Evidence that instrumental conditioning requires conscious awareness in humans

Article (Accepted Version)

Skora, L I, Yeomans, M R, Crom, H S and Scott, R B (2021) Evidence that instrumental conditioning requires conscious awareness in humans. *Cognition*, 208. a104546 1-13. ISSN 0010-0277

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/95578/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Evidence that instrumental conditioning requires conscious awareness in humans.

Lina I. Skora ^{a,b}

Ryan B Scott. ^{a,b}

^a School of Psychology, University of Sussex, Brighton, UK, BN1 9RH

^b Sackler Centre for Consciousness Science, University of Sussex, Brighton, UK, BN1 9RH

Declarations of interest: none reported.

Corresponding author:

Lina Skora, School of Psychology, Pevensey 1 Building, University of Sussex, Falmer, BN1 9QH.

Email: L.Skora@sussex.ac.uk

Abstract

Instrumental conditioning is a crucial substrate of adaptive behaviour, allowing individuals to selectively interact with the stimuli in their environment to maximise benefit and minimise harm. The extent to which complex forms of learning, such as instrumental conditioning, are possible without conscious awareness is a topic of considerable importance and ongoing debate. In light of recent theoretical and empirical contributions casting doubt on the early demonstrations of unconscious instrumental conditioning, we revisit the question of its feasibility in two modes of conditioning. In Experiment 1, we used trace conditioning, following a prominent paradigm (Pessiglione et al., 2008) and enhancing its sensitivity. Success in this task requires participants to learn to approach reward-predictive stimuli and avoid punishment-predictive stimuli through monetary reinforcement. All stimuli were rendered unconscious using forward-backward masking. In Experiment 2, we used delay conditioning to shorten the stimulus-outcome delay, retaining the structure of the original task but presenting the stimuli under continuous flash suppression to allow for an overlap of the stimulus, action, and outcome, as well as replacing monetary reinforcement with primary appetitive reinforcement. In both experiments, we found evidence for absence of unconscious instrumental conditioning, showing that participants were unable to learn to adjust their behaviour to approach positive stimuli and avoid negative ones. This result is consistent with evidence that unconscious stimuli fail to bring about long-term behavioural adaptations, and provides empirical evidence to support theoretical proposals that consciousness might be necessary for adaptive behaviour, where selective action is required.

Keywords: unconscious instrumental conditioning, operant conditioning, unconscious learning, adaptive behaviour, consciousness

1. Introduction

The extent to which learning can proceed without conscious awareness has been an extensively studied question. While there have been successful demonstrations that simple associations can be learned without conscious awareness of the stimuli, even between different modalities (e.g. Scott, Samaha, Chrisley, & Dienes, 2018), the possibility of successful unconscious learning at increasing levels of complexity is unclear. One example of a complex learning process which has previously been suggested to proceed without cue awareness is instrumental conditioning, where agents learn the contingencies between stimuli in the world and their outcomes, and to selectively act upon them (e.g. approach/avoid; Mastropasqua & Turatto, 2015; Palminteri, Clair, Mallet, & Pessiglione, 2012; Palminteri et al., 2009; Pessiglione et al., 2008). However, emerging evidence suggests that more complex forms of learning, including instrumental conditioning (Reber, Samimizad, & Mormann, 2018), contingency learning (Travers, Frith, & Shea, 2018), and fear conditioning (Mertens & Engelhard, 2020) may not be possible unconsciously, as originally thought. Adding fuel to the debate are developments in experimental and statistical methods used to reliably assess absence of awareness (Dienes, 2015a; Rothkirch & Hesselmann, 2017; Shanks, 2017). With this in mind, we revisit the topic of unconscious instrumental conditioning. Instrumental learning is critical to adaptive behaviour (both ontogenetically and phylogenetically). As such, the extent to which it is feasible without conscious awareness is a question of considerable importance. In what follows, we attempt to conceptually replicate the original findings of Pessiglione *et al* (2008), employing more sensitive methods and a more robust statistical approach, in two different paradigms: a trace conditioning task using monetary reinforcements, following the original task (Experiment 1), and a delay conditioning task with appetitive and aversive primary reinforcements (Experiment 2).

Past research into unconscious learning suggests that simple forms of associative learning, ranging from classical conditioning, to emotional, visuospatial, or multisensory learning can proceed without conscious awareness of the stimuli (R. E. Clark & Squire, 1998; Faivre, Mudrik, Schwartz, & Koch, 2014; Knight, Nguyen, & Bandettini, 2003; Lin & He, 2009; Olsson & Phelps, 2004; Rosenthal et al., 2016; Rosenthal, Kennard, & Soto, 2010; Scott et al., 2018; Seitz, Kim, & Watanabe, 2009). Learning is usually indexed with presence or absence of the conditioned response, or with reaction times (e.g. where shorter reaction times to congruent vs incongruent pairings are indicative of learning). Unaware learning (as well as priming) is typically observed when stimuli are presented at short spatiotemporal intervals, or concurrently (van Gaal, de Lange, & Cohen, 2012). For example, unaware classical conditioning was demonstrated in delay scenarios (where stimuli to be associated overlap temporally), but not in trace scenarios (where they are separated by an interval; Clark & Squire, 1998).

In contrast, instrumental conditioning can be classified as a more complex process. Agents must not only learn the associations between different stimuli, or stimuli and their outcomes, but also deploy action selectively (e.g. approach or avoid), and adapt their behaviour in the long-term. This involves integrating information over a long temporal scale and distinct modalities, as involved in processing the visual input, extracting its predictive value, deploying a selective response, processing the reinforcement, and comparing the expected outcome with the actual outcome in order to update the representations of stimulus value. As such, it is a considerably more complex process than the aforementioned simpler forms of associative learning, which do not require selective decisions on whether to act or not, or a behavioural adaptation from trial to trial.

Past theoretical and empirical accounts of conscious versus unconscious processing suggest that such an increased level of complexity should require conscious access. A number of theoretical contributions propose that consciousness is related to long-lasting, long-range connections between distinct brain regions, supporting recurrent information integration across distinct cognitive modules (Dehaene & Naccache, 2001; Dehaene & Changeux, 2011; Dehaene, Charles, King, & Marti, 2014; Lamme, 2006; Mudrik, Faivre, & Koch, 2014). As such, low-level or short-range (spatial or temporal) information integration might be possible without conscious awareness, but consciousness might be necessary at increased levels of complexity – for example, semantic knowledge, complex visual processing, decision-making and problem-solving (Baars, 2002), all of which involve integration of information across longer spatiotemporal intervals or larger spatial distance. Indeed, neuroimaging evidence has found conscious processing to be characterised by global, long-range spread of activity, in contrast to more localised, shorter-range projections when processing is unconscious (Baars, 2002; Baars, Ramsøy, & Laureys, 2003; Dehaene et al., 2014, 2001; Dehaene & Naccache, 2001; Melloni et al., 2007).

Another important consideration is related to the methodology of assessing absence of consciousness during task performance. Although it is a frequent practice in unconscious processing research to infer absence of awareness when a behavioural measure (e.g. conditioning, priming, etc.) is above chance, while an independent measure of awareness is not significantly different from chance (e.g. a separate unconscious discrimination task), this approach has drawn criticism (Dienes, 2015a; Vadillo, Konstantinidis, & Shanks, 2016). Finding that performance on a separate awareness check does not differ from chance is not enough to infer true absence of awareness – a non-significant result cannot disambiguate between absence of an effect (i.e. support for the null hypothesis) and absence of evidence for an effect (i.e. insensitive data). In a meta-analysis, Vadillo, Konstantinidis and Shanks (2016) argue that the seemingly chance-performance on this type of awareness test is more likely to reflect a false negative, especially in low-powered studies with a small sample. This approach, also

adopted in the Pessiglione *et al* study, can be rectified in two ways. One is to ensure that the methods to assess awareness are relevant and sensitive (Berry & Dienes, 1993; Shanks, 2017; Shanks & St. John, 1994). This could be achieved by, for instance, a closer similarity between the awareness test and the measure of interest. The second is to apply statistical methods, such as the Bayes factor, which allows to determine whether a null result indicates support for the null (e.g. absence of awareness) over the alternative hypothesis (presence of awareness), or whether the data are insensitive (Dienes, 2014, 2016; Sand & Nilsson, 2016). Both of those approaches will be used in the present paper to ensure true absence of awareness.

Instrumental conditioning is a crucial substrate of adaptive behaviour. Understanding whether it is achievable without conscious awareness is therefore of vital importance for theoretical and practical research investigating the boundaries of unconscious processing. With recent research casting doubt on feasibility of more complex forms of unconscious learning, as well as the methodological and statistical artifacts which can contribute to falsely assuming absence of awareness, it is imperative to rigorously address the question of feasibility of unconscious instrumental conditioning. Here, we present two experimental attempts to conceptually replicate Pessiglione and colleagues' unconscious instrumental conditioning task (also employed in Mastropasqua & Turatto, 2015; Palminteri et al., 2008, and Palminteri et al., 2012). The original task constituted trace conditioning – participants learned the association between the subliminally presented stimulus and its outcome (presented supraliminally as monetary reinforcement), with an intervening gap of 3 seconds, during which action could have been executed (approach) or not (avoid). Experiment 1 is closely based on this paradigm, similarly adopting trace conditioning, and introducing minor methodological changes to enhance the sensitivity of the task. Experiment 2 seeks to evaluate whether unconscious instrumental conditioning can be achieved under conditions previously found to be more favourable for unconscious learning, namely where there is no temporal delay between the stimulus and outcome. Hence, Experiment 2 is a delay instrumental conditioning task, retaining the general structure of the original task, but ensuring that stimulus presentation overlaps with both the response, and the delivery of the reinforcement. To allow for the overlap of all three (which is not possible in just the visual domain), we replaced the secondary visual (monetary) reinforcement with primary reinforcement (appetitive and aversive tastes), which have been previously shown to be highly effective in achieving learning in humans (Birbaumer, Klucken, Angrilli, Andreatta, & Pauli, 2015; Martin-Soelch, Linthicum, & Ernst, 2007; Parkinson, Cardinal, & Everitt, 2000). The strong spatiotemporal overlap was designed in order to maximise the chance of observing conditioning. Notably, primary reinforcements might also offer a stronger incentive to learn. In both experiments, we used Bayes Factors (including robustness regions) to determine whether the results sensitively

support the null hypothesis (i.e. absence of learning) or the alternative hypothesis (presence of learning).

2. Experiment 1

Experiment 1 sought to conceptually replicate the unconscious instrumental conditioning task (Pessiglione et al., 2008), where participants learned to associate predictive cues with reward and punishment, and to adjust their behaviour accordingly by responding (Go) only to rewarding cues, and refraining from a response (NoGo) to punishing cues. Here, we retained the structure of the task, and introduced a number of changes to increase the sensitivity of the paradigm. Firstly, we used a staircased perceptual discrimination task, allowing to determine the sub-threshold stimulus presentation duration for each participant individually (in contrast to keeping the duration at 33 or 50ms, as in Pessiglione *et al.*). We exploited modern screen technologies with higher refresh rates, allowing to present stimuli closer to participants' actual thresholds. Secondly, stimuli were presented in lower contrast, in order to maximise the potential exposure durations without being consciously perceived (following Scott et al., 2018). Thirdly, we disposed of the neutral cue, which was not predictive of any outcome and used in the original experiment primarily for the purpose of comparisons in functional neuroimaging.

Finally, we used a trial-by-trial awareness check, allowing to monitor the levels of awareness on an ongoing basis, and analyse only the trials which were genuinely outside of conscious awareness. In the original Pessiglione *et al.* study, awareness was assessed with a separate measure, and participants' performance on the task of interest is considered unaware if they perform at chance in the separate awareness check (this has also been a standard approach in the field). However, this does not account for the possibility of transient, undetected moments of awareness during the performance itself contributing to unconscious knowledge. As such, a trial-by-trial check allows us to measure awareness in a more immediate fashion (Berry & Dienes, 1993; Newell & Shanks, 2013). Because our staircase should sensitively settle on sub-threshold conditions for each participant, the number of conscious trials to be excluded should be small. While excluding individual aware trials (or participants) can lead to regression to the mean (Shanks, 2017), our analyses have demonstrated its effect to be negligible if only a small proportion of trials or participants are excluded. Here, we will adopt a threshold of 25%, which corresponds to a maximum conscious contamination of a type I d' of 0.05 (see supplementary material in Skora, Livermore, Dienes, Seth, & Scott (*accepted*) for a method to determine the maximum proportion of conscious knowledge inadvertently contributing to observed unconscious performance at different proportions of trials classified as unconscious).

2.1. Method

2.1.1. Participants

46 participants (15 males, 1 unknown) with a mean age of 23.7 years ($SD = 9.84$) were recruited for participation via the University of Sussex online recruitment system. Sample size was determined using the Bayesian Stopping Rule, with data collection continuing until a sensitive result was obtained in the main (conditioning) task after excluding conscious trials (see 2.2.2. *Evidence of learning* for detail). All participated in exchange for course credit. All reported having normal or corrected-to-normal vision, and no current or history of cardiac or neurological illness. Ethical approval was granted by the School of Psychology ethics committee at the University of Sussex, and the study was conducted in accordance with the Declaration of Helsinki.

2.1.2. Stimuli and Materials

The experiment was conducted in Matlab 2017b (MathWorks, 2017), running Psychophysics Toolbox (Brainard, 1997). All stimuli were presented on a Samsung 2233RZ LCD monitor (1680 by 1050 pixels) with a refresh rate of 120Hz, with the aim of ensuring fast and precise stimulus presentation in line with previous recommendations (Wang & Nikolić, 2011).

The target stimuli included 12 neutral symbols (6 vertically symmetrical, 6 asymmetrical) obtained from Agathodaimon font in the main task, and two circular shapes in the perceptual discrimination task used for threshold finding. All were 180x180 pixels in size, and presented in light grey (RGB: 217 217 217) on white background. The stimuli were forward and backward masked with black and white noise masks, also 180x180 pixels in size, with block size of 3x3 pixels. The forward and backward masks were generated afresh on each trial by randomly re-scrambling a black and white noise image. Low contrast cues and the type of mask were deliberately chosen in order to increase the duration of presentation without conscious awareness, following Scott, Samaha, Chrisley and Dienes (2018).

2.1.3. Procedure

2.1.3.1. Threshold setting.

Participants were seated with their chin on a chin rest placed at 45cm distance from the screen. Each session began with the threshold of visual awareness determined individually for each participant using a masked perceptual discrimination task. Each trial began with a fixation cross (500ms), followed by a mask (300ms), a target cue (either a symmetrical or an asymmetrical circular shape, starting at 600ms), and another mask (300ms). After each sequence, participants were asked to determine whether the target cue was symmetrical or asymmetrical by pressing corresponding arrows. Next, they were asked to assess whether they had any confidence in their judgment, or if they

were guessing, also using corresponding arrows (following Scott et al., 2018). They were explicitly instructed to indicate 'some confidence' if they had any degree of confidence (even a hunch), and 'total guess' only if they felt they did not see the cue and were responding randomly. This distinction was applied as we considered any feelings, including hunches or intuitions, as conscious experience. Each time a correct response was made with confidence, the display duration of the target cue was reduced by 50ms on the following trial. When a duration of 100ms was reached, or the first guess response was made, the display duration returned to the previous level (+50ms), and subsequently reduced in 8.35ms steps on the following trials, corresponding to single screen refresh duration in a 120Hz screen. The reduction continued until a participant indicated guessing again, at which point the display time remained the same until a participants indicated guessing on six consecutive trials, regardless of the accuracy of responses. The cue display duration on those trials was set as their individual threshold of conscious awareness. The average display duration was 208ms ($SD = 64ms$).

2.1.3.2. *Conditioning task.*

The main task was adapted from the subliminal instrumental conditioning task approximating a deterministic instrumental learning scenario used previously (Pessiglione et al, 2008), in which participants learn to approach (Go response) or avoid (NoGo) the presented stimuli. Here, each trial consisted of a fixation cross (500ms), mask (300ms), target cue (display duration determined in the perceptual discrimination task), and mask (300ms), followed by a decision prompt in the form of a question mark, during which participants had 2 seconds to make a response (see *Fig.1.* for task chronology). While the question mark was presented, participants could decide if they want to 'approach' the cue by pressing the spacebar (Go) or 'avoid' the cue by refraining from pressing (NoGo). Pressing the space bar (Go) indicated a decision to take the risk, following which the participant could win 1 token (golden token displayed on the screen) or lose 1 token (a red cross over the golden token displayed), depending on the type of cue presented between the masks. Not pressing the space bar (NoGo) indicated a safe choice, which always resulted in a null outcome (greyed-out token). Participants were instructed that they should follow their instincts or gut feelings in making the decisions, as they should not be consciously aware of the type of cue presented.

Following feedback presentation, participants were asked to report whether they thought the cue was vertically symmetrical or asymmetrical, and their confidence in that judgment on a binary scale (some confidence or total guess). As in the threshold setting task, they were explicitly instructed to indicate 'some confidence' if they had any degree of confidence (even a hunch), and 'total guess' only if they felt they did not see the cue and were responding randomly. This provided a strict criterion of awareness, ensuring that even correct hunches were labelled as aware. Both responses were made

with the arrow keys before proceeding to the next trial. Following 3 correct and confident responses (indicating awareness), display duration was reduced by another single screen refresh (8.35ms). Participants were explicitly instructed that cue symmetry was unrelated to its rewarding or punishing nature. They were also shown examples of vertically symmetrical and asymmetrical stimuli on a different sample of images.

The task contained 2 blocks of 100 trials, each with 50 rewarding and 50 punishing trials in randomised order, as well as a short, 8 trial practice round prior to beginning, using a different set of random stimuli, for participants to become familiar with the general structure of the task. For each block, two target cues (one symmetrical, one asymmetrical) from the pool of 12 were randomly assigned to be rewarding or punishing, without replacement, in order to ensure each block contained a novel pair.

Note that the task was conducted as a control condition of a larger experiment (Skora, 2020; doctoral thesis, also <https://osf.io/b34uv/>), not reported here, where 4 additional blocks were conducted with an added manipulation (600 trials in total). The control blocks reported here are fully independent from the manipulation.

Figure 1. A single trial sequence. Participants are presented with a single cue (rewarding or punishing), rendered unconscious with forward-backward masking. After cue presentation, they are requested to make an approach (Go) or avoid (NoGo) response, following their instincts. In the event of a Go response, the response window closes immediately, but stays open for 2 seconds if no response is recorded. Here, a participant executed the Go response following a rewarding cue, and were rewarded with one token. Each trial ends with a cue symmetry judgment and confidence judgment to assess cue awareness.

2.2. Analyses and Results

2.2.1. *Exclusion criteria*

In order to ascertain that analyses are conducted only on trials that were truly unconscious, all individual trials where participants made a correct symmetry judgment with confidence were marked as aware and excluded (11.75% of all trials). Nineteen participants showed some awareness during the task which resulted in a drop in display duration (mean drop = 17ms; mode = 0; min = 0; max = 216ms). Of those who showed awareness, 7 participants (15%) who were aware on more than 25% (50 of 200) of all trials were excluded from further analysis. Two participants were excluded due to failing to make any Go responses, yielding a final sample of 37 participants.

For the exploratory reaction time (RT) analysis, RTs under 100ms or over 2 standard deviations from individual means were excluded (8%). One participant who lost more than 25% of RT trials (50) to those criteria was excluded, yielding a sample of 36 participants for the RT analysis. Note that NoGo trials yielded no RTs, thus were not included in the analysis. Please see Appendix A for a flowchart of the exclusion criteria for both analyses.

2.2.2. *Evidence of learning: Performance in the conditioning task*

On average, participants were found to execute more Go responses (58%) than NoGo responses (42%), regardless of cue type. To account for this response bias, type I d' (a Signal Detection Theoretic measure of sensitivity to signal versus noise; Stanislaw & Todorov, 1999) was computed for the main conditioning task, treating Go responses to rewarding cues as Hits, and Go responses to punishing cues as False Alarms. The resulting measure of sensitivity can be taken as evidence of successful learning (i.e. discrimination between the cues) if it is significantly above 0.

At the group level, d' scores were entered into a one-sample t-test against 0, which indicates no ability to discriminate the stimuli (no sensitivity between signal versus noise, akin to chance performance). A Bayes Factor (B; Dienes, 2015, 2016) was computed for the difference, with a half-normal distribution, mean specified as 0, and d' of 0.7 as the SD of the mean (corresponding to the expected effect size, based on Pessiglione et al., 2008). 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 unaware sample. To indicate the robustness of Bayesian conclusions, a robustness region will be reported for each B, giving the range of scales that qualitatively support the same conclusion (i.e. evidence as insensitive, or as supporting H_0 , or as supporting H_1), notated as:

$RR_{B>3}[x1, x2]$ or $RR_{B<1/3}[x1, x2]$ or $RR_{1/3>B>3}[x1, x2]$ where $x1$ is the smallest SD that gives the same conclusion and $x2$ is the largest (see Dienes, 2019).

When aware trials were excluded, d' was not significantly different from 0 ($M = 0.02$, $SE = 0.05$; $t(36) = 0.46$, $p = 0.649$, $B_{H(0,0.7)} = 0.10$, $RR_{B<1/3}[0.23, \text{Inf}]$; see *Figure 2*). As such, we found evidence for absence of unconscious instrumental conditioning. In contrast, analysis of the 1081 (11.75%) aware trials shows that aware d' was significantly over 0 ($M = 0.72$, $SE = 0.19$, $t(38) = 3.76$, $p = 0.001$, $B_{H(0,0.7)} = 420.22$, $RR_{B>3}[0.06, 185]$), demonstrating sensitive evidence for learning. In the full sample (without applying awareness exclusions, but excluding the two participants who failed to make any Go responses), total d' was also significantly over 0, albeit B remained insensitive ($M = 0.11$, $SE = 0.05$, $t(43) = 2.10$, $p = 0.042$, $B_{H(0,0.7)} = 1.56$, $RR_{1/3>B>3}[0.4, 3.5]$). We also confirmed that the distribution of aware trials across the trial block did not show an increasing tendency (*Figure 3*), suggesting that conscious knowledge was unlikely to inform subsequent trials in a manner resulting in more conscious trials. However, when assessing the correlation between the number of aware trials and total d' (without awareness-related trial and participant exclusions) we found no significant correlation ($r = 0.154$, $p = 0.319$; see *Figure 3*).

Figure 2. Panel A, left: average type I d' for unaware trials only, with individual data points. **Panel A, right:** average proportions of correct responses (Go to rewarding and NoGo to punishing cues) across the length of the trial block, for unaware trials only. Ribbon represents a regression line (90% CI). Participants' choices throughout the duration of the block remained erratic and around chance level, failing to display any evidence of learning the correct decisions. **Panel B, left:** average type I d' for aware trials only, with individual data points. **Panel B, right:** average proportions of correct responses across the length of the trial block, for aware trials only. Ribbon represents a regression line (90% CI). Note that aware trials were not distributed evenly between participants.

Figure 3. Left: Histogram showing the distribution of aware trials across the length of a trial block (two blocks are averaged together). **Right:** Total d' (all trials, before awareness-related trial and participant exclusions) plotted against the number of aware trials. Ribbon represents 95% CI.

2.2.3. Exploratory analysis: Reaction times

An $RT_{\text{difference}}$ index was computed by subtracting RTs to rewarding cues from RTs to punishing cues. As such, 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 learning (e.g. Atas, Faivre, Timmermans, Cleeremans, & Kouider, 2014). Zero indicates that there was no difference between the two.

The index was entered into a one-sample t-test against 0. B was computed with a half-normal distribution, mean specified as 0, and a value of 34ms as the SD of the mean (obtained from a past study which found a RT difference in the absence of performance effects in a similar task (Atas et al., 2014)). The t-test revealed that $RT_{\text{difference}}$ was not significantly different from 0 ($M = 14.06$, $SE = 17.81$, $t(35) = 0.79$, $p = 0.435$, $B_{H(0,34)} = 0.90$, $RR_{1/3 > B > 3} [0, 120]$).

2.3. Conclusions of Experiment 1

Experiment 1 attempted to conceptually replicate the unconscious instrumental (trace) conditioning task, used by Pessiglione *et al* (2008). The results demonstrate that no unconscious instrumental conditioning was achieved – when participants had no conscious awareness of the rewarding or punishing nature of the stimuli, they were unable to discriminate between them and adjust their behaviour accordingly (i.e. choose to Go or NoGo, respectively). Bayes Factors were used to supplement null hypothesis significance tests, allowing us to assert that the null result obtained

indicates absence of learning. This result is a failure to corroborate the conclusions of Pessiglione et al (2008).

In light of previous research and the theoretical contributions about the limits of unconscious processing (e.g. Mudrik et al., 2014), we considered it plausible that the trace conditioning paradigm used here might make it harder to achieve learning, even if it is possible without cue awareness. In the present task, participants had to learn from two events – the stimulus and its consequence – separated by an interval of up to 2 seconds, during which they may or may not execute an action. If the large temporal interval is preventing the unconsciously presented predictive cue from being integrated with subsequent information, shortening the delay might offer a solution. We address this possibility in Experiment 2 by amending the task to constitute delay, rather than trace, instrumental conditioning.

3. Experiment 2

Experiment 2 sought to conceptually replicate the unconscious instrumental conditioning task again, this time shortening the stimulus-outcome delay. To achieve this, we used a delay conditioning paradigm, while retaining as much of the structure of the original task as possible. As such, we employed the same Go/NoGo task, where participants learned to associate predictive cues with reward and punishment, and to adjust their behaviour accordingly by responding (Go) only to rewarding cues, and refraining from a response (NoGo) to punishing cues. In contrast to Experiment 1, the response and reinforcement delivery occurred during (rather than following) stimulus presentation, constituting delay conditioning. Prolonged stimulus presentation was enabled through continuous flash suppression (CFS; Tsuchiya & Koch, 2005). CFS uses interocular suppression to render the target stimulus unconscious by presenting it to the non-dominant eye, while simultaneously presenting a high-contrast flashing pattern (called the Mondrian pattern) to the dominant eye. Instead of monetary rewards, Experiment 2 used appetitive and aversive tastes as primary rewards, allowing for extended delivery to overlap with stimulus presentation. In a fashion identical to Experiment 1, a trial-by-trial awareness check was used, with the same exclusion threshold of a maximum of 25% of individual trials.

3.1. Method

3.1.1. Participants

39 participants (5 male) with a mean age of 21 years ($SD = 5.2$) were recruited for the study via the University of Sussex online recruitment system. Sample size was determined with the Bayesian Stopping Rule (see section 3.2.2. *Evidence for learning* for detail) after excluding conscious trials. All

participated in exchange for course credit. Participants were required to have normal or corrected-to-normal binocular vision, no current or history of neurological illness and/or diabetes, and no allergies to gluten, lactose, milk or chocolate. Further exclusion criteria included dislike of chocolate milkshakes and/or smoking more than five cigarettes a day, as those factors might affect responsiveness to tastes. Participants were requested to refrain from eating and drinking (apart from water) for 2 hours prior to the experiment. Ethical approval was granted by the School of Psychology ethics committee at the University of Sussex, and the study was conducted in accordance with the Declaration of Helsinki. The study was pre-registered on Open Science Framework at <https://osf.io/jt8m9>.

3.1.2. *Stimuli and Materials*

The experiment was implemented in Matlab 2018b (MathWorks, 2018) with the Cogent toolbox (UCL LoN, 2003). All stimuli were presented on a Dell monitor (1280 by 1024 pixels), with a refresh rate of 60Hz¹. Binocular presentation was ensured with a mirror stereoscope. Responses were collected with a standard keyboard.

The target stimuli were three geometrical shapes, including one symmetrical and two asymmetrical shapes (generated by randomly overlaid shapes). All stimuli were 120 by 120 pixels in size, presented in dark grey (RGB: 80,80,80) on lighter grey (RGB: 128,128,128) background. The Mondrian patterns for the CFS were squares composed of smaller colourful blocks (see Figure 4). On each trial, the Mondrian patterns were dynamically flashing at the rate of 10Hz through random re-scrambling of the order of 10 patterns.

The solution used as the appetitive stimulus was a chocolate milkshake (Shaken Udder Chocolush). The aversive stimulus was a mixture of food-grade quinine monohydrochloride dihydrate, at the FDA-approved maximum concentration (80mg/l; Sigma Aldrich), and standard table salt (NaCl; 12.5g/l). Custom-made syringe pumps were used to administer the solutions to participants' mouths through two clean silastic tubes (one per solution), inserted into the corners of the mouth and placed in the middle of the tongue. The pumps were triggered directly from the Matlab task script, delivering a 0.5ml bolus of the corresponding solution at a time.

3.1.3. *Procedure*

3.1.3.1. *Conditioning task.*

The conditioning task followed a similar Go/NoGo procedure as outlined in Experiment 1, amended for presentation under continuous flash suppression (CFS; Tsuchiya & Koch, 2005). During

¹ A 120Hz screen, as used in Experiment for the purpose of sensitive threshold finding, was no longer necessary for a CFS setup.

the unconscious presentation of each cue, participants were asked to decide if they want to ‘approach’ the cue by pressing the spacebar (Go) or ‘avoid’ the cue by refraining from pressing (NoGo). Approaching the cue resulted in the delivery of a 0.5ml bolus of the corresponding solution – appetitive if they approached the rewarding cue, and aversive if they approached the punishing cue. A NoGo response was again considered a safe choice, which resulted in no taste outcome. In order to further incentivise learning and ensure participants do not default to the safe choice, each correct response additionally resulted in winning a token. Participants were instructed to maximise their token winnings for an extra prize draw entry.

The cues were presented unconsciously using CFS, with the cue itself presented to participants’ non-dominant eye, and a colourful Mondrian pattern flashing at the rate of 10Hz presented simultaneously to the dominant eye through a mirror stereoscope, placed atop a chinrest at a 50cm distance from the screen. Ocular dominance was established prior to beginning with a standard Miles test (Miles, 1930). The unconscious presentation of each cue started with a fade-up period of 500ms. A further 500ms after the cue reached full contrast, the fixation dot (overlaid on displays to both eyes) changed colour from black to white, indicating that response was required. The response window remained open for 1000ms. When the participant executed a Go response, a pump was triggered to deliver the corresponding taste outcome (appetitive or aversive, depending on the type of the cue). The taste delivery continued for 2000ms, while the cue remained on-screen (concurrently with the Mondrian pattern) for 1000ms, ensuring a full second of temporal overlap between the cue and the outcome (reward or punishment), constituting delay conditioning. See *Figure 4* for an illustration of the trial sequence.

Following the taste delivery, participants were asked to report whether the cue was symmetrical or asymmetrical, and their confidence in that judgment on a binary scale (some confidence or total guess; the same instructions were given as in Experiment 1). Both responses were made with the arrow keys before proceeding to the next trial.

The task consisted of 1 block of 100 trials, with a total of 50 rewarding and 50 punishing trials in a randomised order, as well as a short practice round. The task used two stimuli (one symmetrical, one asymmetrical), out of which one was randomly assigned to be appetitive, and one to be aversive, for each participant.

Figure 4. Single trial sequence in Experiment 2, illustrating the binocular presentation with the non-dominant eye (left) receiving the cue (predictive of reward or punishment), and the dominant eye receiving the flashing Mondrian pattern. Following the cue reaching full contrast, participants are required to make an approach (Go) or avoid (NoGo) response, and are rewarded with 0.5ml of chocolate milkshake or punished with a 0.5ml of a quinine/saline solution accordingly. The taste delivery overlaps with cue presentation for 1000ms and continues for another 1000ms. Each trial ends with a cue symmetry judgment and confidence judgment to assess cue awareness on the trial.

3.1.3.2. Liking and familiarity ratings.

Following the conditioning task, the two cues used and the third, previously unseen cue were presented on the screen in a randomised order in a single column, alongside corresponding VAS scales. First, participants were asked to rate their liking of each stimulus, on a scale from 'Not at all' to 'Extremely'. Next, they were asked to rate their familiarity with each, again on the same scale. The scale was presented to participants as unnumbered, but was bounded at 1-600. Following the ratings, participants were thanked and debriefed.

3.2. Results

3.2.1. Exclusion criteria.

The exclusion criteria were identical to Experiment 1. 31.8% of all trials were marked as aware and excluded. 20 (51%) participants who were aware on more than 25% (25) of all trials were classified as aware. The remaining 19 participants were found to be aware on 9.6% of all trials. The proportion of participants classified as aware on more than 25% of the trials was surprisingly high. We suspect that the length of stimulus presentation under CFS (2 seconds) may have made it easier for some participants to see the stimuli on some occasions. Once seen, the stimuli are typically easier to see on subsequent trials, which is what likely resulted in the high rate of aware trials. Because of the risk of regression to the mean following post-hoc selection of a large number of participants (Shanks, 2017), we found that the maximum proportion of affected trials could be 55%, corresponding to a d' of 0.99

(computed with the method used in Experiment 1). Excluding such a large proportion of participants could lead to a false-positive result (presence of learning), where some conscious knowledge is mistakenly included in unconscious knowledge, in this case overinflating the d' by up to 0.99. As we will show below, this possibility is not relevant here, because the result shows absence of learning (even after allowing for the contribution of conscious knowledge).

For maximum rigour and transparency, analysis was conducted on both the unaware sample as planned, and the full sample (those classified as aware and unaware), after excluding individual aware trials.

For the reaction time (RT) analysis, RTs under 100ms or over 2 standard deviations from individual means were excluded (4.49%). No participant lost more than 12% of their trials to this criterion. Again, NoGo trials were not included in the analysis.

3.2.2. *Evidence of learning: Performance in the conditioning task.*

As in Experiment 1, participants were found to execute more Go responses (62%) than NoGo responses (38%), regardless of cue type. Type I d' was computed again to account for this response bias. At the group level, d' scores were entered into a one-sample t-test against 0, which indicates no ability to discriminate between the stimuli (chance performance). B was computed for the difference, with a half-normal distribution, mean specified as 0, and d' of 0.7 as the SD of the mean (corresponding to the expected effect size from Pessiglione et al., 2008). In line with the Bayesian Stopping Rule, data collection continued until a sensitive result was found either in support of the H_0 (absence of unconscious learning; $B < 0.3$) or the H_1 (evidence for unconscious learning; $B > 3$).

For the sample with aware participants excluded ($N = 19$), total d' was not significantly different from 0 ($M = -0.10$, $SE = 0.10$; $t(18) = -1.00$, $p = 0.329$, $B_{H(0,0.7)} = 0.07$, $RR_{B<1/3} [0.17, \text{Inf}]$; see Figure 5). As such, we found evidence for absence of unconscious instrumental conditioning. For the full sample (participants classified both as aware and unaware, but with excluded aware trials; $N = 39$), total d' was also not significantly different from 0, but the B was insensitive ($M = 0.13$, $SE = 0.14$; $t(38) = 0.97$, $p = 0.339$, $B_{H(0,0.7)} = 0.48$, $RR_{1/3>B>3} [0, 1.1]$). In contrast, analysis of the 1240 (31.8%) aware trials only shows that aware d' was significantly over 0 ($M = 0.99$, $SE = 0.27$, $t(32) = 3.60$, $p = 0.001$, $B_{H(0,0.7)} = 250.25$, $RR_{B>3} [0.1, 140]$), demonstrating sensitive evidence for learning. In the full sample (without applying any awareness exclusions), total d' was also significantly over 0 ($M = 0.35$, $SE = 0.12$, $t(38) = 2.63$, $p = 0.012$, $B_{H(0,0.7)} = 21.01$, $RR_{B>3} [0.05, 5.5]$). Again, we confirmed that the distribution of aware trials across the trial block did not show an average increasing tendency (Figure 6). However, we found a significant correlation between the number of aware trials and total d' (without awareness-related trial and participant exclusions; $r = 0.571$, $p < .001$; see Figure 6).

A. Unaware sample ($N = 19$)

B. Full sample ($N = 39$)

C. Aware sample ($N = 33$)

Figure 5. Panel A, left: average type I d' with individual data points (unaware sample, $N=19$). **Panel A, right:** average proportions of correct responses (Go to rewarding and NoGo to punishing cues) across the length of the trial block. Ribbon represents a regression line (90% CI). **Panel B, left:** average type I d' with individual data points (full sample, $N = 39$). **Panel B, right:** average proportions of correct responses across the length of the trial block. Ribbon represents a regression line (90% CI). Participants' choices throughout the duration of the block remained erratic and around chance level, failing to display any evidence of learning the correct decisions. Note that individual aware trials were excluded. **Panel C, left:** average type I d' with individual data point (aware trials only, $N = 33$). **Panel C, right:** average proportions of correct responses across the length of the trial block. Ribbon represents a regression line (90% CI).

Figure 6. Left: Histogram showing the distribution of aware trials across the length of a trial block in Experiment 2. **Right:** Total d' (all trials, before awareness-related trial and participant exclusions) plotted against the number of aware trials. Ribbon represents 95% CI.

3.2.3. Exploratory analysis: Reaction times

The $RT_{\text{difference}}$ index was computed again by subtracting RTs to rewarding cues from RTs to punishing cues (with positive values indicating that participants took a longer time to respond to punishing than to rewarding cues), and entered into a one-sample t-test against 0. B was computed with a half-normal distribution, mean specified as 0, and a value of 34ms as the SD of the mean (obtained from a past study which found a RT difference in absence of performance effects in a similar task (Atas et al., 2014)). For the sample with aware participants excluded ($N = 19$), the t-test revealed

that $RT_{\text{difference}}$ was not significantly different from 0 ($M = 7.98\text{ms}$, $SE = 9.69$, $t(18) = 0.83$, $p = 0.420$, $B_{H(0,34)} = 0.59$, $RR_{1/3>B>3}[0,75]$). The same was true for the full sample ($N = 39$; $M = 2.84$, $SE = 15.69$, $t(38) = 0.18$, $p = 0.857$, $B_{H(0,34)} = 0.48$, $RR_{1/3>B>3}[0,56]$).

3.2.4. *Liking and familiarity ratings*

Repeated-measures analyses of variance (ANOVAs) were conducted for liking of the three presented cues (rewarding, punishing, and unseen) and familiarity with them, on the unaware sample ($N=19$) and the full sample ($N = 39$). Due to the lack of reasonable expected values, B s were computed for all pairwise comparisons with a default Cauchy (0, $r = 1/\sqrt{2}$) prior in JASP (JASP Team, 2020). In the unaware sample, the ANOVA revealed no significant main effect of cue type (rewarding, punishing, unseen) on liking ($F(2,36) = 0.13$, $p = 0.875$), or familiarity ($F(2,36) = 0.51$, $p = 0.606$). No differences were found in any of the pairwise comparisons between the three cues on liking (all $p > 0.05$, all $B_{10} < 0.3$), suggesting that participants in the unaware sample developed no increased liking of the reward-predicting stimulus over the punishment-predicting one. Similarly, no differences were found on familiarity, with the exception of an insensitive difference in familiarity between the rewarding and unseen cues ($B_{10} = 0.40$).

In the full sample (i.e. including participants showing some awareness of the stimuli during the task), the ANOVA showed a significant main effect of cue type on liking ($F(2,76) = 4.74$, $p = 0.012$), with the rewarding cue receiving a significantly higher liking rating ($M_{\text{rew}} = 349.6$, $SE_{\text{rew}} = 28.17$) than the unseen cue ($M_{\text{unseen}} = 234.03$, $SE_{\text{unseen}} = 23.32$; $p = 0.006$, $B_{10} = 17.34$). The difference in liking between the rewarding and punishing ($M_{\text{pun}} = 253.3$, $SE_{\text{pun}} = 27.78$) cues was not sensitive ($p = 0.138$, $B_{10} = 1.16$). There was no difference between the punishing and unseen cues ($p = 1.000$, $B_{10} = 0.19$). The ANOVA also revealed a significant main effect of cue type on familiarity in the full sample ($F(2,76) = 12.35$, $p < 0.001$). Unsurprisingly, both the rewarding ($M_{\text{rew}} = 345.51$, $SE_{\text{rew}} = 33.01$) and punishing ($M_{\text{pun}} = 338.62$, $SE_{\text{pun}} = 32.03$) cues received a higher familiarity rating than the unseen cue ($M_{\text{unseen}} = 157.18$, $SE_{\text{unseen}} = 26.26$; $p < 0.001$, $B_{10} > 100.00$ for both comparisons). There was no difference in familiarity between the rewarding and punishing cues ($p = 1.000$, $B_{10} = 0.18$). All comparisons are Bonferroni-corrected.

3.3. Conclusions of Experiment 2

Experiment 2 made a second attempt to conceptually replicate unconscious instrumental conditioning, this time with a delay conditioning paradigm. The results demonstrate that no unconscious instrumental conditioning was achieved, even when the temporal delay between the presentation of the stimulus and its consequence overlapped in time. Again, when participants had

no conscious awareness of the stimuli, they were unable to adjust their responses accordingly (i.e. choose to Go in response to rewarding stimuli and NoGo to punishing stimuli). Bayes Factors allowed us to assert that the null result obtained indicates lack of learning. This result constitutes a failure to conceptually replicate unconscious instrumental conditioning.

4. General Discussion

The extent to which complex forms of learning, such as instrumental conditioning, are possible without conscious awareness is still an unresolved question. To our knowledge, evidence for unconscious instrumental learning has been limited (Pessiglione et al., 2008; Mastropasqua & Turatto, 2015) and disputed (Reber et al., 2018). Yet, instrumental behaviour is a vital and one of the oldest forms of adaptive behaviour (both phylogenetically and ontogenetically). As such, we consider the extent to which it can or cannot proceed unconsciously an important question to settle. In the present paper, we revisited unconscious instrumental conditioning in two modes of conditioning: trace (Experiment 1) and delay (Experiment 2). We enhanced the sensitivity of the paradigms, and manipulated the temporal delay between stimulus and outcome – previously identified as a key aspect of unconscious learning.

Experiment 1 attempted to revisit the original Pessiglione *et al* (2008) finding obtained in an unconscious instrumental (trace) conditioning task. Our results show evidence for absence of unconscious instrumental conditioning – when participants had no conscious awareness of the rewarding or punishing nature of the stimuli, they were unable to discriminate between them and adjust their behaviour accordingly (i.e. choose to Go or NoGo, respectively). Bayes Factors allowed us to assert that the null result obtained indicates absence of learning. This result is a failure to conceptually replicate the original Pessiglione et al (2008) finding. Nonetheless, it is important to recognise that, as a conceptual replication, the present study introduced a number of changes to the task, thus precluding a more direct comparison. A direct replication can be found in Skora et al. (*accepted*).

Following the failure to demonstrate unconscious learning, we hypothesised that the large temporal gap between the stimulus and its outcome might hinder the ability to learn. Past theoretical and empirical work has shown that unconscious associative learning or priming is more likely to occur at shorter temporal intervals (e.g. van Gaal et al., 2012), or when the events are overlapping in time (delay conditioning; Clark & Squire, 1998). Thus, Experiment 2 made a second attempt to investigate unconscious instrumental conditioning, this time with a delay conditioning paradigm. Again, the results show evidence for absence of unconscious instrumental conditioning, even when the stimulus, response and reinforcement temporally overlapped. As before, when participants had no conscious

awareness of the stimuli, they failed to adjust their responses accordingly (i.e. choose to Go in response to rewarding stimuli and NoGo to punishing stimuli). Bayes Factors again allowed us to assert that the null result obtained indicates absence of learning. As in Experiment 1, this result constitutes a failure to conceptually replicate unconscious instrumental conditioning.

Despite extensively demonstrated effects in simpler associative learning or priming tasks, it appears that unconscious instrumental learning cannot be achieved even if the stimulus-outcome delay is shortened to the point of total overlap. As such, perhaps the problem is not the length of the delay. Instead, it may in fact be the instrumental component of the task itself. Previous theoretical contributions have already suggested that conscious access might be necessary for higher-order processes, including selective decision-making (Baars, 2002; Dehaene & Changeux, 2011; Lamme, 2006). Under those accounts, consciousness allows for information exchange between distinct cognitive modules through long-lasting, long-range recurrent interactions between brain areas. Hence, it might be key for flexible and lasting information processing strategies. Regardless of the temporal intervals between events, instrumental learning is a complex process, necessitating integration of information across distinct modalities involved in processing the visual input, extracting its predictive value from past interactions, deploying the response, processing the reinforcement, and comparing the expected outcome with actual outcome to update the representations of stimulus value. Such integration might simply not be possible when the stimulus is presented unconsciously, as it may not have the capacity to evoke such a broad range of activity. Corroborating this interpretation, past work has already pointed out that unconscious stimuli are less likely to evoke long-term behavioural adaptations than conscious stimuli (e.g. in post-error slowing or conflict adaptation; (de Lange, van Gaal, Lamme, & Dehaene, 2011; Kunde, Reuss, & Kiesel, 2012; Reber et al., 2018; Travers et al., 2018; van Gaal et al., 2012). Instrumental learning might constitute a similar case.

If instrumental learning does indeed require conscious access to proceed, the earlier results supporting unconscious instrumental learning may have resulted from transient moments of awareness during the learning task contributing to the unconscious knowledge. Those may have been undetected due to lack of immediate, sensitive measures (Berry & Dienes, 1993; Newell & Shanks, 2013; Shanks & St. John, 1994), such as trial-by-trial awareness checks. In the two experiments reported here, it was evident that most participants were aware of the nature of the stimuli on at least some of the trials, and that there are individual differences in those proportions. Capturing aware trials and excluding them ensured that the trials analysed were genuinely unconscious, and constituted a considerable improvement over previously employed methods. Crucially, it also avoids the potential for anything other than a negligible effect from regression to the mean (Shanks, 2017), the extent of which we computed for both experiments.

Nonetheless, it is worthwhile to address whether our methodology, although more rigorous, made it more difficult for learning to happen. While a trial-by-trial awareness check was essential to rigour, it introduced a small, variable time lag (around 1 second) between the trials. This might have further disrupted the process of temporal integration across trials, affecting the long-term consolidation of the stimulus-outcome associations learned within them. However, it is the integration of events within the trial that is vital to learning in the first place. As such, we consider any plausible interference of the inter-trial delay to be minor.

Another important consideration is the large number of participants displaying awareness of stimuli in Experiment 2. According to our criteria, we allowed a maximum of 25% of aware trials to be excluded per participant, corresponding to the level at which the risk of aware trials in observed unconscious trials is marginal, amounting to 0.05 of total d' (Skora et al., accepted). Yet, a half of the sample was found to be aware at more than 25% of the trials and hence excluded. Exclusion of this surprisingly large sample poses a risk of regression to the mean, where the remaining sample might show a false-positive (evidence of learning) purely as a statistical artefact (Shanks, 2017). In the interest of maximum transparency, we analysed both samples (with individual aware trials excluded), showing that both exhibited absence of learning – thus rendering the risk of a false-positive on the remaining unaware sample irrelevant.

The surprisingly large proportion of participants showing awareness of the stimuli under CFS on over 25% of trials is likely to have stemmed from the long duration of each CFS trial (2 seconds), which was necessary to allow for cue presentation to temporally overlap with the response and the response-contingent taste reinforcement. With known individual differences in CFS (e.g. Gayet & Stein, 2017), it is plausible that some people may have seen the cue presented under interocular suppression at this duration easier than others. Once seen, the stimulus typically becomes visible easier again to perceive on subsequent occasions. This could be an issue for consideration for future research. Here, we were able to reliably account for that in the analyses presented.

Finally, the absence of unconscious instrumental conditioning reported here might seem intriguing in light of previous evidence that instrumental conditioning can proceed in simpler – presumably unconscious – organisms, such as fruit flies (*Drosophila*), sea slugs (*Aplysia*), or pond snails (*Lymnaea*) (Brembs, 2003, 2011; Brembs & Heisenberg, 2000). One interpretation of this discrepancy could be that unconscious instrumental conditioning is indeed possible in certain forms, perhaps in simpler tasks than those applied here, or in more naturalistic situations. However, the question of presence or absence of consciousness in more primitive animals is far from settled, and far beyond the scope of this paper. Some have argued that consciousness is intimately tied to action, and could be construed as providing a frame of reference for agents' interactions with the world (Clark, 2016;

Land, 2012; Merker, 2005). As such, it may be found (in a very rudimentary form) in simpler organisms (Baars, 2005; Merker, 2005), perhaps including invertebrates (Barron & Klein, 2016; Muñoz et al., 2020). Nonetheless, the nature of consciousness and its functional relationship with action are still a matter of active debate.

Another – more parsimonious – interpretation could be that humans and simpler animals differ in terms of the conditions necessary for instrumental learning. Human instrumental conditioning may have developed new dependencies to replace or extend simple stimulus-response associations, catering for more complex, flexible, or longer-term decision-making. Such new dependencies may make it impossible for humans to be instrumentally conditioned in the absence of conscious awareness, which is argued to permit long-lasting, long-range recurrent processing strategies (Baars, 2002; Dehaene & Changeux, 2011; Lamme, 2006), and in turn to sustain flexible learning and action deployment. It is plausible that simpler animals follow simpler stimulus-action learning routines which are not optimised for more complex or flexible decision-making, and hence could be executed without conscious awareness. It has also been pointed out that simple forms of instrumental conditioning in both vertebrates and invertebrates can arise as a corollary, or a special case, of classical conditioning, rather than instrumental conditioning per se (see Moore, 2004 for a thorough overview of the evolution of the many different types of learning across species). Indeed, as mentioned earlier, classical conditioning and other simple forms of associative learning have been shown to proceed without conscious awareness (e.g. Clark & Squire, 1998).

To summarise, we used a trace conditioning paradigm (Experiment 1, following a method used by Pessiglione et al., 2008) and a delay conditioning paradigm (Experiment 2) to investigate the feasibility of unconscious instrumental conditioning with different temporal delays. We provide rigorous evidence demonstrating that instrumental conditioning cannot be achieved without stimulus awareness, corroborating the emerging evidence that complex types of learning, including instrumental conditioning, may require conscious awareness. This provides strong support for the theoretical proposals that unconscious stimuli fail to bring about long-term behavioural adaptations, and that consciousness might be necessary for complex processes, where selective action is required. As such, this result contributes to mapping out the boundaries of unconscious processing in adaptive behaviour.

5. Acknowledgments

This work was supported by the Sackler Centre for Consciousness Science, funded by the Sir Mortimer and Theresa Sackler Foundation, and the School of Psychology at the University of Sussex.

The authors are also thankful to Dr Hans Crombag and Prof Martin Yeomans at the University of Sussex for providing essential resources.

6. Supplementary Materials

The complete sets of materials both experiments and the corresponding data have been made publically available on the Open Science Framework. Materials from Experiment 1 are available at <https://osf.io/b34uv/>, and the corresponding data at <https://osf.io/bvyf3/>. Materials and data from Experiment 2 are available at <https://osf.io/drmcg/>.

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](https://doi.org/10.1016/S1364-6613(00)01819-2)
- Baars, B. J. (2005). Subjective experience is probably not limited to humans: The evidence from neurobiology and behavior. *Consciousness and Cognition*, 14(1), 7–21.
<https://doi.org/10.1016/j.concog.2004.11.002>
- Baars, B. J., Ramsøy, T. Z., & Laureys, S. (2003). Brain, conscious experience and the observing self. *Trends in Neurosciences*, 26(12), 671–675. <https://doi.org/10.1016/j.tins.2003.09.015>
- Barron, A. B., & Klein, C. (2016). What insects can tell us about the origins of consciousness. *Proceedings of the National Academy of Sciences*, 113, 4900–4908.
<https://doi.org/10.1073/pnas.1520084113>
- Berry, D. C., & Dienes, Z. (1993). *Implicit Learning: Theoretical and Empirical Issues*. Hove, UK: Lawrence Erlbaum.
- Birbaumer, N., Klucken, T., Angrilli, A., Andreatta, M., & Pauli, P. (2015). Appetitive vs. Aversive conditioning in humans. *Front. Behav. Neurosci*, 9, 128.
<https://doi.org/10.3389/fnbeh.2015.00128>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436.
- Brembs, B. (2003). Operant conditioning in invertebrates. *Current Opinion in Neurobiology* *Current Opinion in Neurobiology*, 13, 710–717. <https://doi.org/10.1016/j.conb.2003.10.002>
- Brembs, B. (2011). Spontaneous decisions and operant conditioning in fruit flies. *Behavioural Processes*, 87, 157–164. <https://doi.org/10.1016/j.beproc.2011.02.005>
- Brembs, B., & Heisenberg, M. (2000). The operant and the classical in conditioned orientation of *Drosophila melanogaster* at the flight simulator. *Learning and Memory*, 7, 104–115.
- Clark, A. (2016). *Surfing Uncertainty: Prediction, action and the embodied mind*. Oxford, UK: Oxford University Press.
- Clark, R. E., & Squire, L. R. (1998). Classical conditioning and brain systems: The role of awareness. *Science*, 280, 77–81.
- de Lange, F. P., van Gaal, S., Lamme, V. A. F., & Dehaene, S. (2011). How awareness changes the relative weights of evidence during human decision-making. *PLoS Biology*, 9(11).
<https://doi.org/10.1371/journal.pbio.1001203>
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic

- evidence and a workspace framework. *Cognition*, 79, 1–37. Retrieved from https://www.jsmf.org/meetings/2003/nov/Dehaene_Cognition_2001.pdf
- 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, Stanislas, 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, Stanislas, & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, 79, 1–37. Retrieved from www.elsevier.com/locate/cognit
- Dehaene, Stanislas, Naccache, L., Cohen, L., Bihan, D. Le, Mangin, J.-F., Poline, J.-B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, 4(7), 752–758. <https://doi.org/10.1038/89551>
- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results, 5(781), 1–17. <https://doi.org/10.3389/fpsyg.2014.00781>
- Dienes, Z. (2015a). How bayesian statistics are needed to determine whether mental states are unconscious. In M. Overgaard (Ed.), *Behavioral Methods in Consciousness Research* (pp. 199–220). Oxford: Oxford University Press.
- Dienes, Z. (2015b). How many participants might I need? - YouTube. Retrieved November 14, 2019, from 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>
- 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>
- Gayet, S., & Stein, T. (2017). Between-Subject Variability in the Breaking Continuous Flash Suppression Paradigm: Potential Causes, Consequences, and Solutions. *Frontiers in Psychology*, 8, 437. <https://doi.org/10.3389/fpsyg.2017.00437>
- 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.
- Kunde, W., Reuss, H., & Kiesel, A. (2012). Consciousness and cognitive control. *Advances in Cognitive Psychology*, 8(1), 9–18. <https://doi.org/10.5709/acp-0097-x>

- 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>
- 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>
- Martin-Soelch, C., Linthicum, J., & Ernst, M. (2007). Appetitive conditioning: neural bases and implications for psychopathology. *Neuroscience and Biobehavioral Reviews*, 31, 426–440. <https://doi.org/10.1016/j.neubiorev.2006.11.002>
- Mastropasqua, T., & Turatto, M. (2015). Attention is necessary for subliminal instrumental conditioning. *Scientific Reports*, 5(1290), 1–7. <https://doi.org/10.1038/srep12920>
- 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](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>
- Moore, B. R. (2004). The evolution of learning spaces. *Biological Review*, 79, 301–335. <https://doi.org/10.1145/1449956.1450017>
- 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>
- Muñoz, R. N., Leung, A., Zecevik, A., Pollock, F. A., Cohen, D., Van Swinderen, B., ... Modi, K. (2020). General anesthesia reduces complexity and temporal asymmetry of the informational structures derived from neural recordings in *Drosophila*. *Physical Review Research*, 2. <https://doi.org/10.1103/PhysRevResearch.2.023219>
- 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>
- Olsson, A., & Phelps, E. A. (2004). Learned Fear of “Unseen” Faces After Pavlovian, Observational, and Instructed Fear. *Psychological Science*, 15, 822–828. Retrieved from

- <http://journals.sagepub.com/doi/pdf/10.1111/j.0956-7976.2004.00762.x>
- Palminteri, S., Clair, A.-H., Mallet, L., & Pessiglione, M. (2012). Similar improvement of reward and punishment learning by Serotonin Reuptake Inhibitors in Obsessive-Compulsive Disorder. *Biological Psychiatry*, 72, 244–250. <https://doi.org/10.1016/j.biopsych.2011.12.028>
- Palminteri, S., Lebreton, M., Worbe, Y., Grabli, D., Hartmann, A., & Pessiglione, M. (2009). Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. *Proceedings of the National Academy of Sciences of the United States of America*, 106(45), 19179–19184. <https://doi.org/10.1073/pnas.0904035106>
- Parkinson, J. A., Cardinal, R. N., & Everitt, B. J. (2000). Limbic cortical-ventral striatal systems underlying appetitive conditioning. *Progress in Brain Research*, 126, 263–285. [https://doi.org/10.1016/S0079-6123\(00\)26019-6](https://doi.org/10.1016/S0079-6123(00)26019-6)
- 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., Antoniadis, 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>
- Rosenthal, C. R., Kennard, C., & Soto, D. (2010). Visuospatial sequence learning without seeing. *PLoS ONE*, 5(7), e11906. <https://doi.org/10.1371/journal.pone.0011906>
- Rothkirch, M., & Hesselmann, G. (2017). What we talk about when we talk about unconscious processing - A plea for best practices. *Frontiers in Psychology*, 8(835), 1–6. <https://doi.org/10.3389/fpsyg.2017.00835>
- Sand, A., & Nilsson, M. E. (2016). Subliminal or not? Comparing null-hypothesis and Bayesian methods for testing subliminal priming. *Consciousness and Cognition*, 44, 29–40. <https://doi.org/10.1016/J.CONCOG.2016.06.012>
- 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/10.1016/j.cognition.2018.02.008>
- 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>

- Shanks, D. R. (2017). Regressive research: The pitfalls of post hoc data selection in the study of unconscious mental processes. *Psychonomic Bulletin & Review*, 24(3), 752–775.
<https://doi.org/10.3758/s13423-016-1170-y>
- Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioural and Brain Sciences*, 17, 367–447.
- Skora, L. I., Livermore, J. J. A., Dienes, Z., Seth, A. K., & Scott, R. B. (in-principle accepted). Feasibility of Unconscious Instrumental Conditioning: A Registered Replication. *Cortex*.
<https://doi.org/10.31234/OSF.IO/P9DGN>
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31(1), 137–149.
- 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>
- Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, 8(8), 1096–1101. <https://doi.org/10.1038/nn1500>
- Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review*, 23(1), 87–102.
<https://doi.org/10.3758/s13423-015-0892-6>
- 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>
- Wang, P., & Nikolić, D. (2011). An LCD monitor with sufficiently precise timing for research in vision. *Frontiers in Human Neuroscience*, 5(85), 1–10. <https://doi.org/10.3389/fnhum.2011.00085>

Appendix A.

Exclusion criteria flowchart, illustrating the hierarchical approach to the main analysis (unaware trials) in both experiments.