# Stimulus awareness is necessary for both instrumental learning and instrumental responding to previously learned stimuli

Lina I. Skora, Ryan B. Scott, Gerhard Jocham

Wissen, wo das Wissen ist.

UNIVERSITÄTS- UND
LANDESBIBLIOTHEK
DÜSSELDORF

# Stimulus awareness is necessary for both instrumental learning and instrumental responding to previously learned stimuli

Lina I. Skora [a,b,*], Ryan B. Scott [b,c], Gerhard Jocham [a]

[a] *Biological Psychology of Decision Making, Institute of Experimental Psychology, Heinrich Heine University, 40225 Düsseldorf, Germany*
[b] *School of Psychology, University of Sussex, Brighton BN1 9RH, UK*
[c] *Sussex Centre for Consciousness Science, University of Sussex, Brighton BN1 9RH, UK*

## ABSTRACT

Instrumental conditioning is a crucial part of adaptive behaviour, allowing agents to selectively interact with stimuli in their environment. Recent evidence suggests that instrumental conditioning cannot proceed without stimulus awareness. However, whether accurate unconscious instrumental *responding* can emerge from consciously acquired knowledge of the stimulus-action-outcome contingencies is unknown. We studied this question using instrumental trace conditioning, where participants learned to make approach/avoid decisions in two within-subject modes: conscious (stimuli in plain view) and unconscious (visually masked). Both tasks were followed by an unconscious-only instrumental performance task. We show that even when the contingencies are reliably learned in the conscious mode, participants fail to act upon them in the unconscious responding task. We also replicate the previous results that no instrumental learning occurs in the unconscious mode. Consequently, the absence of stimulus awareness not only precludes instrumental conditioning, but also precludes any kind of instrumental responding to already known stimuli. This suggests that instrumental behaviour is entirely supported by conscious awareness of the world, and corroborates the proposals that consciousness may be necessary for adaptive behaviours requiring selective action.

## 1. Introduction

Instrumental behaviour – execution of appropriate action in order to achieve reward or avoid punishment – is fundamental to flexible, goal-oriented interaction with the world. Control of instrumental behaviour is a complex process, involving not only learning of the associations between stimuli, actions, and their outcomes, but also assigning motivational value to the outcomes to invigorate future behaviour, selective deployment of action, and adaptation in the face of environmental volatility. Mechanistically, this involves integrating information across distinct modalities and long temporal scales, including processing the stimulus and extracting its predictive value (e.g. positive or negative), deploying a selective response (e.g. approach or avoid), and comparing the expected outcome with the actual outcome in order to update the expectations of stimulus values (Balleine, 2011; Balleine and O'Doherty, 2010; Balleine and Ostlund, 2007; Sutton and Barto, 1998).

Recent evidence demonstrates that instrumental learning cannot proceed when the reward- or punishment-predictive stimuli are prevented from entering awareness (Reber et al., 2018; Skora et al., 2021a; Skora et al., 2021b; Skora et al., 2022; Skora et al., 2023). Comparable evidence is emerging for other complex forms of learning, such as fear conditioning (Mertens and Engelhard, 2020) and contingency learning (Travers et al., 2018). In contrast, simpler forms of learning that do not require selective action, such as classical conditioning or simple associative learning (e.g. learning the association between two unrelated stimuli), appear to be feasible even when the stimuli are not consciously perceived (Clark and Squire, 1998; Kim et al., 2015; Knight et al., 2003; Lin and He, 2009; Rosenthal et al., 2016; Scott et al., 2018).

This disparity supports the theoretical perspectives suggesting that processes requiring information integration across time and distinct cognitive modules should require conscious access (Dehaene et al., 2014; Dehaene and Changeux, 2011; Dehaene and Naccache, 2001; Lamme, 2006; Mudrik et al., 2014). Conscious access, under this perspective, reflects complex processing patterns necessary for long-range, long-lasting, feedback and feedforward connections between

distinct brain regions. While low-level or shorter-range information integration, supporting simple forms of learning, may be possible without conscious awareness, increased complexity of processing, such as that observed during problem-solving or decision-making, should require consciousness (Baars, 2002; Melloni et al., 2007; van Gaal et al., 2012). Indeed, rendering stimuli inaccessible to consciousness with visual masking could prevent learning from taking place by disabling information integration across the long-range recurrent network supporting such learning (Skora et al., 2021a; Skora et al., 2021b; Skora et al., 2022, 2023).

Nonetheless, while unconscious instrumental conditioning might not be feasible, it is conceivable that instrumental behaviour can proceed unconsciously once the necessary knowledge of stimulus-action-outcome associations has been acquired consciously. Such an effect has been observed in a multisensory associative learning paradigm, where conscious acquisition of the contingencies proved necessary for accurate performance when stimulus awareness was prohibited (Faivre et al., 2014). Here, we test whether a similar dependency exists for instrumental behaviour. Specifically, we test if instrumental responding to stimuli that are not consciously perceived can occur after conscious learning of the instrumental contingencies. We test this possibility with a two-phase task, conducted in two within-subjects conditions. In the first condition, subjects perform a conscious instrumental conditioning phase, learning to approach a rewarding stimulus and avoid a punishing stimulus (100% deterministic), followed by a performance stage, where they continue to make approach/avoid responses to the same stimuli, now rendered unconscious through forward-backward visual masking. In the second condition, both the learning and the performance stages are unconscious.

Investigating this effect in an instrumental conditioning scenario can shed light on the functional boundaries of consciousness from the perspective of adaptive value. If accurate unconscious instrumental performance is observed following conscious acquisition of the stimulus-action-outcome contingencies, we may conclude that consciousness is vital for acquisition of adaptive behaviours, but not necessarily for their successful execution. A failure to observe successful unconscious performance following conscious conditioning would suggest that instrumental behaviour, both learning and execution, is entirely dependent on conscious access.

## 2. Method

### 2.1. Participants

38 participants (9 males) with a mean age of 22.42 years ($SD = 2.84$, range = 19–31, two participants failed to report their age) were recruited at the [university hidden], through word of mouth and campus advertisement. All had normal or corrected-to-normal vision, no current or history of neurological illness, and participated in exchange for course credit or 6 EUR payment. Ethical approval was granted by the Psychology Ethics committee at the [hidden], and the study was conducted in accordance with the Declaration of Helsinki.

Sample size was determined using the Bayesian Stopping Rule (see *Results-Learning*), whereby data collection continued until a sensitive result was obtained in both conditioning tasks, either in support of the H0 (absence of learning; conventionally indicated by a Bayes Factor smaller than 0.3), expected for the unconscious conditioning phase, or the H1 (evidence of learning; indicated by a Bayes Factor larger than 3), expected for the conscious conditioning phase.

### 2.2. Stimuli and materials

The experiment was ran on Matlab 2019a (MathWorks, 2019) with Psychophysics Toolbox (Brainard, 1997). The stimuli were presented on a 24 in. Asus PG248Q gaming monitor (1920 by 1080 pixels) with a refresh rate of 120 Hz.

The stimuli in the main task included 6 neutral symbols (3 vertically symmetrical, 3 asymmetrical) taken from Agathodaimon font, and two circular shapes in the perceptual discrimination task used for threshold determination. All stimuli were 240 × 240 pixels in size, and presented in light grey (RGB: 217217217) on white background. The stimuli were forward-backward masked with a 240 × 240-pixel black-and-white noise image, generated by randomly re-scrambling the 3 × 3 pixel blocks in the image on the onset of every trial. Low contrast cues and the type of mask were intended to maximise the exposure duration without awareness.

### 2.3. Procedure

Each experimental session was composed of two conditions. Each condition consisted of an initial learning phase, either conscious (C, with stimuli presented in plain view, rendering participants fully aware of them), or unconscious (UC, with the stimuli visually masked and thus presented subliminally, without conscious awareness), followed by a subsequent performance phase (always unconscious, UC). Participants were seated with their chin on a chin rest at a 55 cm distance from the screen, and allowed to choose between English and German as their preferred language of the session. Each session began with the threshold of visual awareness determined individually for each participant using a masked perceptual discrimination task (see *Threshold setting*). The established threshold ($M = 284$ ms, $SD = 108$ ms) was then used as display duration in a brief, 8-trial practice round, and in the first condition (UC learning + UC performance, or C learning + UC performance; conducted in a randomised order). Following the first condition and a self-paced break, participants completed the threshold-finding task again, which was then used to find a display duration during the remaining condition ($M = 239$ ms, $SD = 125$ ms). The second threshold-setting task was applied to counteract the expected visual adaptation to the stimuli over the course of the first condition, especially if C learning, where the stimuli were unmasked and presented in plain view, occurred first.

#### 2.3.1. Threshold setting

Each trial consisted of a fixation cross (500 ms), followed by forward mask (300 ms), a cue (either a symmetrical or an asymmetrical circular shape, starting at 600 ms), and backward mask (300 ms). Next, participants were asked to judge whether the cue was symmetrical or asymmetrical by pressing corresponding arrow-keys, and instructed to guess if they failed to see the stimulus. They were next asked whether they had any confidence in the judgment, or if they were guessing, also using the arrow-keys (following Skora et al., 2021b). They were instructed to indicate 'some confidence' if they had any degree of confidence (even an intuition), and 'total guess' only if they felt they did not see the cue and responded at random. This distinction was applied as any feelings, including hunches or intuitions, were considered conscious experience (Dienes and Seth, 2010). This awareness check, combining an objective discrimination with a subjective confidence report, ensures that conscious experience of the stimulus is directly relevant to the first-order perception of relevant features of this stimulus, which can then guide choices. In other words, the confidence report reflects the conscious status of the relevant perceptual content (Dienes et al., 2010; Dienes and Seth, 2010). Thus, if participants can discriminate the relevant property of the stimulus (here, symmetry) with any degree of confidence, even just an intuition, we classify them as having conscious experience of the stimulus on that trial. Any objective discrimination, regardless if correct or not, made without a corresponding conscious experience reflected in any degree of confidence, would then be classified as an unconscious trial. Every time a correct symmetry response was made with confidence, the presentation duration was reduced by 50 ms on the subsequent trial. Upon reaching a duration of 100 ms, or upon the first guess response, the presentation duration returned to the previous level (+50 ms), and reduced in 8.35 ms (single screen refresh) steps on the

subsequent trials. The reduction continued until the next guess, at which point the presentation duration remained the same if the participant continued guessing. The duration maintained over six consecutive guesses (regardless of the accuracy of responses) was set as the individual threshold of conscious awareness.

### 2.3.2. Conditioning / performance tasks

The conditioning / performance tasks were adapted from a subliminal deterministic instrumental learning scenario used previously (Pessiglione et al., 2008; Skora et al., 2021b). In the task, participants try to learn to approach (Go response) or avoid (NoGo) the presented stimuli, where one stimulus is deterministically associated with a positive outcome if approached, and another with a negative outcome. If the participants approached the stimulus associated with a win, or successfully avoided a stimulus associated with a loss, they were rewarded with 1 token (golden token displayed on the screen as feedback). Conversely, if they approached the stimulus associated with a loss, or avoided a stimulus associated with a win, they were punished with $-1$ token (a red cross over the golden token displayed). Note that in contrast to the previous paradigms using this task, the 'avoid' (NoGo) response was not neutral, and was also leading to stimulus-action-contingent outcomes, thus maximising the chances of learning. Here, the task proceeded in two conditions. In the UC learning + UC performance condition, the learning task consisted of 70 trials of approach/avoid responses to stimuli rendered subliminal using forward-backward masking, with participants' performance tested in a subsequent 70 trials. In the C learning + UC performance condition, the learning task consisted of 70 trials of approach/avoid responses to stimuli presented in plain view, with their performance tested for those stimuli rendered subliminal in the subsequent 70 trials. Each condition (and the practice task) used a different pair of stimuli, chosen from the pool of 6, randomly assigned to be either rewarding or punishing, without replacement, in order to ensure that each condition contained a novel pair. The stimulus-outcome contingencies remained the same between the learning and performance tasks of each condition.

In all unconscious trials (within the UC learning and UC performance phases), each trial consisted of a fixation cross (500 ms), mask (300 ms), target stimulus (duration determined in the perceptual discrimination task), and mask (300 ms), followed by a decision prompt in the form of a question mark, during which participants had two seconds to make a response (Fig.1). While the question mark was on-screen, participants decided between 'approaching' the stimulus by pressing the spacebar (Go) and 'avoiding' the stimulus by refraining from pressing (NoGo). Participants were instructed that they should follow their instincts in making the decisions, as they should not expect to be consciously aware of the stimulus shown.

Following feedback, participants were asked to report whether the stimulus was vertically symmetrical or asymmetrical, and their confidence in that judgment on a binary scale (some confidence or total guess). As in the threshold task, they were instructed to indicate 'some confidence' if they had any degree of confidence in the symmetry of the stimulus (including instincts), and 'total guess' only if they responded randomly. This allowed us to label all correct and confident trials (including correct hunches) as aware, providing a strict criterion of awareness. The responses were made with the arrow keys before proceeding to the next trial. If three correct and confident responses were made in a row (indicating awareness), duration was reduced by another single screen refresh (8.35 ms). Participants were shown examples of the stimuli (from a different sample) prior to beginning, and informed that stimulus symmetry was unrelated to its rewarding or punishing outcome.

In the C learning task, the trial sequence was identical, with the exception of the masks, which were omitted in order to present the stimulus in plain view (Fig. 1.B). Following feedback presentation, participants were also asked to judge the symmetry of the symbol and their confidence in their decision, with the expectation that they should

be able to be largely correct and confident in their judgments. No reduction in display duration occurred following three correct and confident responses.

### 2.4. Analyses and results

#### 2.4.1. Exclusion criteria

In the UC learning/performance tasks, all trials where participants made a correct symmetry judgment with confidence were marked as aware and excluded (17% of all trials), in order to ensure that analyses are conducted only on truly unconscious trials. All participants showed some awareness during the task ($M_{AWARETRIALS}$ = 48, range: 1–198), which resulted in a drop in duration of exposure ($M$ = 40 ms, $M$ = 30 ms for phases 1 and 2, respectively). Of those who showed awareness, 12 participants (31%) who were aware on >25% (52) of all UC trials were excluded from further analysis.

In the C learning task, one participant who was not correct and confident over 25% of time was excluded, on the assumption that they were not paying attention to the task or the symbols when they were clearly visible.

#### 2.4.2. Learning

On average (across all unconscious phases), participants executed more Go responses (56%) than NoGo responses, regardless of stimulus type. To account for this response bias, type I $d'$ (a Signal Detection Theoretic measure of sensitivity to signal versus noise; (Stanislaw and Todorov, 1999) was computed, with Go responses to rewarding cues treated as Hits, and Go responses to punishing cues as False Alarms. The resulting measure of sensitivity can be taken as evidence of successful discrimination between the stimuli – approximating learning - if it is significantly above 0.

Group-level $d'$ scores for both the conscious and unconscious learning tasks were entered into one-tailed one-sample $t$-tests against 0 (no ability to discriminate the stimuli; chance performance). A Bayes Factor (B; Dienes, 2015a, 2016) was computed for the difference, with a half-normal distribution, mean specified as 0, and $d'$ of 1.79 as the SD of the mean (corresponding to the expected effect size should learning take place, obtained in the supraliminal task of Skora et al., 2023)).[1] A robustness region (RR) is reported for each B, giving the range of scales that qualitatively support the same conclusion (i.e. evidence as insensitive, or as supporting H0, or as supporting H1), notated as: $RR_{1/3>B>3}$ [x1, x2], $RR_{B<1/3}$ [x1, x2], and $RR_{B>3}$[x1, x2], where x1 is the smallest SD that gives the same conclusion and x2 is the largest (Dienes, 2019). In line with the Bayesian Stopping Rule (Dienes, 2015b), data collection continued until a sensitive result was found in support of either $H_0$ (absence of learning; by convention indicated by a B smaller than 0.3) or $H_1$ (presence of learning; indicated by a B larger than 3).

In the C learning condition, d' was significantly different from 0 ($M$ = 2.90, $SE$ = 0.28; $t(24)$ = 10.50, $p < 0.001$, Cohen's $d$ = 2.10; $B_{H(0,1.79)}$ = 146,536,301.23, $RR_{B>3}$[0.14, 988][2]; Fig. 2), indicating that participants were able to learn the associations between the stimuli, their actions, and the action-dependent outcomes. In the UC learning condition,

---

[1] Note that the $d'$ value of 1.79 specified as the expected effect size differs from the pre-registered value of 0.7, estimated from Pessiglione et al. (2008). The reason for this deviation is twofold: 1) the learning effect obtained in Pessiglione et al. failed to replicate in independent studies (Skora et al., 2021, 2022, 2023), suggesting that the value of 0.7 as representing unconscious learning is unreliable; 2) the value of 1.79 corresponds to the expected justabove threshold accuracy, where learning was reliable. Robustness regions provided for Bs allow to assess the range for which the same qualitative conclusion holds, independently of the value chosen as the expected effect size. The robustness regions indeed include the initial value of 0.7, showing that the conclusion would not change should the value have been used.

[2] All Bs were computed with a t-distribution, using a calculator available at https://bencepalfi.shinyapps.io/Dienes_BF_calculator/

## A. Unconscious (UC) learning / performance



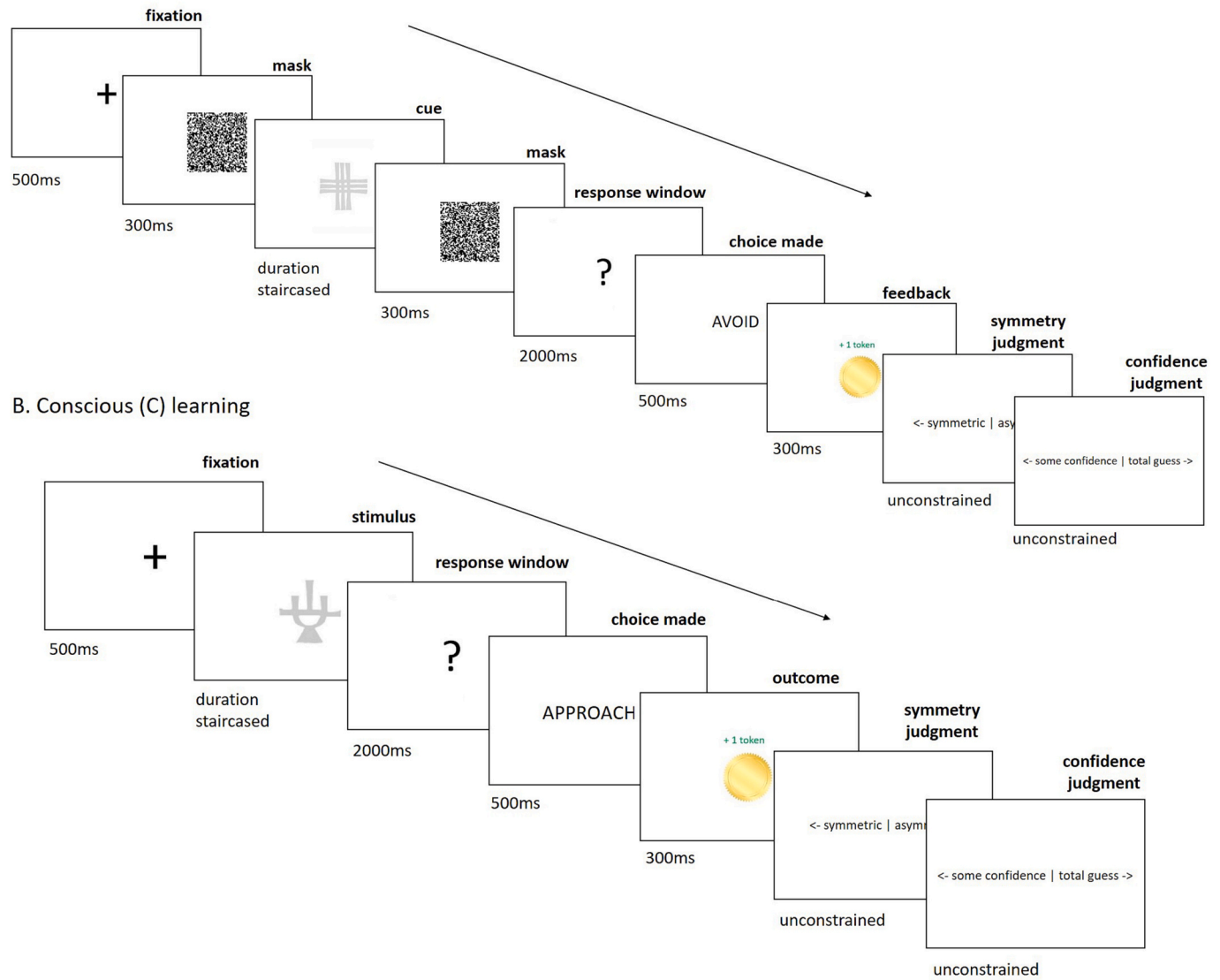## B. Conscious (C) learning

**Fig. 1.** A: Illustration of a single trial in the UC mode of learning and performance (English version). Following a fixation cross, the target stimulus (either rewarding or punishing if approached) was presented between two visual masks. Participants had two seconds to make an approach (Go) or avoid (NoGo) decision, and were rewarded or punished accordingly. In this example, a participant made a correct avoid decision and was rewarded with a golden token. The trial ended with a binary symmetry and confidence judgment. B: Illustration of a single trial in the C mode of learning, where a participant made a correct approach decision. The trial sequence was identical to the UC mode, but the stimuli were presented in plain view, without visual masks.

after excluding aware trials, d' was not significantly different from 0 ($M = 0.03$, $SE = 0.07$; $t(24) = 0.37$, $p = 0.357$, Cohen's $d = 0.07$; $B_{H(0,1.79)} = 0.06$, $RR_{B<1/3}[0.32, \text{Inf}]$; Fig. 2), indicating that participants were not able to learn the associations and, consequently, an absence of unconscious instrumental conditioning.

The magnitude of learning in both conditioning stages (conscious and unconscious) was then compared with a paired *t*-test. Due to the lack of a precise numeric prediction of the expected difference between the conditions, Bayes factors for the paired *t*-tests were computed with a normal distribution, mean specified as 0, and the SD of the distribution specified as half of the maximum expected difference (previously used d' of $1.79/2 = 0.9$; following Dienes, 2019). As expected, C learning gave rise to a significantly higher d' than UC learning ($M_{\text{diff}} = 2.87$, $SE_{\text{diff}} = 0.31$; $t(24) = -9.18$, $p < 0.001$, Cohen's $d = -1.84$; $B_{H(0,0.9)} = 1,062,733.22$, $RR_{B>3}[0.2, 955]$; Fig. 2).

### 2.4.3. Instrumental performance following conscious vs unconscious learning

The statistical approach applied to performance was identical to that applied to the learning blocks. In the UC performance task following C learning, d' was not significantly different from 0 ($M = -0.26$, $SE = 0.27$; $t(22) = -0.94$, $p = 0.822$, Cohen's $d = -0.20$; $B_{H(0,1.79)} = 0.08$, $RR_{B<1/3}[0.45, \text{Inf}]$; Fig. 2). The same was the case in the UC performance task following UC learning, albeit it is noteworthy that accuracy was numerically higher ($M = -0.05$, $SE = 0.16$; $t(24) = -0.29$, $p = 0.611$, Cohen's $d = -0.06$; $B_{H(0,1.79)} = 0.07$, $RR_{B<1/3}[0.40, \text{Inf}]$; Fig. 2), indicating that participants were not able to perform with above-chance accuracy indicative of learning in either condition. There was no difference in performance between the two conditions ($M_{\text{diff}} = 0.21$, $SE_{\text{diff}} = 0.33$; $t(22) = 0.65$, $p = 0.525$, Cohen's $d = 0.14$; $B_{N(0,0.9)} = 0.41$, $RR_{1/3<B<3}[0, 1.3]$), although the B failed to provide strong support for the absence of a difference.

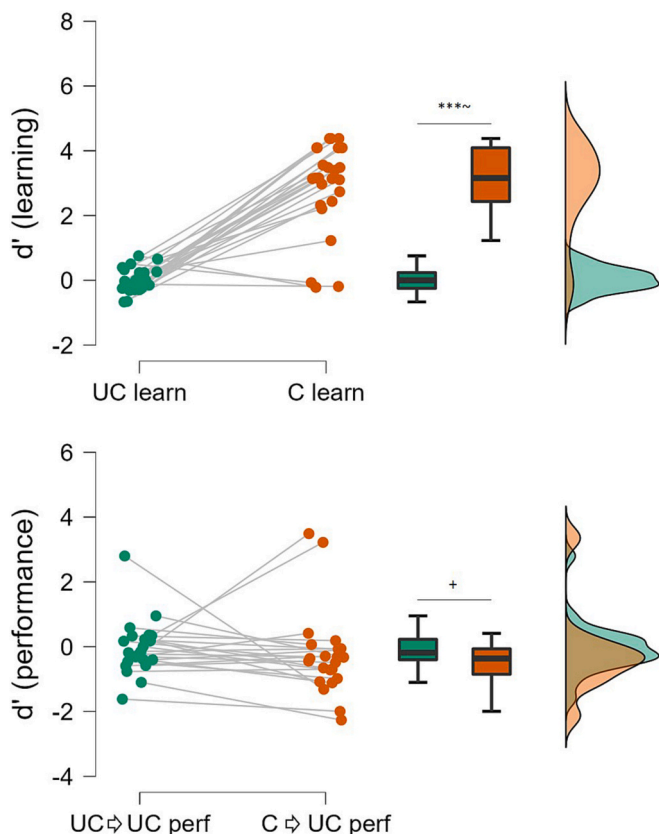Within the conditions, there was a significant decrease in accurate

**Fig. 2.** Top: Type I *d'* for the learning blocks, comparing unconscious instrumental learning (UC learn) with conscious instrumental learning (C learn). Bottom: Type I *d'* for the performance blocks, comparing unconscious instrumental performance following unconscious learning (UC- > UC perf) with unconscious instrumental performance following conscious learning (C- > UC perf). Asterisks indicate significance at: *** = *p* < 0.001. For B: ~ indicates a sensitive B supporting the H1. + indicates a sensitive B supporting H0.

choices (*d'*) from the C learning stage to the corresponding UC performance stage ($M_{diff} = 3.12$, $SE_{diff} = 0.41$; t(22) = 7.52, p < 0.001, Cohen's $d = 1.57$; $B_{N(0,0.9)} = 14,802$, $RR_{B>3}[0.25, 992]$; see Fig. 2). From the UC learning stage to the corresponding UC performance stage, there was no change in accurate choices ($M_{diff} = 0.07$, $SE_{diff} = 0.16$; t(24) = 0.46, $p = 0.651$, Cohen's $d = 0.09$; $B_{N(0,0.9)} = 0.19$, $RR_{B<1/3}[0.6, Inf]$; see Fig. 2).

The same effect held when comparing the *d'* obtained in only the last 20 trials of each stage. There was a significant decrease in accurate choices (*d'*) from the end of the C learning stage ($M = 3.57$) to the end of the UC performance stage ($M = -0.34$; $M_{diff} = 3.23$, $SE_{diff} = 0.43$; t(21) = 8.86, p < 0.001, Cohen's $d = 1.89$; $B_{N(0,0.9)} = 11,064.25$, $RR_{B>3}[0.26, 1100]$). From the end of the UC learning stage ($M = 0.09$) to the end of the corresponding UC performance stage ($M = -0.30$), there was no change in accurate choices ($M_{diff} = 0.39$, $SE_{diff} = 0.25$; t(24) = 1.49, $p = 0.147$, Cohen's $d = 0.30$; $B_{N(0,0.9)} = 0.82$, $RR_{1/3<B<3}[0, 2.8]$). See Fig. 3 for a depiction of proportions of correct choices across the spans of the entire learning and performance stages.

### 2.4.4. Exploratory analysis: Win-stay/ lose-switch behaviour

Given participants' chance-level behaviour in the unconscious blocks, in an exploratory analysis, we computed the proportion of win-stay and lose-switch behaviour (WSLS). WSLS is a heuristic whereby a choice leading to a win is repeated, and a choice resulting in a loss provokes a switch to an alternative choice (Estes, 1950). Such a strategy corresponds to setting the learning rate to one in reinforcement learning algorithms, indicating that choices are only guided by the most recent outcome, rather than by an integrated choice-outcome history. Note that

here, as participants showed no signs of learning, we analysed stay or switch behaviour on the trial immediately following the win or loss outcome (rather than the next trial on which the same stimulus will be presented again). We investigated these performance metrics to obtain exploratory evidence that participants, despite random performance in the absence of knowledge about the stimulus presented to them, were still trying to solve the task by defaulting to a simpler strategy. We calculated win-stay (and lose-shift, respectively) proportions as the percentage of win (loss) trials that were followed by a choice repetition (a choice switch) on the next trial. Above-chance WS and LS values would indicate that participants relied on WSLS as a strategy to guide their decisions in absence of stimulus awareness providing clear stimulus-outcome mappings. One-sample *t*-tests against the chance value of 0.5 were computed for each learning and performance stage (see Table 1). In absence of reasonable expected effect, B was computed for each stage using a Cauchy distribution and a default prior in JASP (JASP Team, 2023). Note that we refrained from computing WSLS values in the conscious learning condition, as participants rapidly attained near-ceiling accuracy, which precludes interpretability of WSLS scores from this block.

Participants showed a tendency to engage in a WS strategy in all three unconscious phases (UC performance following C learning, UC learning and the corresponding UC performance), repeating a rewarded choice on the next trial (see Fig. 4). However, a strong support for this hypothesis from B was only obtained in the UC performance stage following UC learning. In contrast, participants did not engage in the LS strategy more often than expected from chance in any of the learning or performance stages (see Table 1 for full test statistics).

### 2.4.5. Exploratory analysis: reaction times

Reaction times (RT) >2 SD from the mean for each subject or shorter than 100 ms were removed from analysis. Subsequently, $RT_{difference}$ index was computed by subtracting RTs to rewarding cues from RTs to punishing cues. The resulting positive values indicate that participants took a longer time to respond to punishing cues than to rewarding cues, in line with RT-oriented indicators of reward learning (Atas et al., 2014). Zero indicates that there was no difference between the two. The index was entered into a one-sample *t*-test against 0, for learning in both conscious and unconscious conditions. B was computed with a half-normal distribution, mean specified as 0, and a value of 34 ms as the SD of the mean (obtained from a past study which found a RT difference in the absence of performance effects in a comparable task (Atas et al., 2014)). Note that aware trials were excluded, and that trials where no response was made (correctly or incorrectly) yielded no RTs.

In the C learning block, the $RT_{difference}$ was significantly different from 0 ($M = 257.04$ ms, $SE = 60.54$, t(19) = 4.25, p < 0.001, Cohen's $d = 0.95$; $B_{H(0,34)} = 4.58$, $RR_{B>3}[27,30,000]$), showing that participants responded slower to punishing stimuli, as expected. In the UC learning block, the $RT_{difference}$ was not significantly different from 0 ($M = -12.84$ ms, $SE = 21.32$, t(24) = -0.60, $p = 0.553$, Cohen's $d = -0.12$; $B_{H(0,34)} = 0.37$, $RR_{1/3>B>3}[0,45]$).

In the UC performance task following C learning, the $RT_{difference}$ was not significantly different from 0 ($M = -44.05$ ms, $SE = 65.72$, t(20) = -0.67, $p = 0.510$, Cohen's $d = -0.15$; $B_{H(0,34)} = 0.70$, $RR_{1/3>B>3}[0,130]$). Finally, in the UC performance task following UC learning, the index was again not significantly different from 0 ($M = 44.25$ ms, $SE = 23.59$, t(24) = 1.88, $p = 0.073$, Cohen's $d = 0.38$; $B_{H(0,34)} = 1.78$, $RR_{1/3>B>3}[0,430]$). The Bayes factors suggest that the data were insensitive.

## 3. Discussion

In the present study, we investigated whether instrumental responding can proceed unconsciously once the necessary knowledge of stimulus-action-outcome association has been acquired consciously. Subjects performed a two-stimulus, deterministic instrumental trace conditioning task in a conscious and unconscious mode (within-
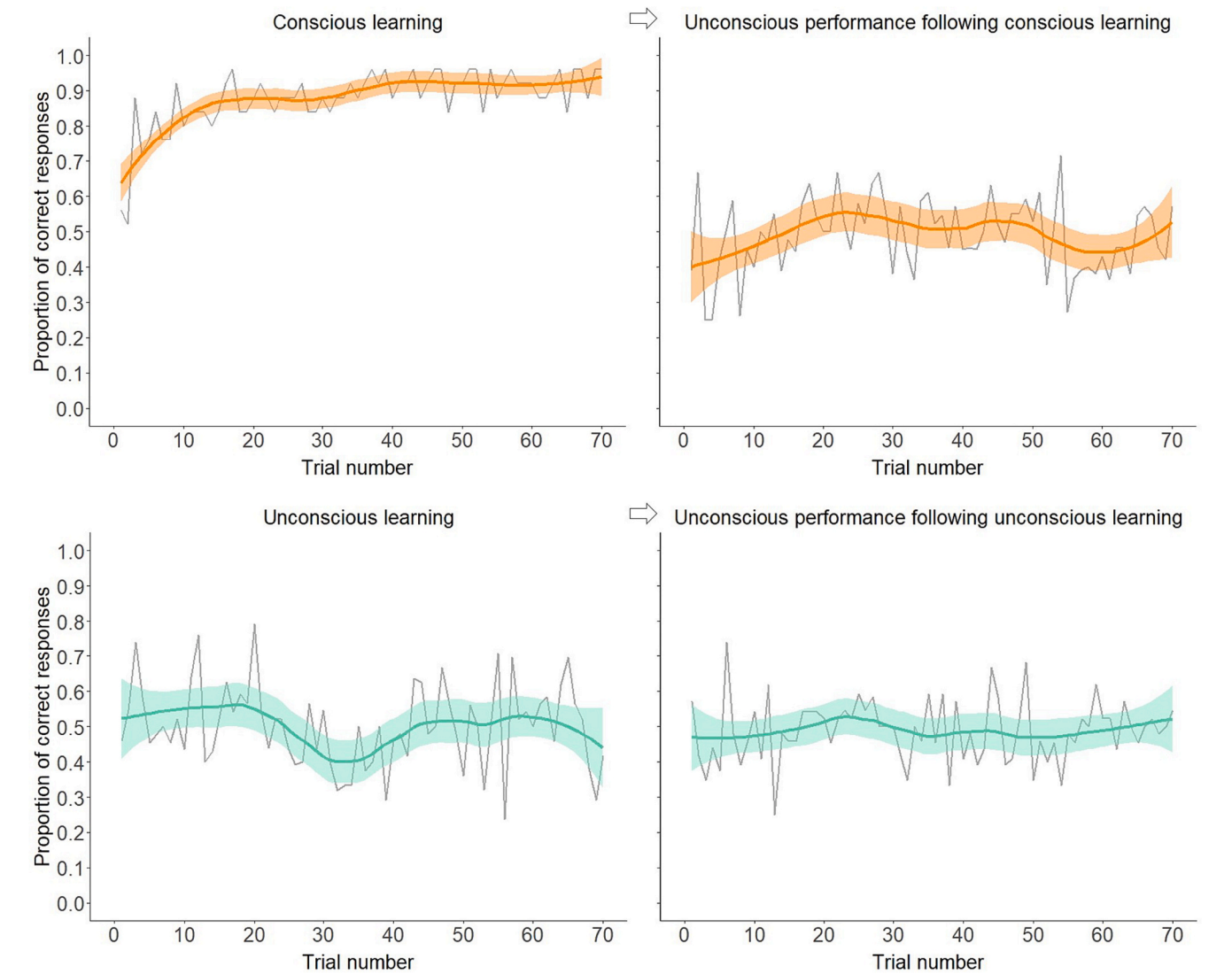
**Fig. 3.** Top: Learning curves for the conscious learning (all trials) - > unconscious performance condition (unaware trials only). Bottom: Learning curves for the unconscious learning - > unconscious performance condition (unaware trials only). Grey lines represent mean (across participants) of correct choices (approach and avoid responses to the rewarding and punishing options, respectively). Coloured lines represent smoothed values (locally weighted regression; ribbon represents +/− 1 SEM).

**Table 1**
WSLS analysis across the learning and performance stages of both conditions (conscious, unconscious learning), representing *t*- and *p*-values obtained from one-sample *t*-tests against 0.5 (chance level), Cohen's *d*, and default-prior Bayes factor B.

| Phase | Win-stay | | | | | Lose-switch | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *M (SE)* | *t*(24) | *p* | *d* | $B_{10}$ | *M (SE)* | *t*(24) | *p* | *d* | $B_{10}$ |
| UC performance following C learning | 0.56 (0.03) | 2.08 | 0.049 * | 0.42 | 1.31 | 0.49 (0.04) | *t*(21) = -0.29 | 0.772 | −0.06 | $0.22^{+}$ |
| UC learning | 0.56 (0.03) | 2.11 | 0.045 * | 0.42 | 1.38 | 0.50 (0.04) | −0.04 | 0.968 | −0.01 | $0.21^{+}$ |
| UC performance following UC learning | 0.60 (0.03) | 2.81 | 0.010 ** | 0.56 | 4.93~ | 0.51 (0.05) | 0.11 | 0.915 | 0.02 | $0.21^{+}$ |

For *p*-values, asterisks represent: *: *p* < 0.05, **: *p* < 0.01. For B, + represents support for H0, and ~ represents support for H1.

subjects), in both cases followed by an unconscious instrumental performance stage using the same stimuli.

Firstly, the results indicate that while subjects were easily able to learn the associations in the conscious mode, there was no evidence of learning in the unconscious mode, reflected in participants' inability to discriminate between the positive and negative stimuli, and approach or avoid accordingly. This replicates previous evidence that both trace and delay instrumental conditioning requires conscious access to the stimuli

(Reber et al., 2018; Skora et al., 2021a, b, 2022, 2023). Secondly, there was no evidence of successful instrumental responding to unconsciously presented cues, even after the associations were reliably learned in the conscious conditioning mode. This demonstrates that it is not only the acquisition of new instrumental associations that requires conscious access – rather, successful instrumental behaviour in general appears to rely on consciousness.

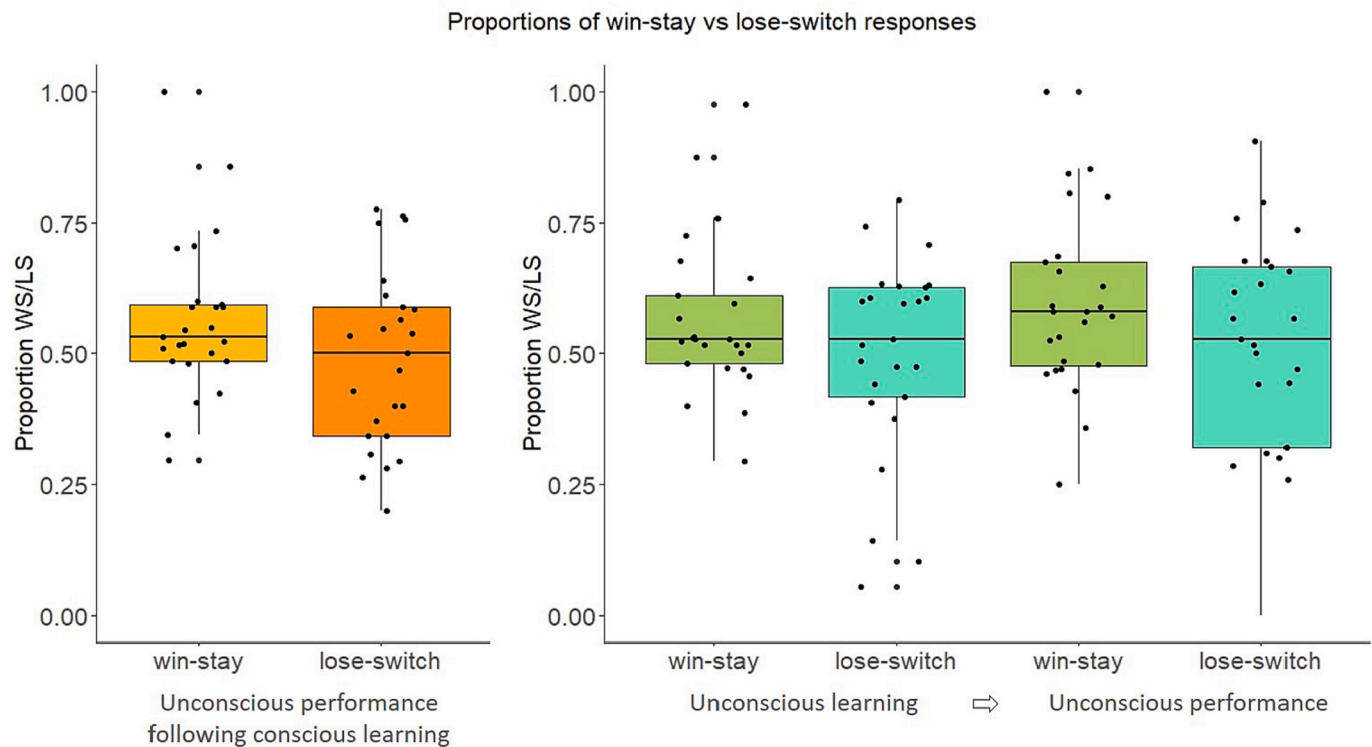Previous work suggested that once information has been integrated

**Fig. 4.** Boxplots presenting the proportions of win-stay/lose-switch responses during the unconscious stages of the task: unconscious performance task following conscious learning (left), and unconscious learning and the corresponding unconscious performance task (right). Conscious learning task is omitted due to poorly computable WSLS proportions caused by participants learning the correct choices quickly with near-ceiling accuracy, thus eliminating the need for relying on WSLS as a decision heuristic (see Fig.3 for learning curves).

consciously, re-activating the relevant processing patterns could proceed as an unconscious process. This has been demonstrated for multisensory integration and simpler kinds of stimulus-stimulus associative learning (Faivre et al., 2014; Mudrik et al., 2014). However, it does not appear to be the case for instrumental associations. Even reliable, conscious knowledge of stimulus-action-outcome associations, trained over the course of 70 trials (35 trials per stimulus), failed to produce above-chance instrumental performance when the same stimuli were blocked from entering awareness with visual masking.

One interpretation of this effect could be that the components of the process are too widely distributed, requiring sequential processing across distinct brain regions, retrieval of the representations of expected stimulus values from memory, and selective action deployment. Processing of this complexity, involving long-range, feedback and feedforward connections, is largely considered to require conscious access (Baars, 2002; Dehaene et al., 2014; Dehaene and Changeux, 2011; van Gaal et al., 2012). A subliminally presented stimulus might not be able to evoke this broad range of activity across regions and time. If this is true, then the unconsciously presented stimulus fails to be integrated with subsequent events, even though it is processed sufficiently in the visual cortex and perhaps beyond (note that similar masking methods produces successful simple associative audio-visual learning; Scott et al., 2018). Possibly, subliminally presented stimuli are not represented in higher-order visual areas that provide input to lateral orbitofrontal regions that have been shown to be crucial for correctly assigning credit for an outcome to the choice that caused it (Jocham et al., 2016; Klein-Flügge et al., 2013; Walton et al., 2010). During learning, even if a correct response is made, the outcome may not be able to be successfully linked with the subliminally presented stimulus and the executed action. Consequently, the stimulus never gets to carry any intrinsic value in order to drive subsequent choices. As such, a failure during this process would explain absence of unconscious instrumental conditioning in the first place.

However, after successful conscious learning, the stimuli should already have acquired action-contingent outcome values. Failure to respond appropriately to an already known but subliminally presented stimulus implies that blocking it from conscious access might prevent it from activating the regions coding for its value. Alternatively, the regions coding for value may be activated, but fail to facilitate selective choices. This scenario, however, is unlikely for a few reasons. Firstly, activating expected value might activate appropriate action purely through automatic, instinctive processes (e.g. through a Pavlovian bias, the tendency to approach rewarding stimuli). Secondly, one would expect to observe indirect markers of stimulus values being activated that are not under volitional control, such as autonomic readouts of performance monitoring or a reaction time difference. Those have not been observed in this context (Skora et al., 2021b; Skora et al., 2022). Nonetheless, neuroimaging and connectivity analyses might be fruitfully applied to investigate the entire process in order to pinpoint where it breaks down.

The result that instrumental performance still requires conscious access even after conscious training invites the question of the need for consciousness in habitual, automatic behaviour. Instrumental behaviour can become habitual (driven by highly automated stimulus-response associations, as opposed to outcome values) when sufficiently deeply ingrained (Dezfouli and Balleine, 2012). It is plausible that extensive conscious instrumental experience with a stimulus could reduce or eliminate the need for conscious access through automatic mapping of stimulus to responses. Still, instrumental behaviour is characterised by high flexibility and selectiveness. As such, it is likely that it should require conscious access in order to allow for rapid and flexible responding or behavioural adjustment if the contingencies change. Once instrumental behaviour becomes habitual, the flexibility is diminished. It is conceivable that consciousness facilitates the flexibility of instrumental behaviour, which becomes rigid and resistant to adjustment without conscious access.

The analysis of WSLS strategy suggests that participants partially relied on decision heuristics to guide their decisions in the absence of stimulus awareness providing clear stimulus-action-outcome mappings. In the unconscious learning and performance tasks, participants had a tendency to repeat rewarded choices (although this result achieved strong support only in the unconscious performance following unconscious learning task). However, there was no evidence for a LS strategy being adopted in any of the unconscious tasks. Interestingly, the unconscious performance task following unconscious learning showed a clear and sensitive tendency for WS – perhaps due to having learned that there is no clear strategy to follow over the course of the task besides WS. This, however, remains a speculation.

Pertaining to reaction times, we were not able to find sensitive evidence against the absence of a difference in response speed to rewarding versus punishing stimuli (see also Skora et al., 2021b). This likely stemmed from an insufficient number of trials (notably, as a feature of the design, avoid responses do not yield any RT). This outcome differs from the previously reported result, where participants were faster to respond to rewarding stimuli than to punishing stimuli in absence of any difference in choices, suggesting that some unconscious knowledge may have been acquired (Atas et al., 2014). However, the difference in methods between the Atas et al. study and ours may account for the discrepancy. For example, conscious access is considered to be graded and different across masking methods (Breitmeyer, 2015; Kouider and Dehaene, 2007; Windey et al., 2014), with crowding (as used by Atas et al.) potentially permitting more conscious access. Importantly, Atas et al. did not control awareness in a trial-wise fashion.

Interestingly, previous research has shown that subliminal (masked) No-Go cues can activate prefrontal regions associated with inhibitory control and trigger behavioural inhibition, resulting in slower RTs to No-Go cues or even complete abortion of a response (Lau and Passingham, 2007; van Gaal et al., 2008; van Gaal et al., 2010). Those results imply that in the performance stage (after successful learning), the subliminal punishing stimuli should be able to trigger some degree of inhibition. This was not observed in our task. One possible reason for this difference could be that the above papers interleaved the subliminal cues with supraliminal cues, potentially priming the subliminal trials with knowledge from the supraliminal trials. Another possible reason stems from the type of behaviour. In the above studies, participants' default action (Go) is meant to be stopped by a No-Go signal (whether masked or not). In our task, each trial requires an active choice between approach and avoid, rather than inhibition of a default response. This key difference might engage different mechanisms: inhibitory control versus active choice. It is possible that inhibitory control of default actions can be triggered without awareness of the cue, but once selective action is needed, cue awareness is necessary.

This evidence, and the above considerations, have implications for understanding the function of consciousness, supporting the notions that adaptive behaviours, where selective and flexible decision-making is involved, require consciousness for successful operation. Complex forms of learning (including trace conditioning, second-order conditioning, or flexible re-learning) have been considered to share overlapping markers with the hallmarks of consciousness, suggesting that the two may be evolutionarily intertwined (Birch et al., 2020; Birch et al., 2021; Ginsburg and Jablonka, 2019). Elsewhere, consciousness has been closely tied to action, providing a frame of reference for individuals' interactions with the world (Clark, 2016; Land, 2012; Merker, 2005; Seth et al., 2016). Consequently, consciousness may enable complex, flexible and longer-term decision-making strategies, going beyond the simple and rigid stimulus-stimulus or stimulus-response associations. Indeed, such forms of learning have been demonstrated without conscious access (Bayley et al., 2005; Clark and Squire, 1998; Knight et al., 2003; Lin and He, 2009; Rosenthal et al., 2016; Scott et al., 2018; Seitz et al., 2009)

Nonetheless, it is noteworthy that d' is a limited means of assessing learning, averaging across the entire block and discounting potential improvements in the learning process. However, inspecting the learning curves suggests that participants failed to improve their responses over the course of the unconscious blocks (learning and performance), in contrast to the conscious learning condition. Secondly, while the trial-by-trial awareness check constitutes an immediate, relevant, and sensitive measure of awareness (Berry and Dienes, 1993; Newell and Shanks, 2013), there is still a possibility that the conscious (albeit excluded) trials affected behaviour on the unconscious trials. We took steps to guard against this possibility by excluding participants with a large number of conscious trials (over 25%). For those with a smaller number of conscious trials, the potential effect of conscious knowledge is negligible (Skora et al., 2023).

Finally, it could be argued that symmetry, used as a first-order property of the stimulus in the objective part of the awareness check, is too high-level of a feature to adequately capture any partial perception of the stimulus. The motivation for choosing symmetry and the answer to this concern lie in the awareness check. Symmetry was the key characteristic predictive of the nature of the stimuli (in each block, one stimulus was always symmetrical and the other asymmetrical). Hence, we assessed participants' confidence reflecting the conscious status of symmetry as the relevant perceptual content – if they could perceive the stimulus sufficiently to determine if it was symmetrical or not with any confidence, they should also be able to approach or avoid it correctly. This is, of course, imperfect, as it may fail to capture partial perception based on another telling property of the stimuli. However, if such uncaptured partial perception was driving choices, it could have only elevated the d'. Since we failed to obtain evidence of learning, we deem the likelihood of uncaptured partial perception unlikely.

To conclude, the present study investigated whether instrumental responding can proceed unconsciously once the necessary knowledge of stimulus-action-outcome associations has been acquired consciously. We demonstrate that even when the instrumental contingencies are reliably learned consciously, participants fail to successfully respond to the same stimuli when they are not consciously perceived. We also replicate the previous evidence that no instrumental conditioning occurs in the unconscious mode. Consequently, the absence of conscious awareness of the stimuli in the environment not only precludes instrumental conditioning, but also precludes any kind of instrumental responding to already known stimuli. This suggests that instrumental behaviour is entirely supported by conscious awareness of the world, and corroborates the proposals that consciousness may be necessary for adaptive behaviours requiring selective action.

**Open practices statement**

The study was preregistered at https://osf.io/d3gwt. The Matlab code, stimuli, and data are publically accessible at https://osf.io/er8f7/.

**CRediT authorship contribution statement**

**Lina I. Skora:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing. **Ryan B. Scott:** Conceptualization, Methodology, Writing – review & editing. **Gerhard Jocham:** Formal analysis, Funding acquisition, Resources, Writing – review & editing.

**Data availability**

Link to data in manuscript.

# References

Atas, A., Faivre, N., Timmermans, B., Cleeremans, A., & Kouider, S. (2014). Nonconscious learning from crowded sequences. *Psychological Science, 25*(1), 113–119. https://doi.org/10.1177/0956797613499591

Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences, 6*(1), 47–52. https://doi.org/10.1016/S1364-6613(00)01819-2

Balleine, B. W. (2011). Sensation, incentive learning, and the motivational control of goal-directed action. In J. A. Gottfried (Ed.), *Neurobiology of sensation and reward (pp. 287–310)*. CRC Press/Taylor & Francis. http://oreilly.com/catalog/errata.csp?isbn=9781449340377.

Balleine, B. W., & O'Doherty, J. P. (2010). Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology Reviews, 35*, 48–69. https://doi.org/10.1038/npp.2009.131

Balleine, B. W., & Ostlund, S. B. (2007). Still at the choice-point: Action selection and initiation in instrumental conditioning. *Annals of the New York Academy of Sciences, 1104*, 147–171. https://doi.org/10.1196/annals.1390.006

Bayley, P. J., Frascino, J. C., & Squire, L. R. (2005). Robust habit learning in the absence of awareness and independent of the medial temporal lobe. In *, Vol. 436, issue 7050. Nature* (pp. 550–553). www.nature.com/nature.

Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. Lawrence Erlbaum.

Birch, J., Ginsburg, S., Jablonka, · Eva.. (2021). The learning-consciousness connection. *Biology and Philosophy, 36*, 36–49. https://doi.org/10.1007/s10539-021-09802-5

Birch, J., Ginsburg, S., & Jablonka, E. (2020). Unlimited associative learning and the origins of consciousness: A primer and some predictions. *Biology and Philosophy, 35* (6), 56. https://doi.org/10.1007/s10539-020-09772-0

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Breitmeyer, B. G. (2015). Psychophysical '"blinding"' methods reveal a functional hierarchy of unconscious visual processing q. *Consciousness and Cognition, 35*, 234–250. https://doi.org/10.1016/j.concog.2015.01.012

Clark, A. (2016). *Surfing uncertainty: Prediction, action and the embodied mind*. Oxford University Press.

Clark, R. E., & Squire, L. R. (1998). Classical conditioning and brain systems: The role of awareness. *Science, 280*, 77–81.

Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70*(2), 200–227. https://doi.org/10.1016/J.NEURON.2011.03.018

Dehaene, S., Charles, L., King, J.-R., & Marti, S. (2014). Toward a computational theory of conscious processing. *Current Opinion in Neurobiology, 25*, 76–84. https://doi.org/10.1016/j.conb.2013.12.005

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79*, 1–37.

Dezfouli, A., & Balleine, B. W. (2012). Habits, action sequences and reinforcement learning. *European Journal of Neuroscience, 35*(7), 1036–1051. https://doi.org/10.1111/j.1460-9568.2012.08050.x

Dienes, Z. (2015a). How Bayesian statistics are needed to determine whether mental states are unconscious. In *Behavioral methods in consciousness research* (pp. 199–220). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199688890.003.0012.

Dienes, Z. (2015b). How many participants might I need? - YouTube. https://www.youtube.com/watch?v=10Lsm_o_GRg.

Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology, 72*, 78–89. https://doi.org/10.1016/j.jmp.2015.10.003

Dienes, Z. (2019). How do I know what my theory predicts? *Advances in Methods and Practices in Psychological Science, 1–14*. https://doi.org/10.1177/2515245919876960

Dienes, Z., Scott, R. B., & Seth, A. K. (2010). Subjective measures of implicit knowledge that go beyond confidence: Reply to Overgaard et al. Consciousness and Cognition, 19(2), 685–686. doi: https://doi.org/10.1016/j.concog.2010.01.010.

Dienes, Z., & Seth, A. K. (2010). Measuring any conscious content versus measuring the relevant conscious content: Comment on Sandberg et al. *Consciousness and Cognition, 19*(19), 1079–1080. https://doi.org/10.1016/j.concog.2010.03.009

Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review, 57*(2), 94–107. https://doi.org/10.1037/h0058559

Faivre, N., Mudrik, L., Schwartz, N., & Koch, C. (2014). Multisensory integration in complete unawareness: Evidence from audiovisual congruency priming. *Psychological Science, 25*(11). https://doi.org/10.1177/0956797614547916

van Gaal, S., de Lange, F. P., & Cohen, M. X. (2012). The role of consciousness in cognitive control and decision making. *Frontiers in Human Neuroscience, 6*(121), 1. https://doi.org/10.3389/fnhum.2012.00121

van Gaal, S., Ridderinkhof, K. R., Fahrenfort, J. J., Scholte, H. S., & Lamme, V. A. F. (2008). Frontal cortex mediates unconsciously triggered inhibitory control. *Journal of Neuroscience, 28*(32), 8053–8062. https://doi.org/10.1523/JNEUROSCI.1278-08.2008

van Gaal, S., Ridderinkhof, K. R., Scholte, H. S., & Lamme, V. A. F. (2010). Unconscious activation of the prefrontal no-go network. *Journal of Neuroscience, 30*(11), 4143–4150. https://doi.org/10.1523/JNEUROSCI.2992-09.2010

Ginsburg, S., & Jablonka, E. (2019). *Evolution of the sensitive soul*. MIT Press.

Jocham, G., Brodersen, K. H. H., Constantinescu, A. O. O., Kahn, M. C. C., Ianni, A. M., Walton, M. E. E., … Behrens, T. E. E. J. (2016). Reward-guided learning with and without causal attribution. *Neuron, 90*(1), 177–190. https://doi.org/10.1016/J.NEURON.2016.02.018

Kim, D., Seitz, A. R., & Watanabe, T. (2015). Visual perceptual learning by operant conditioning training follows rules of contingency. *Visual Cognition, 23*, 147–160. https://doi.org/10.1080/13506285.2015.1015663

Klein-Flügge, M. C., Barron, H. C., Henning Brodersen, K., Dolan, R. J., & Behrens, T. E. J. (2013). *Segregated Encoding of Reward-Identity and Stimulus-Reward Associations in Human Orbitofrontal Cortex*. https://doi.org/10.1523/JNEUROSCI.2532-12.2013

Knight, D. C., Nguyen, H. T., & Bandettini, P. A. (2003). Expression of conditional fear with and without awareness. *Proceedings of the National Academy of Sciences, 100*, 15270–15283.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 362*(1481), 857–875. https://doi.org/10.1098/rstb.2007.2093

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences, 10*(11), 494–501. https://doi.org/10.1016/j.tics.2006.09.001

Land, M. F. (2012). Minireview the operation of the visual system in relation to action. *Current Biology, 22*, 811–817. https://doi.org/10.1016/j.cub.2012.06.049

Lau, H. C., & Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 27*(21), 5805–5811. https://doi.org/10.1523/JNEUROSCI.4335-06.2007

Lin, Z., & He, S. (2009). Seeing the invisible: The scope and limits of unconscious processing in binocular rivalry. *Progress in Neurobiology, 87*(4), 195–211. https://doi.org/10.1016/J.PNEUROBIO.2008.09.002

Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., & Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *Journal of Neuroscience, 27*(11), 2858–2865. https://doi.org/10.1523/JNEUROSCI.4623-06.2007

Merker, B. (2005). The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution. *Consciousness and Cognition, 14*(1), 89–114. https://doi.org/10.1016/S1053-8100(03)00002-3

Mertens, G., & Engelhard, I. M. (2020). A systematic review and meta-analysis of the evidence for unaware fear conditioning. *Neuroscience & Biobehavioral Reviews, 108*, 254–268. https://doi.org/10.1016/j.neubiorev.2019.11.012

Mudrik, L., Faivre, N., & Koch, C. (2014). Information integration without awareness. *Trends in Cognitive Sciences, 18*, 488–496. https://doi.org/10.1016/j.tics.2014.04.009

Newell, B. R., & Shanks, D. R. (2013). Unconscious influences on decision making: A critical review. *Behavioural and Brain Sciences, 37*(1), 1–19. https://doi.org/10.1111/j.1365-2141.1970.tb01443.x

Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., & Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron, 59*, 561–567. https://doi.org/10.1016/j.neuron.2008.07.005

Reber, T. P., Samimizad, B., & Mormann, F. (2018). Cue discriminability predicts instrumental conditioning. *Consciousness and Cognition, 61*, 49–60. https://doi.org/10.1016/J.CONCOG.2018.03.013

Rosenthal, C. R., Andrews, S. K., Antoniades, C. A., Kennard, C., Soto Correspondence, D., & Soto, D. (2016). Learning and recognition of a non-conscious sequence of events in human primary visual cortex. *Current Biology, 26*, 834–841. https://doi.org/10.1016/j.cub.2016.01.040

Scott, R. B., Samaha, J., Chrisley, R., & Dienes, Z. (2018). Prevailing theories of consciousness are challenged by novel cross-modal associations acquired between subliminal stimuli. *Cognition, 175*, 169–185. https://doi.org/10.1111/j.1538-7836.2011.04331.x

Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. *Neuron, 61*, 700–707. https://doi.org/10.1016/j.neuron.2009.01.016

Seth, A. K., Verschure, P. F. M. J., Blanke, O., Butz, M. V., Ford, J. M., Frith, C. D., … Regan, J. K. O. (2016). Action-oriented understanding of consciousness and the structure of experience. In A. K. Engel, K. J. Friston, & D. Kragic (Eds.), *The pragmatic turn: Toward action-oriented views in cognitive science* (pp. 261–281). MIT Press.

Skora, L. I., Livermore, J. J. A., Dienes, Z., Seth, A. K., & Scott, R. B. (2023). Feasibility of unconscious instrumental conditioning: A registered replication. *Cortex, 159*, 101–117. https://doi.org/10.1016/J.CORTEX.2022.12.003

Skora, L. I., Livermore, J. J. A., Nisini, F., & Scott, R. B. (2022). Awareness is required for autonomic performance monitoring in instrumental learning: Evidence from cardiac activity. *Psychophysiology, e14047*. https://doi.org/10.1111/PSYP.14047

Skora, L. I., Seth, A. K., & Scott, R. B. (2021a). Sensorimotor predictions shape reported conscious visual experience in a breaking continuous flash suppression task. *Neuroscience of Consciousness, 7*, 1–13.

Skora, L. I., Yeomans, M. R., Crombag, H. S., & Scott, R. B. (2021b). Evidence that instrumental conditioning requires conscious awareness in humans. *Cognition, 208*, Article 104546. https://doi.org/10.1016/j.cognition.2020.104546

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31*(1), 137–149.

Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning*. MIT Press.

Travers, E., Frith, C. D., & Shea, N. (2018). Learning rapidly about the relevance of visual cues requires conscious awareness. *Quarterly Journal of Experimental Psychology, 71* (8), 1698–1713. https://doi.org/10.1080/17470218.2017.1373834

Walton, M. E. E., Behrens, T. E. J., Buckley, M. J., Rudebeck, P. H., & Rushworth, M. F. S. (2010). Separable learning Systems in the Macaque Brain and the role of orbitofrontal cortex in contingent learning. *Neuron, 65*.

Windey, B., Vermeiren, A., Atas, A., & Cleeremans, A. (2014). The graded and dichotomous nature of visual awareness. *Philosophical Transactions of the Royal Society B, 369*, 1–11. https://doi.org/10.1098/rstb.2013.0282