forgotten. Deutsch, Gawronski, and Strack (2006) trained participants to evaluate words as positive or negative, e.g., "party" has a positive association. As expected, participants' response time to this task reduced with practice as the task became increasingly automated. Participants also had to evaluate negated words such as "no party." With practice the time taken to process the negation remained the same, suggesting that negating concepts did not become automated. Thus the associative process of evaluating the word can become automatic but integrating it with a negation cannot.

An apparent exception to the principle of integration is the application of the modus ponens rule which requires the integration of major and minor premise to draw a conclusion. Reverberi et al. (2012) suggest that this inference can occur without conscious awareness. However, the explanation for this exception is that the inference schema for modus ponens is highly practiced to the point of automaticity. Therefore the limitation that unconscious thought is not sufficient to integrate information only applies to the integration of concepts in novel or less highly practiced ways.

Overall, these studies all make claims that unconscious thought is sophisticated and can effectively integrate symbolic, abstract concepts to make decisions, form judgments etc. However, closer methodological analysis of these studies typically casts doubt on the reliability of the findings e.g., they are difficult to replicate or it is difficult to demonstrate participants were not thinking consciously at some point during the experiment. Whilst continued efforts to explore the limits of unconscious thought are valuable, at present the evidence suggests a limiting factor that unconscious thought is not sufficient to solve problems that require concepts to be integrated in novel or unfamiliar ways.

Discussion

Each of the three mechanisms are distinct and operate in different ways, but they share the basic feature that they exploit extensive experience in the environment that has been learnt over many repeated encounters with similar problems. As a result they are, in the long run, highly adaptive. For example, the computational models of reinforcement learning that are similar to reward-based learning in humans are described as normative solutions to learning associations (Niv, 2009). Through experience the mechanisms generate effective solutions to familiar environments, and do so efficiently by requiring limited attention and mental effort.

The proposal of three distinct mechanisms raises the question of how they might interact. For example, if a choice is made repeatedly because it leads to the optimal reward, the repetition would also result in the process becoming automated. After a period of time, especially in a naturalistic environment such as the study of habitual behaviours, the choice becomes an automatic, reward-based behaviour. Unconscious processes also interact with conscious thought processes. For example the Aha! moment of insight is when the unconscious process becomes conscious. Incubation suggests a serial process in which unconscious thought precedes conscious thought. However, it could be that conscious and unconscious thought operate in parallel, either in competition, generating different responses, or in cooperation, providing mutual support for a response. The theoretical approach to conscious awareness and attention proposed in the model presented here, that attention is a continuous property that triggers qualitative shifts in cognition at a threshold, also suggests how conscious and unconscious thinking may interact. The unconscious process may well become conscious at some stage, for example when a solution is found, but this will occur when the process requires or attracts a level of attention above the threshold.

The relationship between conscious and unconscious thought mirrors the debate about the relationship between system 1 and system 2 (or type 1 and type 2) thinking within dual process theories of thinking (De Neys, 2017; Evans, 2008; Kahneman and Frederick, 2002). Only Lieberman, Gaunt, Gilbert, and Trope (2002) directly equate unconscious thought with the automatic X system and conscious thought with the deliberate C system. Most dual process theorists claim that both type 1 and type 2 could have conscious and unconscious components

(Evans and Stanovich, 2013). Instead, type 2 processes are defined as those requiring working memory, implying that type 1 processes are those that do not require working memory (Evans and Stanovich, 2013). This leaves open the question of what those type 1 mechanisms might be. It has been suggested that unconscious processes do not involve working memory (Baars and Franklin, 2003) and so the mechanisms proposed in the model could represent the three fundamental mechanisms behind system 1 or type 1 processing.

The model also suggests a number of fruitful directions for research. First, specifying the goals and mechanisms of unconscious thought defines what can and cannot occur unconsciously. The principle of integration proposes that concepts cannot be combined in novel ways, yet it is possible for simple automatic schemas to be applied unconsciously (Reverberi et al., 2012). How complex are the schemas that can be learnt, given sufficient practice? Research on expert decision making suggests that schemas can become very complex, and so the limited capacity in studies of non-expert reasoning may simply be a lack of practice with complex problems. Second, as all of the mechanisms are based in experience of familiar environments, the role of learning is central to all of them. Yet most studies rely on the presentation of novel tasks and so do not examine thought processes that have been acquired over time. Research on decisions by experience, in which the probability of outcomes are learnt through repeated trials, has highlighted how they might differ from decisions by description (Hertwig, Barron, Weber, and Erev, 2004), and the role of learning may well influence other types of thinking too.

Conclusion

Unconscious thought occurs when the level of attention to a task falls below a threshold of conscious awareness but above a minimum threshold where no processing occurs at all. The purpose of this article was to review the mechanisms of unconscious thought that have been proposed across a range of domains of thinking — reasoning, decision making, judgment, insight problem solving, and creativity. The aim was to find common mechanisms that have been used in each of these areas and combine the evidence for them to form an overall account of unconscious thought. Across all of these domains, three types of mechanism recur: automaticity, reward-based association, and spreading activation. The mechanisms are triggered by cues in the environment or internal states and the output of the mechanisms are either specific outputs or affective responses. A second aim was to find the limits of the unconscious thought. Some argue that unconscious thought is very limited and others argue that it can solve problems as well as, and sometimes better, than conscious thought. The review identified a limiting factor referred to as the principle of integration. Unconscious thought is not sufficient in tasks or problems that require concepts to be integrated in novel or unfamiliar ways. Where theories made stronger claims, the evidence supporting those claims was found to be equivocal. Unconscious thought is frequently used and the three underlying mechanisms adapt well to frequently encountered situations. These mechanisms provide the limits of unconscious thought but also the capacity for highly effective thinking.

References

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22,* 261–295.

Anderson, J. R. (1992). Automaticity and the ACT theory. *American Journal of Psychology, 105,* 165–180.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., and Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111,* 1036–1060.

Anderson, J. R., Fincham, J. M., and Douglass, S. (1997). The role of examples and rules in the acquisition of a cognitive skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 932–945.

Baars, B. J. (1997). In the theatre of consciousness. Global workspace theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies, 4,* 292–309.

Baars, B. J., and Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences, 7,* 166–172.

Baddeley, A. D. (2007). *Working memory, thought, and action.* Oxford: Oxford University Press.

Banks, A. P., and Hope, C. (2014). Heuristic and analytic processes in reasoning: An event-related potential study of belief bias. *Psychophysiology, 51,* 290–297.

Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology, 105,* 181–199.

Baumeister, R. F., and Masicampo, E. J. (2010). Conscious thought is for facilitating social and cultural interactions: How mental simulations serve the animal–culture interface. *Psychological Review, 117,* 945–971.

Baumeister, R. F., Masicampo, E. J., and Vohs, K. D. (2011). Do conscious thoughts cause behavior? *Annual Review of Psychology, 62,* 331–361.

Bechara, A., Damasio, H., Tranel, D., and Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275,* 1293–1295.

Betsch, T., Plessner, H., Schwieren, C., and Gütig, R. (2001). I like it but I don't know why: A value-account approach to implicit attitude formation. *Personality and Social Psychology Bulletin, 27,* 242–253.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences, 18,* 227–247.

Bolte, A., and Goschke, T. (2005). On the speed of intuition: Intuitive judgments of semantic coherence under different response deadlines. *Memory and Cognition, 33,* 1248–1255.

Bowden, E. M., and Beeman, M. J. (1998). Getting the right idea: Semantic activation in the right hemisphere may help solve insight problems. *Psychological Science, 9,* 435–440.

Bowers, K. S., Regehr, G., Balthazard, C., and Parker, K. (1990). Intuition in the context of discovery. *Cognitive Psychology, 22,* 72–110.

Braine, M. D., Reiser, B. J., and Rumain, B. (1984). Some empirical justification for a theory of natural propositional logic. In G.H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 18, pp. 313–371). Oxford: Academic Press.

Breitmeyer, B., and Öğmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision (No. 41).* Oxford: Oxford University Press.

Bryant, P. E., and Trabasso, T. (1971). Transitive inferences and memory in young children. *Nature, 232,* 456–458.

Calvillo, D. P., and Penaloza, A. (2009). Are complex decisions better left to the unconscious? Further failed replications of the deliberation-without-attention effect. *Judgment and Decision Making, 4,* 509–517.

Chase, W. G., and Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55–81.

Chein, J. M., and Schneider, W. (2012). The brain's learning and control architecture. *Current Directions in Psychological Science, 21,* 78–84.
Cleeremans, A. (2005). Computational correlates of consciousness. *Progress in Brain Research, 150,* 81–98.
Collins, A. M., and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82,* 407–428.
Crandall, B., and Getchell–Reiter, K. (1993). Critical decision method: A technique for eliciting concrete assessment indicators from the intuition of NICU nurses. *Advances in Nursing Science, 16,* 42–51.
Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London B, 351,* 1413–1420.
Dane, E., Rockmann, K. W., and Pratt, M. G. (2012). When should I trust my gut? Linking domain expertise to intuitive decision-making effectiveness. *Organizational Behavior and Human Decision Processes, 119,* 187–194.
Davis, H. (1992). Transitive inference in rats *(Rattus norvegicus). Journal of Comparative Psychology, 106,* 342–349.
Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience, 8,* 1704–1711.
Dayan, P., and Niv, Y. (2008). Reinforcement learning: The good, the bad and the ugly. *Current Opinion in Neurobiology, 18,* 185–196.
De Groot, A. D. (1946). *Het denken van den schaker.* Amsterdam: Noord Hollandsche.
De Groot, A. D. (1965). *Thought and choice in chess.* The Hague: Mouton Publishers.
De Neys, W. (2012). Bias and conflict: A case for logical intuitions. *Perspectives on Psychological Science, 7,* 28–38.
De Neys, W. (Ed.). (2017). *Dual process theory 2.0.* Oxford: Routledge.
De Neys, W., and Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition, 106,* 1248–1299.
De Neys, W., Moyens, E., and Vansteenwegen, D. (2010). Feeling we're biased: Autonomic arousal and reasoning conflict. *Cognitive, Affective, and Behavioral Neuroscience, 10,* 208–216.
De Vries, M., Witteman, C. L., Holland, R. W., and Dijksterhuis, A. (2010). The unconscious thought effect in clinical decision making: An example in diagnosis. *Medical Decision Making, 30,* 578–581.
De Wit, S., Corlett, P. R., Aitken, M. R., Dickinson, A., and Fletcher, P. C. (2009). Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. *Journal of Neuroscience, 29,* 11330–11338.
Dehaene, S., and Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70,* 200–227.
Dehaene, S., Kerszberg, M., and Changeux, J.P. (2001), A neuronal model of a global workspace in effortful cognitive tasks. *Annals of the New York Academy of Sciences, 929,* 152–165.
Dehaene, S., and Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79,* 1–37.
Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene–Lambertz, G., van de Moortele, P-F., and Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature, 395,* 597–600.
Deutsch, R., Gawronski, B., and Strack, F. (2006). At the boundaries of automaticity: Negation as reflective operation. *Journal of Personality and Social Psychology, 91,* 385–405.
Dijksterhuis, A. (2004). Think different: The merits of unconscious thought in preference development and decision making. *Journal of Personality and Social Psychology, 87,* 586–598.
Dijksterhuis, A., and Aarts, H. (2010). Goals, attention, and (un)consciousness. *Annual Review of Psychology, 61,* 467–490.
Dijksterhuis, A., Bos, M. W., Nordgren, L. F., and Van Baaren, R. B. (2006). On making the right choice: The deliberation-without-attention effect. *Science, 311,* 1005–1007.
Dijksterhuis, A., and Nordgren, L. F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science, 1,* 95–109.
Dunn, B. D., Dalgleish, T., and Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience and Biobehavioral Reviews, 30,* 239–271.
Ericsson, K. A., Krampe, R. T., and Tesch–Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406.

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review Psychology, 59,* 255–278.
Evans, J. S. B., Barston, J. L., and Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition, 11,* 295–306.
Evans, J. S. B., and Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8,* 223–241.
Festinger, L. (1962). *A theory of cognitive dissonance.* Stanford: Stanford University Press.
Finucane, M. L., Alhakami, A., Slovic, P., and Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making, 13,* 1–17.
Frank, M. J., Rudy, J. W., Levy, W. B., and O'Reilly, R. C. (2005). When logic fails: Implicit transitive inference in humans. *Memory and Cognition, 33,* 742–750.
Gamblin, D. M., Banks, A. P., and Dean, P. J. (2020). Affective responses to coherence in high and low risk scenarios. *Cognition and Emotion, 34,* 462–480.
Gawronski, B., Deutsch, R., Mbirkou, S., Seibt, B., and Strack, F. (2008). When "just say no" is not enough: Affirmation versus negation training and the reduction of automatic stereotype activation. *Journal of Experimental Social Psychology, 44,* 370–377.
Gilhooly, K. J., Fioratou, E., Anthony, S. H., and Wynn, V. (2007). Divergent thinking: Strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology, 98,* 611–625.
Gilhooly, K. J., Georgiou, G. J., Garrison, J., Reston, J. D., and Sirota, M. (2012). Don't wait to incubate: Immediate versus delayed incubation in divergent thinking. *Memory and Cognition, 40,* 966–975.
Gobet, F., and Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31,* 1–40.
Gore, J., and Sadler–Smith, E. (2011). Unpacking intuition: A process and outcome framework. *Review of General Psychology, 15,* 304–316.
Greenwald, A. G. (1992). New Look 3: Unconscious cognition reclaimed. *American Psychologist, 47,* 766–799.
Greenwald, A. G., Draine, S. C., and Abrams, R. L. (1996). Three cognitive markers of unconscious semantic activation. *Science, 273,* 1699–1702.
Guilford, J. P. (1967). *The nature of human intelligence.* New York: McGraw–Hill.
Handley, S. J., Newstead, S. E., and Trippas, D. (2011). Logic, beliefs, and instruction: A test of the default interventionist account of belief bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 28–43.
Handley, S. J., and Trippas, D. (2015). Dual processes and the interplay between knowledge and structure: A new parallel processing model. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 62, pp. 33–58). Oxford: Academic Press.
Harris, J. L., Bargh, J. A., and Brownell, K. D. (2009). Priming effects of television food advertising on eating behavior. *Health Psychology, 28,* 404–413.
Hassin, R. R. (2013). Yes it can: On the functional abilities of the human unconscious. *Perspectives on Psychological Science, 8,* 195–207.
Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.
Hélie, S., and Sun, R. (2010). Incubation, insight, and creative problem solving: A unified theory and a connectionist model. *Psychological Review, 117,* 994–1024.
Hertwig, R., Barron, G., Weber, E. U., and Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, 15,* 534–539.
Hesselmann, G., and Moors, P. (2015). Definitely maybe: Can unconscious processes perform the same functions as conscious processes? *Frontiers in Psychology, 6,* 584.
Johnson–Laird, P. N., Mancini, F., and Gangemi, A. (2006). A hyper-emotion theory of psychological illnesses. *Psychological Review, 113,* 822–841.
Kahneman, D., and Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, and D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). Cambridge: Cambridge University Press.
Kahneman, D., and Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.
Kaplan, C. A., and Simon, H. A. (1990). In search of insight. *Cognitive Psychology, 22,* 374–419.
Karpinski, A., Yale, M., and Briggs, J. C. (2016). *Retracted:* Unconscious arithmetic processing: A direct replication. *European Journal of Social Psychology, 46,* 384–391.

Kaushal, N., and Rhodes, R. E. (2015). Exercise habit formation in new gym members: A longitudinal study. *Journal of Behavioral Medicine, 38,* 652–663.
Kiefer, M. (2012). Executive control over unconscious cognition: Attentional sensitization of unconscious information processing. *Frontiers in Human Neuroscience, 6,* 61.
Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G.A. Klein, J. Orasanu, R. Calderwood, and C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 138–147). Norwood, New Jersey: Ablex.
Klein, G. A. (1998). *Sources of power: How people make decisions.* Cambridge, Massachusetts: MIT Press.
Klein, G. A., Calderwood, R., and Clinton–Cirocco, A. (1986, September). Rapid decision making on the fire ground. In *Proceedings of the Human Factors Society Annual Meeting* (Vol. 30, No. 6, pp. 576–580). Los Angeles: SAGE Publications.
Lassiter, G. D., Lindberg, M. J., González–Vallejo, C., Bellezza, F. S., and Phillips, N. D. (2009). The deliberation-without-attention effect: Evidence for an artifactual interpretation. *Psychological Science, 20,* 671–675.
Lea, R. B. (1995). On-line evidence for elaborative logical inferences in text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1469–1482.
Lea, R. B., O'Brien, D. P., Fisch, S. M., Noveck, I. A., and Braine, M. D. (1990). Predicting propositional logic inferences in text comprehension. *Journal of Memory and Language, 29,* 361–387.
Leo, P. D., and Greene, A. J. (2008). Is awareness necessary for true inference? *Memory and Cognition, 36,* 1079–1086.
Lerner, J. S., Li, Y., Valdesolo, P., and Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology, 66,* 799–823.
Lieberman, M. D., Gaunt, R., Gilbert, D. T., and Trope, Y. (2002). Reflexion and reflection: A social cognitive neuroscience approach to attributional inference. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 199–249). San Diego: Academic Press.
Lisman, J., and Sternberg, E. J. (2013). Habit and nonhabit systems for unconscious and conscious behavior: Implications for multitasking. *Journal of Cognitive Neuroscience, 25,* 273–283.
Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.
Macdonald, J. S., and Lavie, N. (2008). Load induced blindness. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1078–1091.
Mednick, S. (1962). The associative basis of the creative process. *Psychological Review, 69,* 220–232.
Miller, J. (2000). Measurement error in subliminal perception experiments: Simulation analyses of two regression methods. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1461–1477.
Morsanyi, K., and Handley, S. J. (2012). Logic feels so good — I like it! Evidence for intuitive detection of logicality in syllogistic reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 596–616.
Moors, A. (2016). Automaticity: Componential, causal, and mechanistic explanations. *Annual Review of Psychology, 67,* 263–287.
Moors, A., and De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132,* 297–326.
Moors, P., and Hesselmann, G. (2017). A critical reexamination of doing arithmetic nonconsciously. *Psychonomic Bulletin and Review, 25,* 472–481.
Mudrik, L., Faivre, N., and Koch, C. (2014). Information integration without awareness. *Trends in Cognitive Sciences, 18,* 488–496.
Mussweiler, T., and Englich, B. (2005). Subliminal anchoring: Judgmental consequences and underlying mechanisms. *Organizational Behavior and Human Decision Processes, 98,* 133–143.
Mussweiler, T., and Strack, F. (2001). The semantics of anchoring. *Organizational Behavior and Human Decision Processes, 86,* 234–255.
Neal, D. T., Wood, W., Wu, M., and Kurlander, D. (2011). The pull of the past: When do habits persist despite conflict with motives? *Personality and Social Psychology Bulletin, 37,* 1428–1437.
Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106,* 226–254.
Newell, B. R., and Shanks, D. R. (2014). Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences, 37,* 1–19.

Newell, B. R., Wong, K. Y., Cheung, J. C., and Rakow, T. (2009). Think, blink or sleep on it? The impact of modes of thought on complex decision making. *Quarterly Journal of Experimental Psychology, 62,* 707–732.

Nieuwenstein, M. R., Wierenga, T., Morey, R. D., Wicherts, J. M., Blom, T. N., Wagenmakers, E. J., and van Rijn, H. (2015). On making the right choice: A meta-analysis and large-scale replication attempt of the unconscious thought advantage. *Judgment and Decision Making, 10,* 1–17.

Nisbett, R. E., and Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology, 53,* 139–154.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., and Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron, 38,* 329–337.

Öllinger, M., and von Müller, A. (2017). Search and coherence-building in intuition and insight problem solving. *Frontiers in Psychology, 8,* 827.

Otto, A. R., Gershman, S. J., Markman, A. B., and Daw, N. D. (2013). The curse of planning: Dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychological Science, 24,* 751–761.

Panksepp, J. (2005). Affective consciousness: Core emotional feelings in animals and humans. *Consciousness and Cognition, 14,* 30–80.

Pennycook, G., Fugelsang, J. A., and Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology, 80,* 34–72.

Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., and Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron, 59,* 561–567.

Peters, E. (2006). The functions of affect in the construction of preferences. In S. Lichtenstein and P. Slovic (Eds.), *The construction of preference.* (pp. 454–463). Cambridge: Cambridge University Press.

Plous, S. (1989). Thinking the unthinkable: The effects of anchoring on likelihood estimates of nuclear war. *Journal of Applied Social Psychology, 19,* 67–91.

Reingold, E. M., Charness, N., Schultetus, R. S., and Stampe, D. M. (2001). Perceptual automaticity in expert chess players: Parallel encoding of chess relations. *Psychonomic Bulletin and Review, 8,* 504–510.

Reverberi, C., Pischedda, D., Burigo, M., and Cherubini, P. (2012). Deduction without awareness. *Acta Psychologica, 139,* 244–253.

Rader, A. W., and Sloutsky, V. M. (2002). Processing of logically valid and logically invalid conditional inferences in discourse comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 59–68.

Ric, F., and Muller, D. (2012). Unconscious addition: When we unconsciously initiate and follow arithmetic rules. *Journal of Experimental Psychology: General, 141,* 222–226.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking.* Cambridge, Massachusetts: MIT Press.

Schmidt, H., Norman, G., and Boshuizen, H. (1990). A cognitive perspective on medical expertise: Theory and implications. *Academic Medicine, 65,* 611–621.

Shanks, D. R. (2017). Regressive research: The pitfalls of post hoc data selection in the study of unconscious mental processes. *Psychonomic Bulletin and Review, 24,* 752–775.

Shea, N., and Frith, C. D. (2016). Dual-process theories and consciousness: The case for 'Type Zero' cognition. *Neuroscience of Consciousness,* 1–10.

Shiffrin, R. M. (1988). Attention. In R. C. Atkinson, R. J. Hernstein, G. Lindzey, and R. D. Luce (Eds.), *Stevens' handbook of experimental psychology* (Vol. 2, pp. 739–811). New York: Wiley.

Simon, H. A. (1987). Making management decisions: The role of intuition and emotion. *Academy of Management Perspectives, 1,* 57–64.

Simon, H. A. (1992). What is an "explanation" of behavior? *Psychological Science, 3,* 150–161.

Sio, U. N., and Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin, 135,* 94–120.

Skinner, B. F. (1963). Operant behavior. *American Psychologist, 18,* 503–515

Sklar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., and Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. *Proceedings of the National Academy of Sciences, 109,* 19614–19619.

Skurnik, I., Yoon, C., Park, D. C., and Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research, 31,* 713–724.

Slovic, P., Finucane, M. L., Peters, E., and MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis, 24,* 311–322.
Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning, 3,* 9–44.
Taatgen, N. A., and Wallach, D. (2002). Whether skill acquisition is rule or instance based is determined by the structure of the task. *Cognitive Science Quarterly, 2,* 163–204.
Thagard, P. (2002). *Coherence in thought and action.* Cambridge, Massachusetts: MIT press.
Thorndike, E. L. (1911). *Animal intelligence: Experimental studies.* New York: Macmillan.
Topolinski, S., and Strack, F. (2008). Where there's a will — there's no intuition. The unintentional basis of semantic coherence judgments. *Journal of Memory and Language, 58,* 1032–1048.
Topolinski, S., and Strack, F. (2009). Scanning the "fringe" of consciousness: What is felt and what is not felt in intuitions about semantic coherence. *Consciousness and Cognition, 18,* 608–618.
Tricomi, E., Balleine, B. W., and O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *European Journal of Neuroscience, 29,* 2225–2232.
Trippas, D., Handley, S. J., Verde, M. F., and Morsanyi, K. (2016). Logic brightens my day: Evidence for implicit sensitivity to logical validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42,* 1448–1457.
Tsuchiya, N., and Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8,* 1096–1101.
Van Opstal, F., de Lange, F. P., and Dehaene, S. (2011). Rapid parallel semantic processing of numbers without awareness. *Cognition, 120,* 136–147.
Wallas, G. (1926). *The art of thought.* New York: Harcourt Brace.
Wason, P. C. (1959). The processing of positive and negative information. *Quarterly Journal of Experimental Psychology, 11,* 92–107.
Weaver, J. E., Steirn, J. N., and Zentall, T. R. (1997). Transitive inference in pigeons: Control for differential value transfer. *Psychonomic Bulletin and Review, 4,* 113–117.
Wegener, D. T., Petty, R. E., Detweiler–Bedell, B. T., and Jarvis, W. B. G. (2001). Implications of attitude change theories for numerical anchoring: Anchor plausibility and the limits of anchor effectiveness. *Journal of Experimental Social Psychology, 37,* 62–69.
Wegner, D. M., Wenzlaff, R., Kerker, R. M., and Beattie, A. E. (1981). Incrimination through innuendo: Can media questions become public answers? *Journal of Personality and Social Psychology, 40,* 822–832.
Westen, D. (1998). The scientific legacy of Sigmund Freud: Toward a psychodynamically informed psychological science. *Psychological Bulletin, 124,* 333–371.
Wood, W., Quinn, J. M., and Kashy, D. A. (2002). Habits in everyday life: Thought, emotion, and action. *Journal of Personality and Social Psychology, 83,* 1281–1297.
Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35,* 151–175.
Zhong, C. B., Dijksterhuis, A., and Galinsky, A. D. (2008). The merits of unconscious thought in creativity. *Psychological Science, 19,* 912–918.