

## Opinion

## Free will without consciousness?

Liad Mudrik <sup>1,2,18,\*,@</sup> Inbal Gur Arie,<sup>1</sup> Yoni Amir,<sup>2</sup> Yarden Shir,<sup>1</sup> Pamela Hieronymi,<sup>3</sup> Uri Maoz,<sup>4,5,6,7</sup> Timothy O'Connor,<sup>8</sup> Aaron Schurger,<sup>4,5,9,10</sup> Manuel Vargas,<sup>11</sup> Tillmann Vierkant,<sup>12</sup> Walter Sinnott-Armstrong,<sup>13,14,15,16</sup> and Adina Roskies<sup>17</sup>

Findings demonstrating decision-related neural activity preceding volitional actions have dominated the discussion about how science can inform the free will debate. These discussions have largely ignored studies suggesting that decisions might be influenced or biased by various unconscious processes. If these effects are indeed real, do they render subjects' decisions less free or even unfree? Here, we argue that, while unconscious influences on decision-making do not threaten the existence of free will in general, they provide important information about limitations on freedom in specific circumstances. We demonstrate that aspects of this long-lasting controversy are empirically testable and provide insight into their bearing on degrees of freedom, laying the groundwork for future scientific-philosophical approaches.

### A threat to free will? Unconscious influences on decision-making

Are our actions free? The origins of this question (reviewed in [1–3]) date back to ancient and medieval times, making it one of the longest and most fervently debated in history. More recently, the debate has been strongly influenced by neuroscientific findings suggesting that subjects' conscious decisions are preceded by decision-related neural activity [4–6]. The meaning of such findings has been contested theoretically and empirically [7–10] and the controversy is ongoing [11,12].

Yet, the cognitive sciences have introduced another line of research that has so far been relatively overlooked in the free will debate (but see [13–17]): a large body of literature claims that decisions might be influenced or biased, sometimes heavily, by various unconscious processes. Would such influences render our decisions unfree?

Here, we critically examine the potential of these findings to pose a threat to free will. Importantly, such threats cut across **compatibilist** (see [Glossary](#)) and **incompatibilist** accounts, as well as reductionist and non-reductionist accounts: unconscious effects on decision-making could be either deterministic or nondeterministic, and, similarly, reducible or not (see [Box 1](#) for an explanation of why these threats are orthogonal to questions about the role of consciousness in decision-making). Cognizant that some reports have been heavily criticized and results called into doubt ([Box 2](#)), we identify the type of findings that could serve as empirical evidence against free will. Thus, we lay the groundwork for a combined scientific-philosophical approach that may advance this age-old question.

### A typology of threats to free will in studies reporting unconscious influences

Imagine the following scenario: it is election day, and you enter the polling booth. You intended to vote for one candidate, but now you reconsider. After deliberating for a few minutes, you decide to vote for a different candidate. Like many decisions, this one probably felt free: it appeared to be you who voluntarily changed your mind, without any external compulsion. Now, assume you

### Highlights

A growing body of literature argues for unconscious effects on decision-making.

We review a body of such studies while acknowledging methodological limitations, and categorize the types of unconscious influence reported.

These effects intuitively challenge free will, despite being generally overlooked in the free will literature. To what extent can decisions be free if they are affected by unconscious factors?

Our analysis suggests that unconscious influences on behavior affect degrees of control or reasons-responsiveness. We argue that they do not threaten the existence of free will in general, but only the degree to which we can be free in specific circumstances.

<sup>1</sup>School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel

<sup>2</sup>Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel

<sup>3</sup>Department of Philosophy, UCLA, Los Angeles, CA 90095, USA

<sup>4</sup>Department of Psychology, Chapman University, Orange, CA 92866, USA

<sup>5</sup>Institute for Interdisciplinary Brain and Behavioral Sciences, Chapman University, Irvine, CA 92618, USA

<sup>6</sup>Anderson School of Management, University of California Los Angeles, Los Angeles, CA 90095, USA

<sup>7</sup>Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

<sup>8</sup>Department of Philosophy, Cognitive Science Program, Indiana University, Bloomington, IN 47401, USA

<sup>9</sup>INSERM U992, Cognitive Neuroimaging Unit, Neurospin Center, Gif-sur-Yvette 91191, France

<sup>10</sup>Commissariat à l'Energie Atomique, Direction des Sciences du Vivant, I2BM, NeuroSpin Center, Gif sur Yvette 91191, France

Box 1. Other challenges to free will

**Determinism**

Many accounts of free will hinge on the truth of **determinism**. Incompatibilist accounts suggest that an act of will is free only if it is not determined by antecedent causes [90–95]. By contrast, compatibilist accounts deny that free will and determinism are incompatible [96–98]. The findings reviewed here are orthogonal to the question of determinism as our discussion does not explore whether unconscious factors are deterministic or **indeterministic**. Indeed, some consider the truth or falsity of determinism to be a question that will never be settled by any neuroscientific findings, given its meta-physical nature [99].

**Reductionism**

Another purported threat to incompatibilist accounts of free will comes from reductionism, which is the claim that everything mental, including decisions, choices, and wills, can be fully reduced to their physical realizers [100]. Opponents of reductionism include those who see wills or decisions as **emergent** [101], as events or properties that cannot be fully reduced to their constituent physical mechanisms in the brain [100]. Some **emergentists** argue that irreducible mental properties can cause changes in the physical world [102], but others are **epiphenomenalists**, who deny that emergent mental properties cause any such changes [103,104]. If our wills have no effects, then, regardless of whether they are free, we never act as a result of free will. Since the studies reviewed here do not attempt to reduce decisions to neural mechanisms, they do not bear upon this particular threat to free will.

<sup>11</sup>Department of Philosophy, University of California, San Diego, La Jolla, CA 92093, USA

<sup>12</sup>School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, Edinburgh, EH8 9AD, UK

<sup>13</sup>Philosophy Department, Duke University, Durham, NC 27708, USA

<sup>14</sup>Kenan Institute for Ethics, Duke University, Durham, NC 27708, USA

<sup>15</sup>Psychology and Neuroscience Department, Duke University, Durham, NC 27708, USA

<sup>16</sup>Duke University School of Law, Durham, NC 27708, USA

<sup>17</sup>Department of Philosophy, Dartmouth College, Hanover, NH 03755, USA

<sup>18</sup>[www.mudriklab.tau.ac.il](http://www.mudriklab.tau.ac.il)

\*Correspondence:

[mudrikli@tauex.tau.ac.il](mailto:mudrikli@tauex.tau.ac.il) (L. Mudrik).

Twitter: @Liad\_Mudrik

discover that your decision can be at least partly explained by the difference in the width-to-height ratio of the candidates' faces or by the location of their names on the ballot, two factors that have been shown to influence other **deliberate decisions** [18,19]. You would never have imagined that seemingly irrelevant details, which you view as providing no reason to favor any of the candidates, would affect such an important decision. Does your decision now appear less free?

Surprisingly, despite the increasing number of studies reporting different types of unconscious effects on decision-making, the significance of such findings for free will has been mostly neglected (but see again [13–16]). Here, we classify these studies into four categories (neither

Box 2. Key criticisms of the evidence for unconscious effects

The studies presented here have been criticized on different grounds and, accordingly, are used here as examples of potential findings that could threaten free will, rather than as actual evidence against free will. Some effects were not replicated or shown to be less pronounced (e.g., priming with food odors did not influence food choice [105]; anchoring effects were markedly smaller in replications [106]; and a recent meta-analysis of nudging effects suggested they were not as effective as in the original studies [107]). A prominent example relates to ego depletion: two large-scale, multilaboratory replication attempts, across 23 [108] and 36 laboratories [109], together encompassing more than 5000 participants, failed to reproduce the key finding [110] that subjects show less self-control in the ego-depleted group. Given this and other work (e.g., [111]), there is currently no conclusive evidence for ego depletion [112].

Other studies have been methodologically criticized. For example, the finding that judges are more likely to approve paroles at the beginning of the workday, or after food break [50] was originally taken as evidence for the effect of hunger on decision-making. However, a later study showed that, in each session, prisoners represented by an attorney went first, providing the presence of an attorney as a compelling alternative explanation for the results [51].

A prominent critique of unconscious effects [113] highlights their methodological limitations. It claims that they failed to convincingly demonstrate that subjects relied on unconscious processes, or that they were unaware of the critical stimulus. The latter is especially relevant for subliminal priming, in which a stimulus is rendered invisible and its effect on behavior is measured. Here, the devil truly is in the details, as minor methodological modifications could yield very different results [89,114,115]. In the studies reviewed in this paper, awareness was not always measured on a trial-by-trial basis [68,69,116], or not in the same group of subjects [116,117]. Thus, it is hard to ascertain whether subjects were indeed unaware of the presented stimuli. Other criticism suggests that somewhat similar measurement problems also apply to studies in which participants are asked to report their awareness of factors affecting their decision [113].

A final line of criticism highlights that many of the findings are obtained only under very specific conditions. For example, while subjects can be biased toward choosing one brand of beverage over others, this occurs only when they are thirsty [118]. Similarly, anchoring effects are found only when the anchor is presented as informative [119,120]. It is accordingly rather difficult to generalize from such special-case results.

exclusive nor exhaustive) based on the experimental manipulations (see [Figure 1](#) for a schematic description of some key findings). Studies in all these categories have attracted substantial attention as well as highly convincing criticism ([Box 2](#)). Importantly, we bracket such methodological concerns here and ask the following more fundamental question: assuming these findings are veridical, what kind of threat, if any, do they pose to free will?

#### Studies manipulating the context of the decision, or the way in which it is framed

Here, the contents of the decision, including the decision alternatives, remain constant, while the context or presentation of those alternatives is manipulated. For example, in Framing Effects, the framework within which the information is presented influences the decision. For example, seminal work by Kahneman and Tversky famously showed that people were less willing to buy another theater ticket after having lost one, compared with losing the same amount of money as the worth of the ticket [20]. In Nudging Effects ([Figure 1A](#)), at least one of the decision components (the alternatives, their outcomes, or the correspondence between them) is presented in a manner that favors a specific alternative [21]. For instance, presenting one alternative as the default, or changing the presentation order of the alternatives, increases the probability that people will choose it, even in pivotal decisions such as whether to become an organ donor [22], take an HIV test [23], reduce meat consumption [24], or transfer personal data to a third party [25].

Another line of work focuses on Anchoring Effects ([Figure 1B](#); e.g., [26,27]), where the value of some comparison item influences subjects' judgments, affecting, for example, the price people are willing to pay for a product [28], or the sentence they find fitting for a crime [29]. In Placement Effects, the physical placement of the alternatives affects the decision ([Figure 1C](#)): for example, a product located in the center of a display is selected more often than the same product placed to the side [18].

Mere Exposure Effects ([Figure 1D](#)) occur when prior exposure to a stimulus makes it appear more favorable (e.g., [30,31]). Action-related exposure can also bias future decisions: when subjects are instructed to respond to a sound, and that sound is arbitrarily paired with specific items, these items are preferred in a subsequent session [32], even after 6 months [33].

#### Studies manipulating the content of the decision in irrelevant ways

These studies manipulate the information available to subjects about aspects of the decision typically considered irrelevant to the decision (e.g., the race of a defendant in a criminal trial). This allows one to study the effect of stereotypes (assumptions about the attributes, behaviors, and characteristics of different groups [34]) on judgments about individuals in the group [35]. Such studies are typically carried out in the context of medical treatments, jury trials, or work environments ([Figure 1E](#)). In judicial decisions, race and gender were found to affect sentencing in both meta-analyses of actual trials and in lab experiments. The race and gender of the defendants, victims, and jurors were reported to influence the given sentence for different crimes to different degrees (e.g., [36–41]). Some of these effects might be at least partly due to perceptual biases; for example, Black men are perceived as bigger and more physically threatening compared with white men [42]. Similarly, gender was found to influence people's assessment of their peer's professional success. One study reported that grant funding success rate is higher for males when applications include an explicit description of the investigator, but not when such descriptions are omitted [43] (see also [44]).

Appearance also biases decisions: candidates perceived as more beautiful won more votes in actual elections [45], and face-based trustworthiness judgments correlate with the chance of getting a death sentence [46]. Even more implicit measures, such as the face width–height ratio

#### Glossary

**Agency:** capacity to act autonomously; the exercise of a capacity to act.

**Arbitrary decisions:** decisions that are not based on reasons; sometimes referred to as 'picking'.

**Compatibilism:** determinism is consistent with free action and free will.

**Control:** ability to bring something (an object, some state of affairs, or your own future action) to be as you would have it to be.

**Deliberate decisions:** reasons-based decisions; sometimes referred to as 'choosing'.

**Determinism:** every event, including every decision and action, has causes that together guarantee that this event will occur instead of any alternative event.

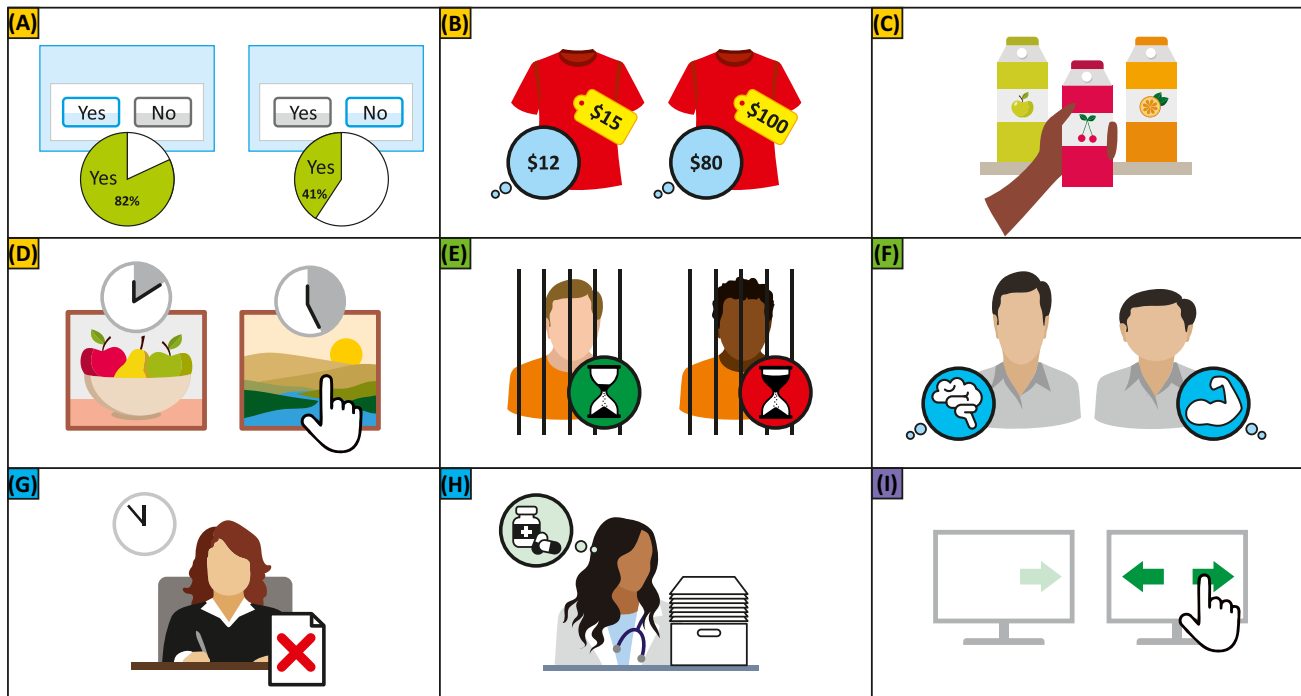
**Emergentism:** (about minds) some mental properties or events are not identical with any physical property or event (or any property or event at a lower level), but they still depend or supervene on (or are grounded in) some physical realizer or physical basis, either by being realized by it (weak emergence) or merely caused by it (strong emergence). The relation of emergence is not symmetrical, because the mental emerges from the physical but not vice versa.

**Incompatibilism:** determinism is inconsistent with free action and free will.

**Indeterminism:** for some events, including some decisions and actions, there is no set of causes that guarantees or determines that they will occur instead of any alternative. Causes of such events might still increase the probability of their occurrence.

**Reasons-responsiveness:** decision, act, or agent is responsive to reasons to the extent that the person, or a relevant subpersonal mechanism, recognizes and reacts to reasons when they are present.

**Reductionism:** (about minds) every mental property or event is identical with a physical property or event (or some property or event at a lower level). The relation of identity is reflexive, symmetric, and transitive.



Trends in Cognitive Sciences

**Figure 1. Different types of unconscious effects on decisions.** Schematic of typical findings of unconscious effects on decisions. (A–D) (yellow): experiments manipulating the context of the decision. (A) Nudging: marking one option as the default heavily biases the likelihood of subjects choosing it [22]. (B) Anchoring: subjects pay more for the same product based on its price tag [28]. (C) Placement effects: subjects are more likely to choose the product in the middle [18]. (D) Mere exposure: preference depends on exposure time; thus, subjects prefer items to which they were exposed for longer [31]. (E, F) (green): experiments manipulating the content of the decision. (E) Stereotype effects: Black/Hispanic defendants are more likely to get a higher sentence than are white defendants [39,40], at least for some crimes [41]. (F) Appearance effects: people with a lower face width–height ratio (fWHR) are considered more intelligent, while those with higher fWHR are considered more physically able [19]. (G, H) (blue): experiments manipulating the state of the agent making the decision. (G) Physiological effects: judges were more likely to approve paroles at the beginning of the workday or after their lunch break, arguably due to being less hungry or tired [50] (but see [51]). (H) Fatigue effects: toward the end of the day, after making many decisions, physicians tend to prefer more conservative treatments [53]. (I) (purple): experiments manipulating awareness of decision-related factors. Masking: an invisible prime (left) biases a subsequent decision on a visible choice (right) [59]. Importantly, for this review, we take these findings at face value; for criticism, see [Box 1](#) in the main text.

(fWHR; [Figure 1F](#)), influence character ascriptions, or even how ‘human’ one is held to be [47]. For example, people with higher fWHR are associated with less sophisticated crimes and are more likely to be selected as team members for football than for chess [19,47].

In another line of studies, subjects made different decisions depending on the manipulation of some decision factors, but, when asked which factors guided their decisions, subjects’ responses did not reflect the factors that were manipulated [48]. Thus, subjects appeared unaware of the factors causally influencing their decisions, and/or have little insight into the actual processes affecting decision-making [49] (although see [Box 2](#)).

#### Studies manipulating the state of the agent making the decision

Here, Physiological Effects on decisions are examined, focusing on physical states of the decider that should be irrelevant to the decision, such as fatigue or hunger ([Figure 1G](#)). For example, judges were more likely to award parole at the beginning of the day or after a lunch break [50] (but see [51] and [Box 2](#)). Similarly, sleep-deprived subjects acted less fairly toward others [52].

Fatigue can also influence decision processes; for example, in Decision Fatigue studies (Figure 1H), making numerous mentally taxing decisions requiring attention is claimed to affect subsequent decisions (e.g., resort to heuristics). Toward the end of the day, surgeons are less likely to schedule a medical operation and prefer more conservative, default treatments [53]. These effects were claimed to support Ego Depletion, the theory that we have limited resources for volitional actions, such that later decisions are harmed when resources are drained [54]. Arguably, depleted participants tend to trust others less [55], act less ethically [56,57], and use heuristics more when making their decisions [58]. The phenomenon has been contested (Box 2), but our question here is whether it threatens free will if these effects are genuine.

#### Studies manipulating awareness of some of the decision components

Here, studies manipulate whether the subject is aware of critical information in the decision content or context (Figure 1I). In the aforementioned studies, participants were aware of all relevant information but were arguably unaware of its influence on their decision-making. Yet, here people are unaware of the influencing factor itself (Figure 1I). These manipulations have been done both for meaningless, **arbitrary decisions** (e.g., whether to press a right or left button [59–61]) and with meaningful, deliberate ones (e.g., consumer/moral decisions; see [62–64] for how this distinction matters for the free will debate).

Substantial evidence shows that arbitrary decisions can be unconsciously biased. A more complicated picture emerges for meaningful, deliberative decisions. In consumer studies, people can typically be biased toward a certain brand or alternative, but only when deprived (e.g., [65–67]). Some reported subliminal (and typically small) effects in the political domain: national flags led participants to vote for more central parties [68], pro-social word primes increased donations by socially oriented participants [69], disgusting faces reduced moral ratings of utilitarian decisions [70], and direct-gaze faces primes affected behavior in prisoner's dilemma games [71].

#### Do these findings threaten free will?

The aforementioned studies suggest that humans are not always conscious of all the factors affecting their decisions (but see Box 2), including factors that even they would deem completely irrelevant to the decision, or even morally problematic. Does this render those decisions less free?

The concept of free will has been extensively discussed and reformulated over many years in philosophy (reviewed in [1,72]), yielding a plethora of theories. We reiterate that our discussion does not concern two more often discussed threats to freedom: **determinism** and **reductionism** (Box 1). Instead, we focus on the independent challenge posed by unconscious psychological processes, asking how, if at all, the reviewed findings might inform accounts of free will. Similarly, our discussion also does not concern the subjective aspect of free will (i.e., whether we experience our actions as free, or feel a sense of **agency**; for discussions about the relations between agency and free will, see [73,74]). Instead, our discussion here focuses on the metaphysical question of whether our decisions and actions are indeed freely willed. For the purposes of this discussion, we assume that free will is a metaphysical possibility, and ask how findings of unconscious effects on decision-making might influence free will in actuality.

#### The importance of consciousness for free will

Consciousness is intuitively necessary for free will, and some philosophers have defended this view [75]. Free will cannot be simply identified with conscious decision-making, since some conscious decisions are not free (e.g., if coerced). Libet's studies [4] have been famously interpreted as showing that decisions precede the conscious experience of deciding (although see [11]), and were accordingly held to be unfree [76]. However, what kind of consciousness is necessary for

free will, and why? We identify five ways in which consciousness can be involved in decision-making; all but the first refers to the object of consciousness (i.e., what are we aware of). Similarly, all but the first refer to the access we have to the experienced decision-related information, rather than to the qualitative nature of that experience [77], or how it feels to have it [78]. That is, we focus on the effects of not having conscious access to various aspects of the decision-making process, rather than the qualitative character of the experience of deciding.

The first kind of consciousness we identify is 'creature consciousness': the ability to have experiences at all (as opposed to being incapable of experiencing things, or losing that ability, for example, while in a coma). This arguably necessary precondition for freedom was of course present in all the studies we canvassed: an agent must be conscious while making a free choice (P). Next, we ask, for that conscious agent, what are they conscious of, when making a decision? We divide the possible answers into four categories.

Category A: awareness of the external (A1)/internal (A2) factors influencing the decision. For example, is the agent aware of the skin color of the defendant, or of their own fatigue? While being aware of some of the factors involved in decisions appears important for deliberate decisions, to require that agents be consciously aware of them all is excessive, for that would include, for instance, all prior causes of the immediate factors that affect the decision.

Category B: awareness of the outcome of the decision (i.e., its content) or of performing the action by which this decision is implemented. This too appears important to free will, because having decided appears to require knowing that a decision has been reached and implemented and what that decision is.

Category C: awareness of the relations between influencing factors and the outcome (i.e., between A and B); that is, the way in which factors influencing the decision affect its resolution. For instance, a customer might be aware that a certain item was placed in the middle and also aware that they opted for that item but still be oblivious to the fact that its location influenced their choice (i.e., the relation between the influence and the outcome).

Category D: awareness of the decision process itself. For example, you may turn left out of habit while driving a very familiar route and thinking of other things. In that case, due to inattention, you probably were not consciously aware that you were deciding (D1). Similarly, we can ask whether the agent needs to be consciously aware of consciously making the decision (i.e., to have a conscious metacognitive judgment of making a decision; D2). Such higher-order awareness does not appear necessary for free will (but see [79]); in fact, in all the reviewed studies, this was not even tested. Importantly, in real life, we rarely make such meta-judgments over our decisions, including those we regard as free.

Using these categories, we classified studies from ~60 empirical papers to get a rough estimate of the ways in which consciousness is manipulated in such studies (Figure 2; and see the supplemental information online). Note that the reviewed papers are not an exhaustive representation of the literature, but a collection of key papers from the different categories. Accordingly, their assignment to categories should be taken as an illustration of the way in which the types of consciousness come into play in such studies, not an accurate quantification encompassing all studies in the field. Still, this illustration does demonstrate the different roles that consciousness can have in the decision-making process: while, in some studies, subjects were unaware of some of the factors affecting their decisions (external/internal, category A), in most studies (of almost all experimental manipulations), subjects were unaware of the influence of some factors



Trends In Cognitive Sciences

Figure 2. How is consciousness involved in decisions? An illustration of our classification of reviewed papers according to the ways in which consciousness can be involved in the decision-making process. At the top left is a precondition criterion: being conscious while making the decision (P). Then, there are two criteria for being conscious of either the external (A1) or internal (A2) factors weighing into the decision. These are followed by a criterion for being conscious of executing the decision (B), and for being conscious of the relations between A and B (C). Finally, being conscious of deciding (D1), and of consciously deciding (D2). Blue shades mark consciousness, red shades mark its absence, gray denotes a situation in which there was no information in the study that allowed us to determine whether subjects met the criterion, and white represents that this criterion was not relevant to the research question (which was the case for all reviewed studies with respect to meta-cognition of consciously deciding; P3). Next to each pie slice is a list of all types of paper in that category (no list appears when there is only one slice, as all papers are included). A certain type of studies can appear for more than one slice. As the figure shows, the most widely shown effect concerns the lack of consciousness of the relations between the factors affecting the decision, and its outcome (for the full articles and their classifications, see the supplemental information online). Note again that this is just a rough estimation of the types of consciousness relevant to each category of studies, rather than an exhaustive analysis.

on their decision (category C). Thus, we posit that lack of awareness of how decisions are affected by contextual factors constitutes the major potential threat to free will posed by these studies.

How lack of consciousness could threaten free will

There are two main reasons that consciousness could be important to free will (see [80] for further discussion). First, freely willing something implies a certain amount of **control**. Being conscious of

our motivations, reasons, decisions, and actions appears to endow us with a kind of executive control that we lack when these are not conscious. For example, being aware of the result of a decision process is arguably necessary for the ability to re-evaluate or change one's mind.

Second, various types of consciousness may be necessary for **reasons-responsiveness**. Here, a 'reason' is a consideration for/against acting in a certain way, rather than a factor explaining why the person in fact acted the way they did. According to some [81], an action or decision can be free even if it is causally determined, as long as the person is sufficiently responsive to reasons; that is, if the person has the ability to recognize and react to reasons for/against acting [81–83]. Having a reduced/compromised ability to respond to reasons [84,85], or suffering from compulsions/delusions that make one unable to recognize or react appropriately to reasons, can render an action unfree.

### Do the reviewed findings undermine either control or reasons-responsiveness?

Evidence that some decisions can be unconsciously biased implies that they are not always guided by the reasons we would take to justify our decision. This poses a challenge to free will if these factors sufficiently impair our ability either to respond to reasons or to control our activity. Given that reasons-responsiveness and control vary with context and in degree, freedom may also come in degrees [83,86]: some agents are more responsive to reasons than are others, and can be more/less reasons-responsive in different situations (e.g., less while drunk) [87]. Similarly, circumstances may leave us more/less in control of our own activity, rendering it more/less free. Importantly, we can be reasons-responsive even if our decisions are not based solely on reasons, as long as enough other reasons are, or can be, taken into account (e.g., when sentencing is affected by not only the defendant's race, but also their danger to society), at least if these other reasons are not also adversely affected by unconscious influences (such as implicit bias) or are illusory or based on false information. Likewise, we can be in control of our actions even if those actions are influenced by factors we do not control. Accordingly, the reviewed studies cannot refute the existence of free will conceived of as dependent upon reasons-responsiveness or control, provided that agents are sufficiently reasons-responsive and in control. To refute the general existence of free will, thus understood, one would need to show that agents are never able to respond to reasons or that they lack control over their behavior. This, of course, has not been demonstrated by any of the reviewed experiments. However, embracing the notion of degrees of freedom [83,86], these experiments demonstrate ways in which control or reasons-responsiveness are compromised or imperfect. They can accordingly help us delineate both the limits and the extent of our freedom, by distilling the contributions of reasons-related and non-reasons-related factors on decisions.

Subliminal effects appear to stand out in that respect (Figure 1): arguably, if our decisions can be biased by factors of which we have no awareness at all, how can we control them? Yet, this seemingly strong threat is highly mitigated upon closer inspection. Even disregarding methodological challenges (which are not trivial: Box 2), the relevance of these findings for freedom is limited, and depends both on one's account of freedom and on the size of the effect. For example, those who think of freedom as reasons-responsiveness will want to differentiate between arbitrary decisions, in which the strongest effects were found, and deliberate decisions, in which effects are smaller and depend on the agent's state. If most effects are on arbitrary decisions, the threat to reasons-responsiveness is smaller, as by definition, responding to reasons will not resolve an arbitrary decision. Furthermore, current studies provide no evidence for subliminal priming overcoming convictions that subjects strongly hold. If primes simply strengthen, activate, or bring to awareness pre-existing reasons, that does not show a lack of reason-responsiveness, although it may threaten control. However, threats to control are



mitigated by the effects being small and typically short-lived [88,89]. Thus, currently, this line of research does not provide strong evidence against free will.

The preceding discussion highlights the importance of understanding the factors influencing behavior. While some factors may partially undermine control, coming to be aware of them ought to enhance it: arguably, the more we are aware of the factors affecting our decisions, the more we can implement strategies to control them, if we like, and, thus, the freer we are. The more ignorant or deluded we are about the causes of our actions, the less control and freedom we have. Thus, while subliminal effects can somewhat degrade freedom, coming to understand them may allow greater control over our responses. Being more attuned to the different biases that might affect decisions can enhance freedom, especially if this knowledge allows us to circumvent these effects. Thus, rather than undermining freedom, learning about different types of unconscious influences on decisions can increase it.

### Can any empirical evidence threaten free will?

We argued that existing evidence for unconscious influences on decision-making does not undermine free will. First, it does not speak to concerns about determinism or reductionism. Second, it does not show that we lack reasons-responsiveness or control. Nonetheless, these studies can reveal ways in which specific decisions and actions are less than fully free, when either control or reasons-responsiveness are mitigated. We now ask whether any empirical findings about unconscious influences show a particular action to be unfree. We characterize two ideal candidates here:

**Subliminal priming:** if an invisible prime prompted subjects to select an option that is incompatible with their conscious preferences in a nonarbitrary decision, this would convincingly render the decision unfree. Of course, as noted earlier, manipulations that simply push subjects toward existing attitudes, or facilitate the activation of a reason, would not suffice to demonstrate a complete lack of reasons-responsiveness. What needs to be shown is that subjects act in a way, or based on a reason, that is contrary to their own. Finally, to remove all hope for control, subjects should be unable to overcome priming, even when made aware of its existence (for an opposite result, see [66]). Of course, unawareness of the prime must be convincingly demonstrated (Box 2).

**Nudging:** if subjects were unable to consciously overcome a known nudging effect when they had a reason to overcome it (i.e., continuing to choose the middle product even when knowing about this bias, and having a reason to not act on the bias), that would be strong evidence that these choices are not reasons-responsive and, thus, not free.

### Concluding remarks

Current findings of unconscious effects on decision-making do not threaten the existence of free will in general. Yet, the results still show ways in which our freedom can be compromised under specific circumstances. More experimental and philosophical work is needed to delineate the limits and scope of these effects on our freedom (see [Outstanding questions](#)). We have evolved to be the decision-makers that we are; thus, our decisions are affected by biases, internal states, and external contexts. However, we can at least sometimes resist those, if we want, and this ability to resist influences contrary to our preferences and reasons is considered a central feature of freedom. As long as this ability is preserved, and the reviewed findings do not suggest otherwise, we are still free, at least usually and to a significant degree.

### Acknowledgments

This publication was made possible through the support of a joint grant from the John Templeton Foundation and the Fetzer Institute. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the

### Outstanding questions

How many of the reviewed studies survive methodological scrutiny and prove to be robust and reproducible?

To what extent do findings generalize to real-life, everyday decisions? Although some of the above-mentioned studies did pertain to decisions made in everyday life, many of them were based on laboratory findings. More work is needed to determine whether, and to what extent, these translate to real-life situations.

To what degree can subjects resist different unconscious effects on decision-making? Are certain effects harder to resist than others?

Can we quantify the degrees of freedom for a given decision, and can those be parametrically manipulated, to enable a systematic investigation of freedom?

Are arbitrary decisions indeed more prone to unconscious influences compared with deliberate or meaningful ones?

John Templeton Foundation or the Fetzer Institute. L.M. is CIFAR Tanenbaum Fellow in the Brain, Mind, and Consciousness program.

**Declaration of interests**

None declared by authors.

**Supplemental Information**

Supplemental information associated with this article can be found online at <http://doi.org/10.1016/j.tics.2022.03.005>.

**References**

1. O'Connor, T. and Franklin, C. (2021) Free will. In *The Stanford Encyclopedia of Philosophy* (Zalta, E., ed.), Metaphysics Research Lab
2. Frede, M. (2011) *A Free Will: Origins of the Notion in Ancient Thought*, University of California Press
3. Dohle, A. (2020) *The Theory of Will in Classical Antiquity*, University of California Press
4. Libet, B. et al. (1983) Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain* 106, 623–642
5. Soon, C.S. et al. (2013) Predicting free choices for abstract intentions. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6217–6222
6. Fried, I. et al. (2011) Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron* 69, 548–562
7. Vargas, M. (2013) *If Free Will Doesn't Exist*, Neither Does Water, Lexington Books
8. Schlegel, A. et al. (2013) Barking up the wrong free: readiness potentials reflect processes independent of conscious will. *Exp. Brain Res.* 229, 329–335
9. Schurger, A. et al. (2012) An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2904–E2913
10. Roskies, A.L. (2010) How does neuroscience affect our conception of volition? *Annu. Rev. Neurosci.* 33, 109–130
11. Schurger, A. et al. (2021) What is the readiness potential? *Trends Cogn. Sci.* 25, 558–570
12. Travers, E. et al. (2021) The Readiness Potential reflects planning-based expectation, not uncertainty, in the timing of action. *Cogn. Neurosci.* 12, 14–27
13. Doris, J.M. (2002) *Lack of Character: Personality and Moral Behavior*, Cambridge University Press
14. Nelkin, D.K. (2005) Freedom, responsibility and the challenge of situationism. *Midwest. Stud. Philos.* 29, 181–206
15. Vargas, M. (2013) Situationism and moral responsibility. In *Decomposing the Will* (Clark, A. et al., eds), pp. 325–350, Oxford University Press
16. Nahmias, E. (2014) Is free will an illusion? Confronting challenges from the modern mind sciences. In *Moral Psychology* (Sinnott-Armstrong, W., ed.), pp. 1–26, MIT Press
17. Nahmias, E. (2007) In *Autonomous agency and social psychology. In Cartographies of the Mind* (Maraffa, M. et al., eds), pp. 169–185, Springer
18. Keller, C. et al. (2015) Nudging product choices: the effect of position change on snack bar choice. *Food Qual. Prefer.* 41, 41–43
19. Hehman, E. et al. (2015) Picking teams: when dominant facial structure is preferred. *J. Exp. Soc. Psychol.* 59, 51–59
20. Kahneman, D. and Tversky, A. (2013) Choices, values, and frames. In *Handbook of the Fundamentals of Financial Decision Making: Part I* (MacLean, L.C. and Ziemba, W.T., eds), pp. 269–278, World Scientific
21. Thaler, R.H. and Sunstein, C.R. (2011) *Nudge: Improving Decisions about Health, Wealth, and Happiness*, HeinOnline
22. Johnson, E.J. and Goldstein, D. (2003) Do defaults save lives? *Science* 302, 338–339
23. Montoy, J.C.C. et al. (2016) Patient choice in opt-in, active choice, and opt-out HIV screening: randomized clinical trial. *BMJ* 352, h6895
24. Kurz, V. (2018) Nudging to reduce meat consumption: Immediate and persistent effects of an intervention at a university restaurant. *J. Environ. Econ. Manag.* 90, 317–341
25. Sakshaug, J.W. et al. (2016) Evaluating active (opt-in) and passive (opt-out) consent bias in the transfer of federal contact data to a third-party survey agency. *J. Surv. Stat. Methodol.* 4, 382–416
26. Andersson, H. et al. (2021) Anchoring effect in judgments of objective fact and subjective preference. *Food Qual. Prefer.* 88, 104102
27. Yoon, S. et al. (2019) The robustness of anchoring effects on preferential judgments. *Judgm. Decis. Mak.* 14, 470–487
28. Ariely, D. et al. (2003) “Coherent arbitrariness”: Stable demand curves without stable preferences. *Q. J. Econ.* 118, 73–106
29. Englich, B. and Soder, K. (2009) Moody experts—how mood and expertise influence judgmental anchoring. *Judgm. Decis. Mak.* 4, 41
30. Armel, K.C. et al. (2008) Biasing simple choices by manipulating relative visual attention. *Judgm. Decis. Mak.* 3, 396–403
31. Mrkva, K. and Van Boven, L. (2020) Salience theory of mere exposure: Relative exposure increases liking, extremity, and emotional intensity. *J. Pers. Soc. Psychol.* 118, 1118
32. Schonberg, T. et al. (2014) Changing value through cued approach: an automatic mechanism of behavior change. *Nat. Neurosci.* 17, 625–630
33. Salomon, T. et al. (2018) The cue-approach task as a general mechanism for long-term non-reinforced behavioral change. *Sci. Rep.* 8, 1–13
34. Hilton, J.L. and Von Hippel, W. (1996) Stereotypes. *Annu. Rev. Psychol.* 47, 237–271
35. Hamilton, D.L. and Sherman, J.W. (2014) *Stereotypes*, Psychology Press
36. Doerner, J.K. and Demuth, S. (2014) Gender and sentencing in the federal courts: are women treated more leniently? *Crim. Justice Policy Rev.* 25, 242–269
37. Holcomb, J.E. et al. (2004) White female victims and death penalty disparity research. *Justice Q.* 21, 877–902
38. Williams, M.R. et al. (2007) Understanding the influence of victim gender in death penalty cases: the importance of victim race, sex-related victimization, and jury decision making. *Criminology* 45, 865–891
39. Mitchell, T.L. et al. (2005) Racial bias in mock juror decision-making: a meta-analytic review of defendant treatment. *Law Hum. Behav.* 29, 621–637
40. Steffensmeier, D. and Demuth, S. (2000) Ethnicity and sentencing outcomes in US federal courts: who is punished more harshly? *Am. Sociol. Rev.* 65, 705–729
41. Mazzella, R. and Feingold, A. (1994) The effects of physical attractiveness, race, socioeconomic status, and gender of defendants and victims on judgments of mock jurors: a meta-analysis 1. *J. Appl. Soc. Psychol.* 24, 1315–1338
42. Wilson, J.P. et al. (2017) Racial bias in judgments of physical size and formidability: from size to threat. *J. Pers. Soc. Psychol.* 113, 59
43. Wittman, H.O. et al. (2019) Are gender gaps due to evaluations of the applicant or the science? A natural experiment at a national funding agency. *Lancet* 393, 531–540
44. Mengel, F. et al. (2019) Gender bias in teaching evaluations. *J. Eur. Econ. Assoc.* 17, 535–566

45. Berggren, N. *et al.* (2010) The looks of a winner: beauty and electoral success. *J. Public Econ.* 94, 8–15
46. Wilson, J.P. and Rule, N.O. (2015) Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychol. Sci.* 26, 1325–1331
47. Deska, J.C. *et al.* (2018) Facing humanness: facial width-to-height ratio predicts ascriptions of humanity. *J. Pers. Soc. Psychol.* 114, 75–94
48. Nisbett, R.E. and Bellows, N. (1977) Verbal reports about causal influences on social judgments: private access versus public theories. *J. Pers. Soc. Psychol.* 35, 613
49. Nisbett, R.E. and Wilson, T.D. (1977) Telling more than we can know: verbal reports on mental processes. *Psychol. Rev.* 84, 231
50. Danziger, S. *et al.* (2011) Extraneous factors in judicial decisions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6889–6892
51. Weinshall-Margel, K. and Shapard, J. (2011) Overlooked factors in the analysis of parole decisions. *Proc. Natl. Acad. Sci. U.S.A.* 108, E833
52. Dickinson, D.L. and McElroy, T. (2017) Sleep restriction and circadian effects on social decisions. *Eur. Econ. Rev.* 97, 57–71
53. Persson, E. *et al.* (2019) The effect of decision fatigue on surgeons' clinical decision making. *Health Econ.* 28, 1194–1203
54. Baumeister, R.F. *et al.* (2000) Ego depletion: a resource model of volition, self-regulation, and controlled processing. *Soc. Cogn.* 18, 130–150
55. Ainsworth, S.E. *et al.* (2014) Ego depletion decreases trust in economic decision making. *J. Exp. Soc. Psychol.* 54, 40–49
56. Gino, F. *et al.* (2011) Unable to resist temptation: how self-control depletion promotes unethical behavior. *Organ. Behav. Hum. Decis. Process.* 115, 191–203
57. Mead, N.L. *et al.* (2009) Too tired to tell the truth: self-control resource depletion and dishonesty. *J. Exp. Soc. Psychol.* 45, 594–597
58. Masicampo, E.J. and Baumeister, R.F. (2008) Toward a physiology of dual-process reasoning and judgment: lemonade, willpower, and expensive rule-based analysis. *Psychol. Sci.* 19, 255–260
59. Furstenberg, A. *et al.* (2015) Evidence of change of intention in picking situations. *J. Cogn. Neurosci.* 27, 2133–2146
60. Schlaghecken, F. and Eimer, M. (2004) Masked prime stimuli can bias “free” choices between response alternatives. *Psychon. Bull. Rev.* 11, 463–468
61. Kiesel, A. *et al.* (2006) Unconscious manipulation of free choice in humans. *Conscious. Cogn.* 15, 397–408
62. Maoz, U. *et al.* (2019) Neural precursors of decisions that matter—an ERP study of deliberate and arbitrary choice. *eLife* 8, e39787
63. Travers, E. and Haggard, P. (2021) The Readiness Potential reflects the internal source of action, rather than decision uncertainty. *Eur. J. Neurosci.* 53, 1533–1544
64. Vierkant, T. *et al.* (2019) Responsibility without freedom? Folk judgements about deliberate actions. *Front. Psychol.* 10, 1133
65. Smarandescu, L. and Shimp, T.A. (2015) Drink coca-cola, eat popcorn, and choose powerade: testing the limits of subliminal persuasion. *Mark. Lett.* 26, 715–726
66. Verwijmeren, T. *et al.* (2013) Warning: you are being primed! The effect of a warning on the impact of subliminal ads. *J. Exp. Soc. Psychol.* 49, 1124–1129
67. Verwijmeren, T. *et al.* (2011) The workings and limits of subliminal advertising: the role of habits. *J. Consum. Psychol.* 21, 206–213
68. Hassin, R.R. *et al.* (2007) Subliminal exposure to national flags affects political thought and behavior. *Proc. Natl. Acad. Sci. U.S.A.* 104, 19757–19761
69. Andersson, O. *et al.* (2017) Subliminal influence on generosity. *Exp. Econ.* 20, 531–555
70. Lim, J. *et al.* (2017) Moral judgment modulation by disgust priming via altered fronto-temporal functional connectivity. *Sci. Rep.* 7, 1–14
71. Luo, Y. *et al.* (2016) The power of subliminal and supraliminal eye contact on social decision making: An individual-difference perspective. *Conscious. Cogn.* 40, 131–140
72. Maoz, U. and Sinnott-Armstrong, W., eds (2021) *Free Will: Philosophers and Neuroscientists in Conversation*, Oxford University Press
73. Pereboom, D. (2014) *Free Will, Agency, and Meaning in Life*, Oxford University Press
74. Mele, A.R. (2017) *Aspects of Agency: Decisions, Abilities, Explanations, and Free Will*, Oxford University Press
75. Levy, N. (2014) *Consciousness and Moral Responsibility*, Oxford University Press
76. Wegner, D.M. (2002) *The Illusion of Conscious Will*, MIT Press
77. Block, N. (1995) On a confusion about a function of consciousness. *Behav. Brain Sci.* 18, 227–287
78. Nagel, T. (1974) What is it like to be a bat? *Philos. Rev.* 83, 435–450
79. Vierkant, T. (2018) Choice in a two systems world: picking & weighing or managing & metacognition. *Phenomenol. Cogn. Sci.* 17, 1–13
80. Roskies, A.L. (2010) Freedom, neural mechanism, and consciousness. In *Free Will and Consciousness: How Might They Work* (Baumeister, R.F. *et al.*, eds), pp. 153–171, Oxford University Press
81. Fischer, J.M. and Ravizza, M. (2000) *Responsibility and Control: A Theory of Moral Responsibility*, Cambridge University Press
82. Brink, D.O. and Nelkin, D.K. (2013) Fairness and the architecture of responsibility. *Oxf. Stud. Agency Responsib.* 1, 284–313
83. Vargas, M. (2013) *Building Better Beings: A Theory of Moral Responsibility*, Oxford University Press
84. McKenna, M. (2013) Responsibility & globally manipulated agents. In *The Philosophy of Free Will: Essential Readings from the Contemporary Debates* (Russell, P. and Deery, O., eds), pp. 342–359, Oxford University Press
85. Sripada, C.S. (2012) What makes a manipulated agent unfree? *Philos. Phenomenol. Res.* 85, 563–593
86. Nelkin, D.K. (2016) Difficulty and degrees of moral praiseworthiness and blameworthiness. *Noûs* 50, 356–378
87. Coates, D.J. and Swenson, P. (2013) Reasons-responsiveness and degrees of responsibility. *Philos. Stud.* 165, 629–645
88. Greenwald, A.G. *et al.* (1996) Three cognitive markers of unconscious semantic activation. *Science* 273, 1699–1702
89. Rothkirch, M. and Hesselmann, G. (2017) What we talk about when we talk about unconscious processing—a plea for best practices. *Front. Psychol.* 8, 835
90. Suppes, P. (1993) The transcendental character of determinism. *Midwest. Stud. Philos.* 18, 242–257
91. Popper, K. and Bartley, I. (2012) *The Open Universe: An Argument for Indeterminism from the Postscript to the Logic of Scientific Discovery* Routledge
92. Kane, R. (1998) *The Significance of Free Will*, Oxford University Press
93. Pereboom, D. (2006) *Living without Free Will*, Cambridge University Press
94. Clarke, R. (1993) Toward a credible agent-causal account of free will. *Noûs* 27, 191–203
95. O'Connor, T. (2002) *Persons and Causes: The Metaphysics of Free Will*, Oxford University Press
96. Dennett, D.C. (1984) *Elbow Room: The Varieties of Free Will Worth Wanting*, MIT Press
97. Fischer, J.M. (2007) Compatibilism. In *Four Views on Free Will* (Fischer, J.M. *et al.*, eds), pp. 44–84, Blackwell
98. Wolf, S. (1990) *Freedom Within Reason*, Oxford University Press
99. Roskies, A.L. (2006) Neuroscientific challenges to free will and responsibility. *Trends Cogn. Sci.* 10, 419–423
100. Newsome, W.T. (2009) Human Freedom “Emergence”. In *Downward Causation and the Neurobiology of Free Will* (Murphy, N. *et al.*, eds), pp. 53–62, Springer
101. O'Connor, T. (2021) Emergent properties. In *The Stanford Encyclopedia of Philosophy* (Zalta, E., ed.), Metaphysics Research Lab
102. Sinnott-Armstrong, W. (2021) Contrastive mental causation. *Synthese* 198, 861–883
103. Robinson, W. (2019) Epiphenomenalism. In *The Stanford Encyclopedia of Philosophy* (Zalta, E., ed.), Metaphysics Research Lab

104. Huxley, T.H. (1874) On the hypothesis that animals are automata, and its history. *The Fortnightly Review* 16, 555–580
105. Mors, M.R. *et al.* (2018) Can food choice be influenced by priming with food odours? *Food Qual. Prefer.* 66, 148–152
106. Maniadis, Z. *et al.* (2014) One swallow doesn't make a summer: new evidence on anchoring effects. *Am. Econ. Rev.* 104, 277–290
107. Hummel, D. and Maedche, A. (2019) How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. *J. Behav. Exp. Econ.* 80, 47–58
108. Hagger, M.S. *et al.* (2016) A multilab preregistered replication of the ego-depletion effect. *Perspect. Psychol. Sci.* 11, 546–573
109. Vohs, K.D. *et al.* (2021) A multi-site preregistered paradigmatic test of the ego depletion effect. *Psychol. Sci.* 32, 1566–1581
110. Sripada, C.S. *et al.* (2014) Methylphenidate blocks effort-induced depletion of regulatory control in healthy volunteers. *Psychol. Sci.* 25, 1227–1234
111. Carter, E.C. *et al.* (2015) A series of meta-analytic tests of the depletion effect: self-control does not seem to rely on a limited resource. *J. Exp. Psychol. Gen.* 144, 796
112. Friese, M. *et al.* (2019) Is ego depletion real? An analysis of arguments. *Personal. Soc. Psychol. Rev.* 23, 107–131
113. Newell, B.R. and Shanks, D.R. (2014) Unconscious influences on decision making: a critical review. *Behav. Brain Sci.* 37, 1–19
114. Shanks, D.R. (2017) Regressive research: the pitfalls of post hoc data selection in the study of unconscious mental processes. *Psychon. Bull. Rev.* 24, 752–775
115. Peters, M.A.K. *et al.* (2017) Does unconscious perception really exist? Continuing the ASSC20 debate. *Neurosci. Conscious.* 2017, nix015
116. Karremans, J.C. *et al.* (2006) Beyond Vicary's fantasies: the impact of subliminal priming and brand choice. *J. Exp. Soc. Psychol.* 42, 792–798
117. Strahan, E.J. *et al.* (2002) Subliminal priming and persuasion: striking while the iron is hot. *J. Exp. Soc. Psychol.* 38, 556–568
118. Bermeitinger, C. *et al.* (2009) The hidden persuaders break into the tired brain. *J. Exp. Soc. Psychol.* 45, 320–326
119. Shanks, D.R. *et al.* (2020) Do incidental environmental anchors bias consumers' price estimations? *Collabra. Psychology* 6, 19
120. Ioannidis, K. *et al.* (2020) On the effect of anchoring on valuations when the anchor is transparently uninformative. *J. Econ. Sci. Assoc.* 6, 77–94