# Enhanced performance in perceptual decision making task without deprivation in mice using medial forebrain bundle stimulation

Antonin Verdier[1], Anna Aldanondo[1], Brice Bathellier[1*], Sophie Bagur[1*]
[1]*Institut de l'Audition, Institut Pasteur, Université de Paris, INSERM, Paris, France.*
** Equal contribution*
*Corresponding author : sophie.bagur@pasteur.fr*

Abstract (150-200 words):
Training animals on perceptual decision making tasks is essential to many fields of neuroscience. Current protocols generally use water as a reward in water-deprived animals. This can be challenging since it requires balancing animals' deprivation level with their well-being. Moreover, trial number is limited once they reach satiation. Here, we present electrical stimulation of the medial forebrain bundle (MFB) as an alternative reward in mice that avoids deprivation entirely while yielding stable motivation for thousands of trials. We conducted a comparison of MFB rewarded and water-deprived animals on two auditory discrimination tasks of increasing difficulty. MFB rewarded mice learnt the hard task faster and with higher overall accuracy levels than their water-deprived counterparts. Moreover, they performed up to 4600 trials within a session without loss of motivation. Importantly, MFB stimulation does not impact the sensory behaviour under study since psychometric parameters and response times are preserved. Finally, in accordance with the lack of deprivation, MFB mice lack signs of metabolic or behavioural stress compared to water-deprived mice (open field, home cage behaviour). Overall, using MFB stimulation as a reward is a highly promising tool for task learning since it enhances task performance whilst avoiding deprivation.

## Introduction

Probing the sensory abilities and decision-making processes of animals in order to model these faculties and link them to neural activity is central to systems neuroscience. This relies on training animals to perform a range of instrumental tasks. In mice, the most common protocols water deprive animals to motivate them to lick for small water rewards. Although this approach has enabled multiple successful studies, it suffers from certain experimental drawbacks. First, water deprivation requires careful and time-consuming monitoring of water consumption in order to avoid discomfort to animals while maintaining their motivation. Despite well-established procedures, motivation levels can be hard to control depending on the individual animal, the housing conditions or the experimental environment, leading to variable results. Finally, once animals reach satiation they stop performing, thus limiting the number of trials per session. This can impede the possibility of probing fine psychophysics [1] and curtail analysis through lack of data, particularly the ability to link noisy neural data to behaviour and perception.

Here we propose a novel reinforcement method which addresses these issues by using intracranial stimulation of the medial forebrain bundle (MFB) as a reward. The MFB courses via the lateral hypothalamus between multiple regions of the brain's reward system, most notably the ventral tegmental area and the nucleus accumbens. It is known to provoke strong self-stimulation without motor side effects [2]. Previously, it has mainly been used in order to study motivation itself but some studies have used this rewarding effect to perform spatial conditioning [3]–[5]. This suggested to us that this method could provide a reward system without the aforementioned drawbacks of satiation or prior deprivation.

Here we demonstrate that MFB stimulation can be used in head-fixed mice to flexibly train animals on demanding perceptual tasks with high and stable performance over thousands of trials per session. We compare performance and learning with that of mice trained on the same tasks using water deprivation and find accelerated learning on hard tasks and improved peak performance characterised by less impulsive behaviour. Moreover, we show that MFB stimulation reduces physiological and behavioural stress in comparison to water deprivation. Therefore can contribute to reducing the impact of experiments on animal welfare, when compatible with the underlying scientific question. Overall, this novel technique provides an easy-to-implement alternative to water deprivation with demonstrable experimental advantages.

## Results

We trained head-fixed mice on different auditory discrimination tasks using MFB intracranial stimulation as a reward. Before comparing these results with the classical water deprivation strategy, we provide an overview of the protocol using MFB stimulation (Fig. 1). First, during a short (~1 hour) surgery a bipolar electrode is stereotaxically implanted into the MFB (Fig. 1A, S1). The MFB is a long fibre tract of around 200µm in diameter and so can be targeted with a high success rate despite stereotaxic variability. In our hands, 90% (21/23) of mice responded positively to stimulation (i.e. learnt to respond to the sound). The efficacy of stimulation can easily be tested using an object preference assay before beginning head-fixated (see methods). In this study, we chose to train mice to press a custom-designed 3D printed lever in order to indicate their response (Fig. 1B). This allowed total dissociation from any water deprivation, often necessary to initiate licking in mice in head-restrained conditions. Animals were first trained to press the lever at least once during a 1.5s response window following a Go sound to receive an immediate 100ms MFB stimulation. Voltage was initially fixed at 1.5V. Once this behaviour was acquired, we calibrated the minimal voltage that elicited robust responding and used this level during subsequent training (Fig. 1C). Mice were then trained on the Go/NoGo task (Fig. 1D,E). They had both to respond to the Go sound and to restrain from pressing following a NoGo sound. False alarms were punished by a 5 to 7s timeout. Each trial was separated by a random interval during which mice had to restrain from pressing to initiate the next trial. In order to facilitate the use of this protocol, we supply on our Github page (https://github.com/AntoninVerdier/MFBasic) designs for the 3D models as well as instructions and code to implement a simple version of the task.

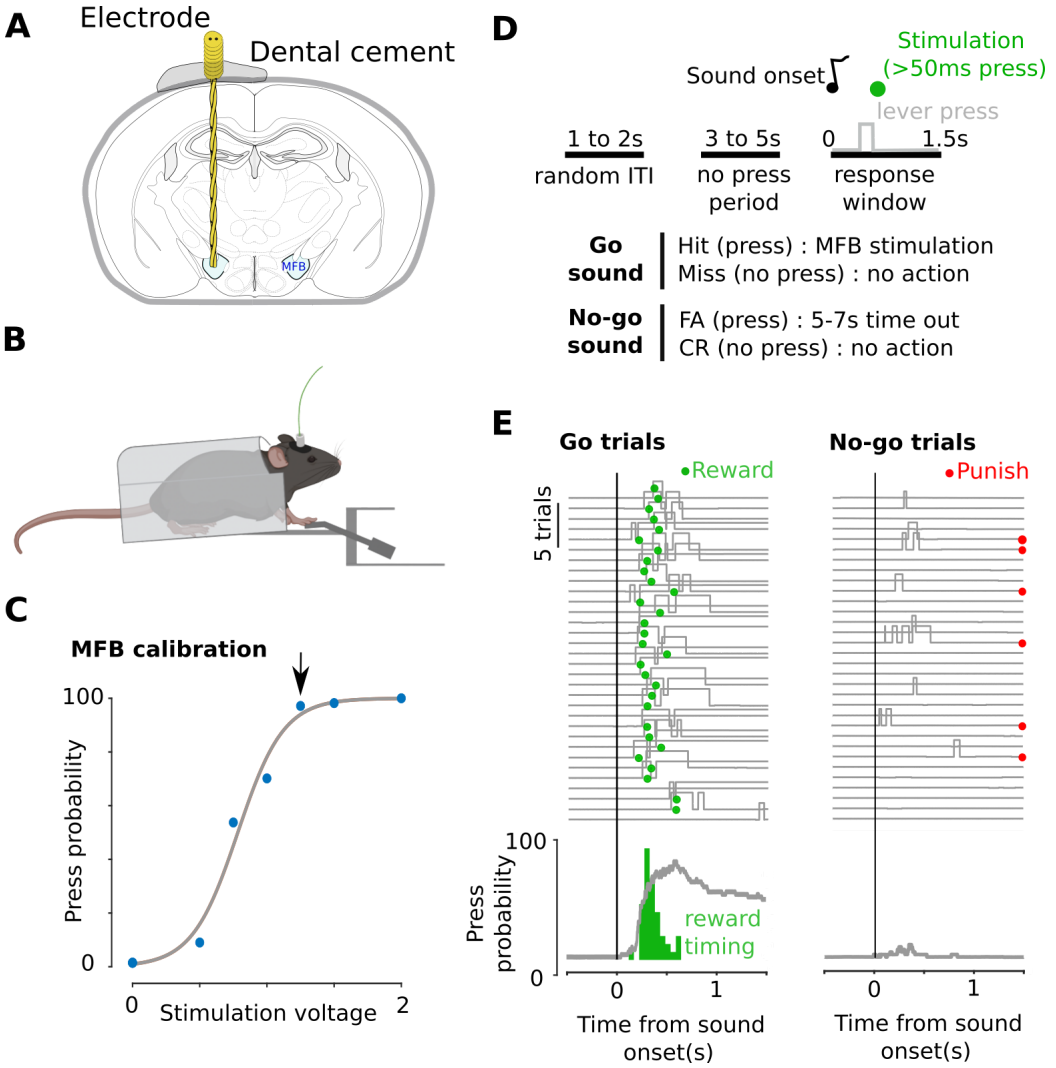


**Figure 1 Protocol for perceptual discrimination learning with MFB reward**

*A. Medial forebrain bundle implantation site of bipolar electrode fixed with dental cement to skull*

*B. Scheme of head fixed mouse during task : pressing the lever with the forepaw after Go sound elicits stimulation via the implanted electrode*

*C. Calibration curve for an example mouse. Measured probability of pressing the lever during 1.5s response window after Go sound for MFB stimulation of different voltages (blue dots) and sigmoid curve fit to data (grey line). Arrow indicates voltage used for learning.*

*D. Structure of the discrimination task*

*E. Traces of lever press for 30 Go and 30 No-go trials from an example mouse after training (top) and the mean lever press trace for all trials (bottom). For each Go trial, the time of reward (50ms after lever press) is shown by a green dot and for each No-go trial, trials punished by a time out are indicated by a red dot. The distribution of reward timings is shown in green (mean 354ms).*

MFB accelerates learning on hard tasks

In order to assess learning using the novel MFB procedure, we compared results from MFB-rewarded and water-deprived mice on identical tasks (the full protocol is presented in Fig. S2). Water-deprived mice were trained to lick for a reward instead of lever pressing but stimulus contingencies, experimenters and training apparatus were otherwise identical.

We first trained one group of mice for each reward type on a classical frequency discrimination task in order to establish frequency psychometric curves. Note that in all tasks used the valence of the sounds was counterbalanced across mice. During the first step, mice were trained to discriminate between two pure tones of low and high frequency (6kHz and 16kHz respectively). The task was considered to be learnt when accuracy was above 80% over a full session. No difference in the speed of learning was observed between the two groups, either in terms of the number of trials (Fig. 2A) or the number of sessions (Fig. 2D).

Following this first step, mice were introduced to a psychometric task where they were required to categorise 16 logarithmically-scaled sounds into low and high frequency categories (threshold frequency : 9800kHz). As in the previous step, the learning rate did not differ between the two groups in the number of trials to reach the criterion. However, the number of experimental sessions necessary was higher for the water-deprived group (Fig. 2B,E) given that MFB-rewarded mice perform more trials per session (see below). Both groups reached 90% accuracy and yielded the same psychometric curves (Fig. 2G)**.** MFB therefore did not alter the measured acoustic abilities of mice that can be probed independently of the reward type once the task has been learnt.

Finally, we asked if a harder task would reveal a difference between the two methods. We trained a sub-group of 4 mice (MFB) and a group of 8 mice (water-deprived), to discriminate between a pure tone and a frequency sweep with identical start frequency. This task is known to be more demanding [6] since mice cannot make a decision based on the salient sound onset but must integrate the sound over time in order to reach a decision. In this task, we found a marked difference between the two reward strategies: water-deprived mice showed very variable learning rates and some animals didn't reach learning criterion (75%) even after 30 training sessions whereas all MFB-rewarded mice learnt the task within 8 training sessions (Fig. 2C,F). Moreover, peak accuracy was higher in MFB-rewarded mice (Fig. 2H).

These results highlight the robustness of the MFB stimulation technique in increasingly complex tasks where water deprivation may fail to provide enough motivation for efficient learning.

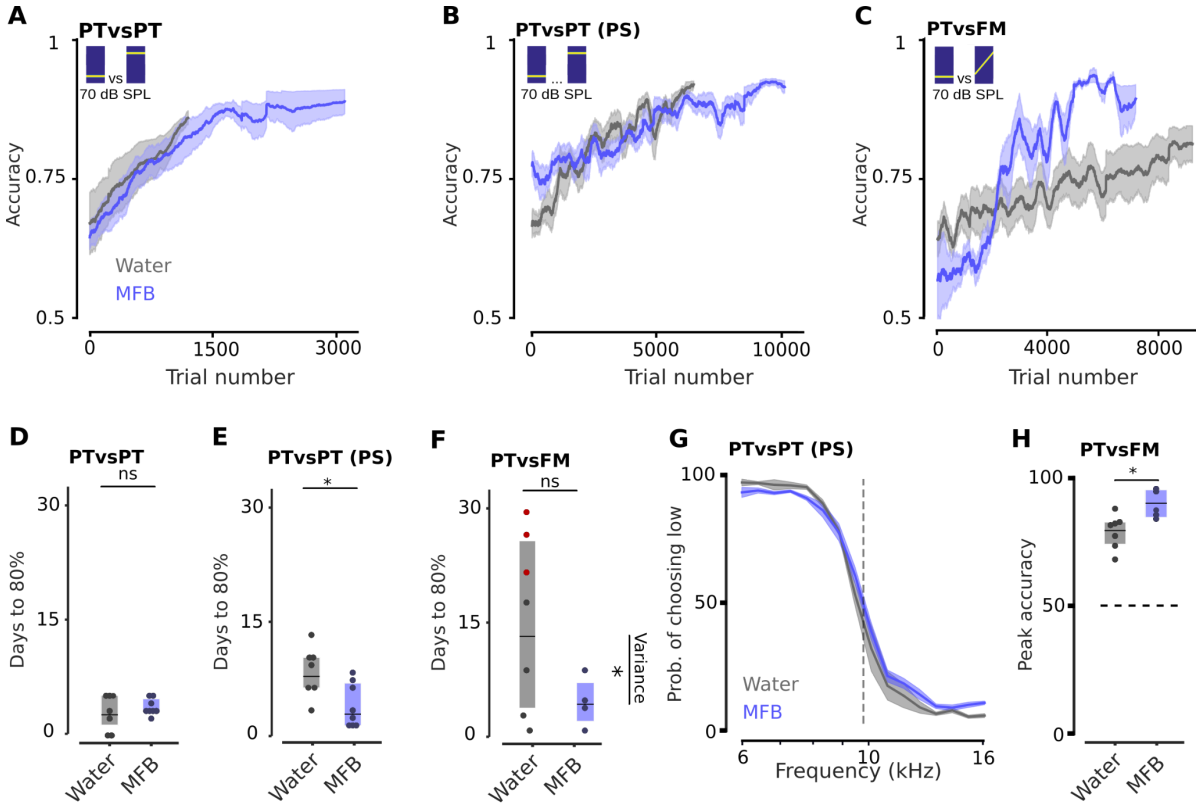


**Figure 2 MFB-rewarded mice learn tasks in less sessions and reach higher peak performance on hard tasks**

*A.B.C. Learning curves for water-deprived (grey) and MFB-rewarded (blue) mice performing the pure tone (PTvsPT) discrimination task (A - n = 7&8), the pure tone psychometric task requiring classification of high and low pure tones (B - n= 7&8) and the pure tone vs frequency modulated (PTvsFM) discrimination task (C - n = 7&4).*

*D.E.F. Number of days necessary to reach criterion accuracy for water-deprived and MFB-rewarded mice for the three tasks as in A,B,C (Wilcoxon rank sum test, pval : 0.74, 0.019, 0.29, ranksum statistic : 52.5, 75.5, 48 n=7&8,7&8,7&4). The vertical bar in F indicates that variability was higher for PT vs FM task for water-deprived mice (Levene test, pval: 0.0156, Levene statistic: 8.8). For water-deprived mice that never reached criterion in the PT vs FM task, the maximum day of training is shown in red.*

*G. Average psychometric function for pure tone frequency classification, drawn from the best 3 sessions above 80% accuracy of each mouse for the water-deprived and MFB-rewarded mice. Shaded areas indicate SEM.*

*H. Accuracy of the best session for water-deprived and MFB-rewarded mice for the PT vs FM discrimination task. (Wilcoxon rank sum test, pval : 0.02, ranksum statistic: 30 n=7&4)*

MFB yields stable performance for thousands of trials per session

Training that relies on water as a reward is intrinsically limited by the satiation of the animal: once the animal has consumed a certain amount of water they stop performing the task. This is reflected in the decline of Go accuracy during the session (Fig. 3A,B). Moreover, excessive consumption of water on a given day can have negative effects on motivation on the following days. In our hands, we therefore limited animals to 320 trials per day in order to maintain motivation during the session and throughout the week. Based on our literature search, depending on the lab, the number of trials compatible with sustained motivation can

vary between a few hundreds of trials and about 1000 trials in some cases (see table Supp.1 & Fig. 3D).

Using MFB reward, we observed during training that mice never showed signs of an irreversible drop in motivation even when performing up to 2000 trials (Fig. 3A,C) The number of trials was never limited by the animal's motivation but by experimental time constraints. We therefore investigated how many trials could be obtained in a single MFB session.

In order to compare with values from the literature, we examined sessions from a range of lengths relevant to laboratory experiments (1h30 to 3h) after learning (performance >80%) on the psychometric task. The number of trials scaled with the session duration, without loss of accuracy and reached up to 2500 trials, a far larger number than typically reported in the literature (Fig. 3D). Since sessions cannot be arbitrarily long, to increase trial yield, the only solution is to reduce inter-trial interval. We therefore measured how fast MFB-rewarded mice can perform this psychometric task while maintaining sufficient accuracy. We obtained high performance even up to 30 trials/minute, obtained with a trial window of 1500ms and total inter-trial interval (no press window + ITI) of 500ms (Fig. 3E). This is likely enabled by the high and stable motivation of the mice but also the lack of time required for reward consumption (i.e. licking to receive water). Finally, we evaluated the maximum number of trials that could be obtained in a 4h duration using a fast (2s ITI) to provide a benchmark for the novel MFB technique. We found that mice performed on average 3800 trials, with some mice achieving more than 4000 trials, with high and consistent accuracy throughout (Fig. 3F, all sessions are shown in Fig. S3).

These results show that MFB allows for well-trained mice to perform very large numbers of trials with high performance, allowing to probe within a single session many repeats of multiple stimuli properties, a key ingredient both for psychometry and linking physiology with perceptual behaviour. Moreover, although in simpler tasks learning rate was not changed, the number of days required to reach the criterion is reduced thanks to an increased number of trials per session.

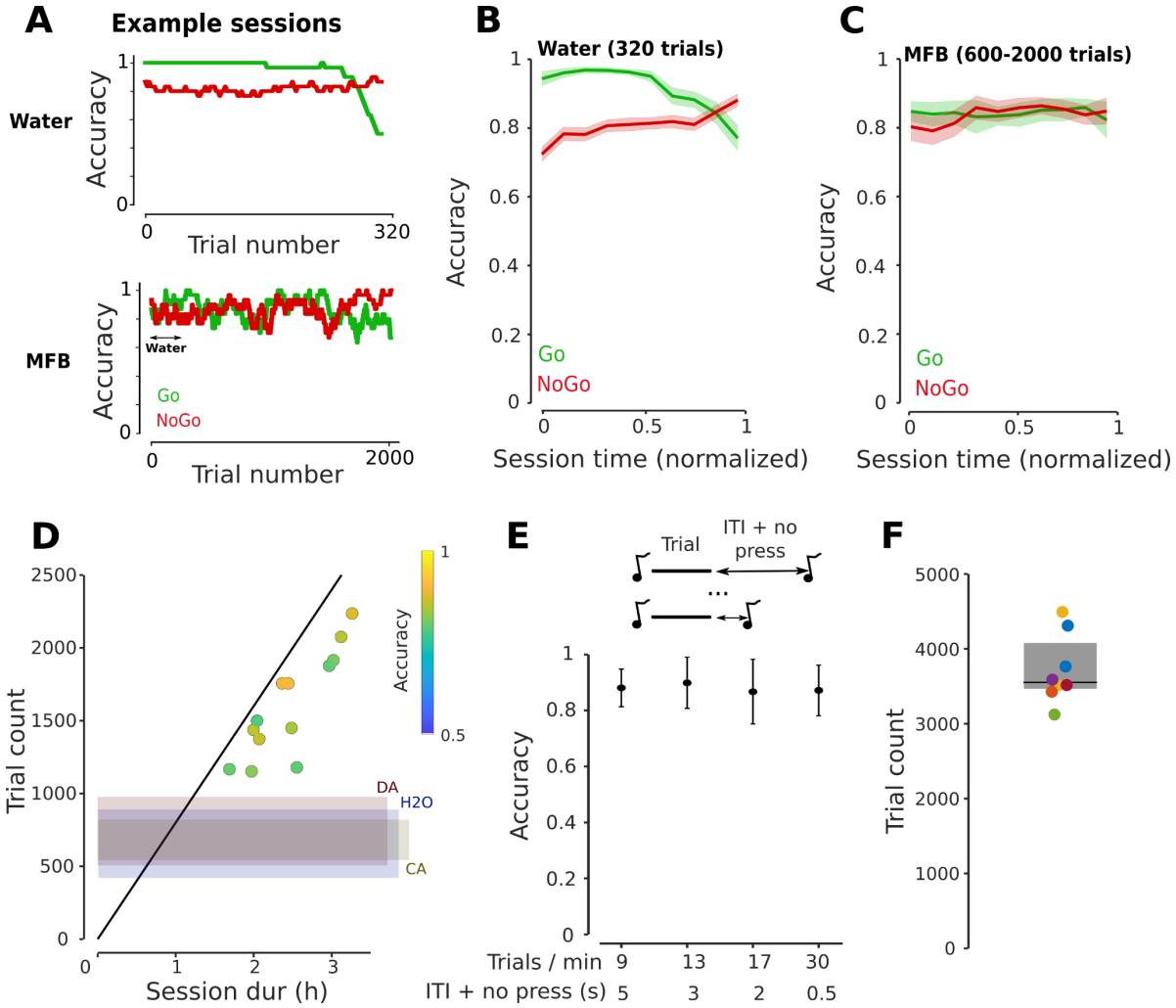


**Figure 3 MFB-rewarded mice perform thousands of trials per session**

*A. Example sessions performance on Go (green) and NoGo (red) trials for a water-deprived and a MFB-rewarded mouse. Go performance drops for the water-deprived mouse after 240 trials whereas the MFB-rewarded mouse is still performing above the criterion after 2000 thousands trials.*

*B.C. Average of Go (green) and NoGo (red) performances for water-deprived (B) and MFB-rewarded (C) mice. For MFB mice, in order to cumulate sessions of varying duration (600-2000 trials), time was normalised between 0 and 1.*

*D. Trial count of several sessions with varying trial duration, accuracy is provided by the colour of the dot. Shaded areas provide the range of trial numbers found with other strategies in the literature: optogenetic stimulation of dopaminergic fibres (DA) [7] water deprivation [8], water deprivation by adding citric acid to drinking water [9].*

*E. Accuracy over 200 trials for different trial speeds. Note that 500ms inter-trial intervals provide the same accuracy as 5000ms-spaced trials (n=4).*

*F. Trial count for 4h session with 2s ITI. Each point in one session, colour indicates mouse identity.*

MFB provides strong motivation with less agitation, impulsivity or signs of stress

So far we have demonstrated advantages of MFB-reward in terms of learning rate and more strikingly in terms of trial yield per session. However, the repeated use of intracranial stimulation could raise concerns of potential detrimental effects which we will now address.

First, we assessed during the task whether the MFB led to agitated behaviour, as might be expected given the strength of the reward and the motor role of the dopaminergic system. We quantified the level of body movement with a piezoelectric sensor placed under the animal as a proxy for agitation due to motor response and to reward. We contrasted hit and false alarm trials during which the mouse executed the same motor action to measure the increase in agitation due to the reward. Both protocols increased the amount of body movement during the response, yet to a much smaller extent with MFB stimulation (Fig. 4A,B). Both protocols also slightly increased movement after reward (Fig. 4A,B). This increase in agitation was however much smaller than that induced simply by the animal's response itself and of comparable amplitude for MFB and water-rewarded mice. To evaluate whether MFB reward was associated with more impulsive behaviour, we measured two classical indicators of impulsivity in the Go/NoGo paradigm [10]: false alarm probability, and premature responses during the inter-trial interval. Whereas water-deprived mice showed a higher false alarm than hit rate, no significant difference appeared in MFB-rewarded mice (Fig. 4C). Moreover, in the MFB animals inter-trial wait periods, which are prolonged by early responding, were shorter than in the water deprivation task (Fig 4D).

Second, we compared signs of physiological or behavioural stress. Mice in the water-deprivation protocol underwent an initial and sustained weight loss as well as a progressive loss of weight every week before the weekend when they had one day ad-libitum access to water (Fig. 5A,B). MFB-rewarded mice however showed a trend towards increasing weight as expected by ageing mice and stable weight during the week (Fig. 5A,B). We also assessed activity levels in the homecage during the night using a running wheel. MFB-rewarded mice ran for over a quarter of the time and water-deprived mice spent less time active, as previously shown (Fig. 5C) [11]. To directly measure stress levels, we performed an open field test prior to any behavioural training and after learning of all tasks. In both groups, we found a reduction in exploratory behaviour after training. However, reductions both in distance travelled and entries in the field centre were significantly stronger in water-deprived mice compared to MFB-rewarded mice (Fig. 5D). The behavioural change common to both groups may be due to the ageing of mice [12], the test-rest effect [13] or the repeated schedule of head-fixation and training although the literature suggests that head-fixation itself does not modify behaviour in the open-field [14].

Overall, we have shown that the use of the MFB stimulation does not induce any motor agitation and leads to less impulsive behaviour than the water deprivation approach. Moreover, it attenuates the impact of learning protocols on markers of stress relative to water deprivation. The novel protocol is therefore compatible with the experimental demands of well-controlled and calm behaviour and can contribute to increase animal well-being.

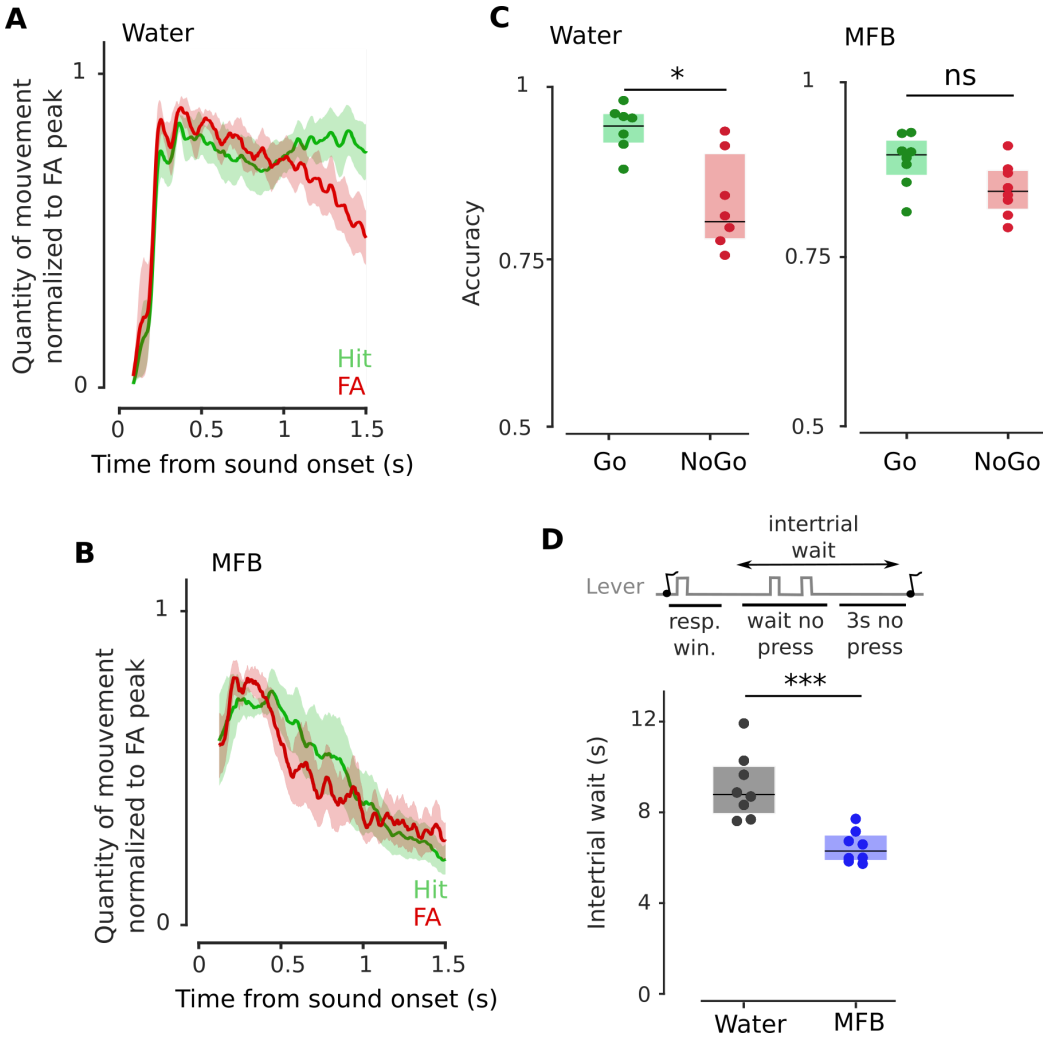


**Figure 4 MFB stimulated mice are less agitated and impulsive than water-rewarded mice**

*A.B Movement level in arbitrary units for hit and false alarm (FA) trials. The difference between the two curves is indicative of the movement induced by the reward since hit trials contain movement induced by the response (licking / lever-press) and reward (water delivery or mfb stimulation) whereas FA trials only contain movement induced by the response. The movement level for each mouse is normalised by the peak of FA movement and therefore the absolute levels of movement between mfb and water cannot be compared (see materials and methods for details, n=8 for mfb and n=4 for water, error bars are sem)*

*C. Average accuracy on Go and NoGo trials in psychometrics task during the three best sessions for water-deprived (left) and MFB-rewarded (right) mice. Note that water-deprived mice are biassed towards go-responding whereas the effect is much weaker in MFB mice (Wilcoxon signed rank test, pval = 0.04, 0.19 signed rank statistic : 26, 28; n=8,8 )*

*D. Average wait period between trials for water-deprived and MFB-rewarded mice that captures the tendency of mice to lick / lever press during the wait period. MFB-rewarded mice are significantly more restrained during the wait period (Wilcoxon rank sum test, pval : 6.2E-4, ranksum statistic : 98, n=8,8)*

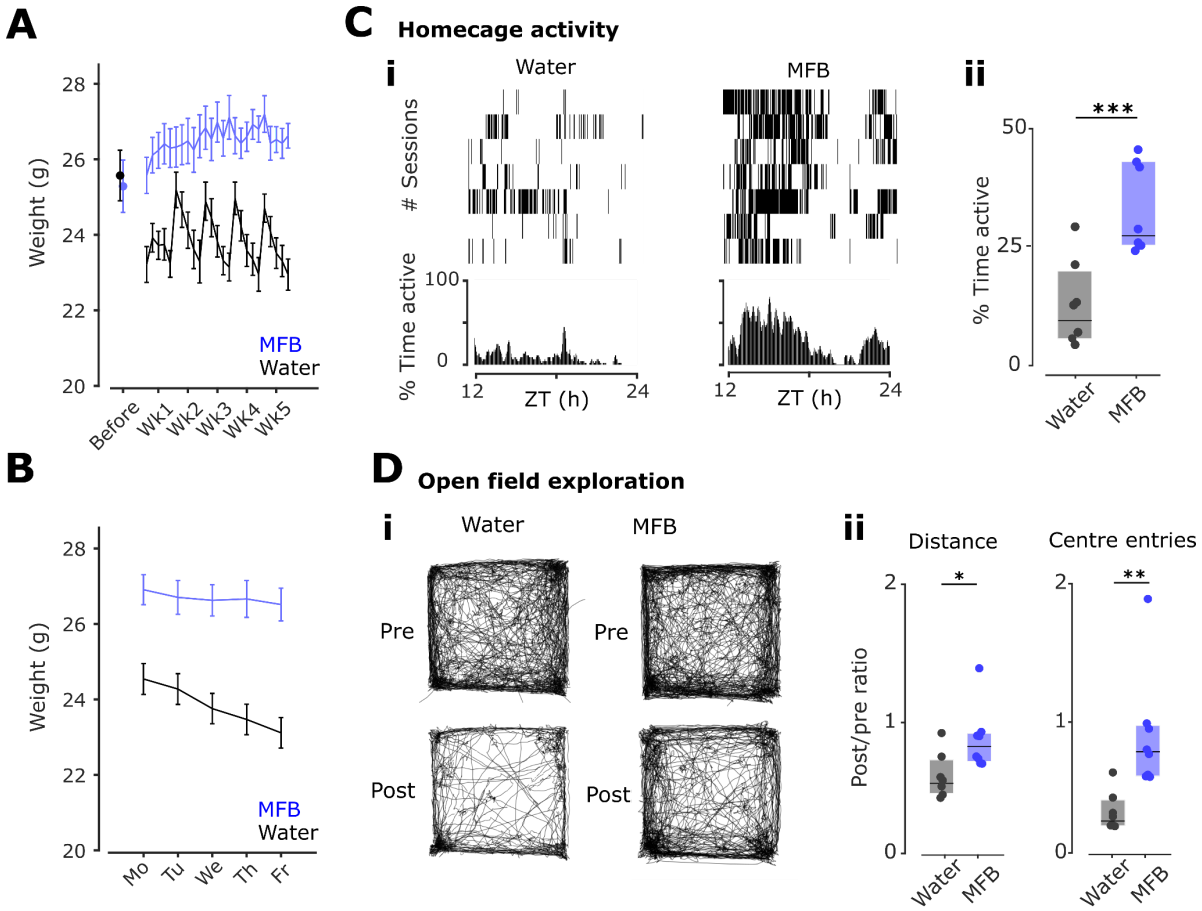


**Figure 5 Improved mouse well-being using MFB stimulation relative to water deprivation protocol**

*A. Weight evolution starting from the week prior to any behavioural experiments (isolated point) and across the first five weeks during which mice were trained five days a week. (n=8, 8 mice, error bar are sem)*

*B. Weight evolution averaged across days of the week. (n=8, 8 mice, error bar are sem)*

*C. i. Top : temporal distribution of periods of activity during the lights-off period for each session (black bar indicates a period during which wheel running was detected). Bottom : mean activity throughout the lights-off period. ii. Percent of time spent active throughout the whole night (Wilcoxon rank sum test, pval : 0.007, ranksum statistic : 73, n=7,7 sessions)*

*D. i. Superposed trajectories for mice during the open field session. ii. Ratio of total distance travelled (left) and number of entries into the centre of the field (right) before and after training during 5min exploration of the open field (Wilcoxon ranksum test, pval : 0.039, 0.028, 0.0022, ranksum statistic : 83,89, n=8,8 mice)*

Response latency and discrimination times in MFB stimulated mice

We finally assessed whether the MFB reward task produces similar sensory-motor behaviour to water-rewarded paradigms by investigating response timing. In a recent report, it was shown that discrimination time is longer for the difficult PTvsFM task than for the easier PTvsPT task [6]. We assessed with the same methodology, the earliest time at which the Go and the NoGo sound could be significantly distinguished based on the animal's response (Fig. 6A illustrates the method). Entirely consistent with previous observations established using water-deprivation, we found longer discrimination times for the PTvsFM task for both MFB-reward and water deprivation (Fig. 6). Moreover there was a

good quantitative match between the two methods despite using differences in reward and motor response (median for MFB and water : PTvsPT: 240/220ms, PTvsFM: 490/580ms).

Interestingly, as reported previously [6], longer discrimination times in the PTvsFM task were not reflected in response latencies between the two tasks for either reward contingency (Fig. 7A-D).

In the case of water deprivation, the early licking that is captured by reaction time quantification is suppressed in the simpler PTvsPT task for the NoGo sounds : mice can discriminate on the first lick. During the PTvsFM task mice elicit a similar number of early licks at sound onset for the Go and the NoGo sound, where onset is ambiguous. It is only later that they discriminate between the two sounds by suppressing licking to the NoGo sound. This is facilitated by the structure of the protocol : a permissive lick response threshold is used (NoGo trials are successful if the number of licks is less than 3-5) to ease the production of discriminative responses despite this early non-discriminatory behaviour.

In the MFB protocol, we used a more stringent response criterion and NoGo trials are considered successful only if there is no lever press. Yet, in the PTvsFM task non-discriminative early responses are still observed (Fig. 6A). Interestingly, we also found that response times were significantly shorter for false alarm NoGo trials than for Go trials (Fig. 7C,D). Based on our observations with water-deprivation, we reasoned that this difference was driven by early, non-discriminative responses perhaps associated with an excessive motivation to perform reward collection [15]–[17]. If this is the case, we would expect behavioral epochs with short lever press latencies on Go trials to be associated with low discrimination performance (i.e. high FA rates). Indeed, we found that FA rate and lever press latencies in Go trials were positively correlated (Fig. 7E-G), consistent with earlier reports showing that correct trials are associated with longer latencies [15], [17].

These results exemplify how our MFB paradigm can be easily transferred to the study of perceptual discrimination since it reproduces known results from the literature including subtler aspects of behaviour such as temporal dynamics. First, consistent with a broad literature indicating that animals have different modes of task engagement that impact both their performance and their timing [15]–[18], we have shown that these two variables covary as expected using MFB reward. Moreover we verify that the properties of the stimulus have the same impact on response timing using MFB reward and water-deprivation protocols. Therefore, our direct stimulation of the reward system appears to leave intact much of the core mechanisms underlying perception and decision making, allowing for their continued study with this new approach.

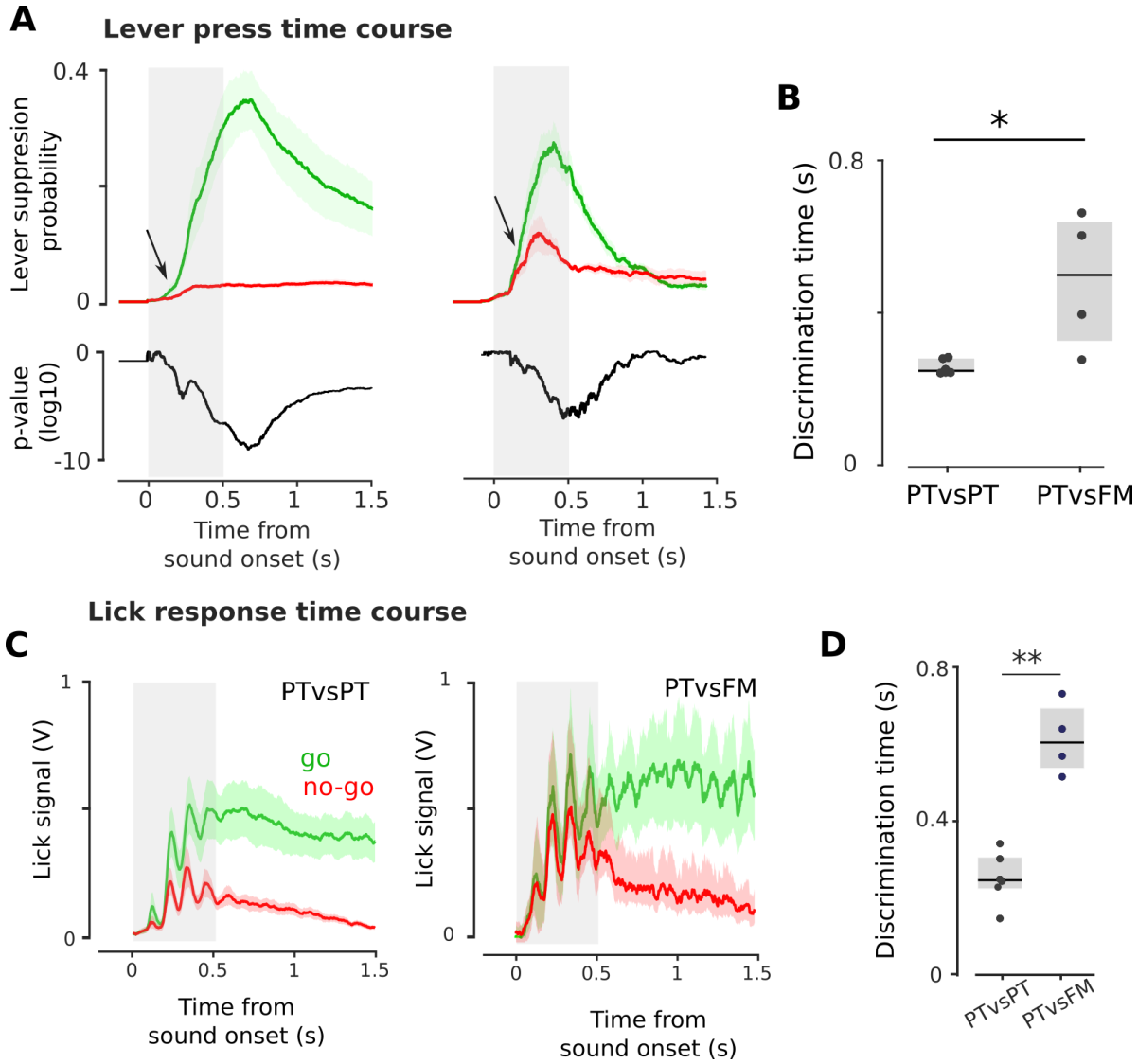


**Figure 6 Similar task dependance of discrimination times for MFB and water-reward protocols**

*A. Average lever press probability for PTvsPT and PTvsFM tasks across all go and no-go trials. Arrow shows the estimated discrimination time defined as the earliest time at which p-value of a t-test for the difference between the go and no-go traces (bottom pannel) is below 0.01. Error bars are sem.*

*B. Discrimination time as defined in A for each animal. (n=8,4 Wilcoxon rank sum test, pval = 0.03, ranksum statistic : 23)*

*C. Average licking signal for PTvsPT and PTvsFM tasks across all go and no-go trials.*

*D. Discrimination time, defined as earliest time where Go and NoGo lick signals are significantly different, for the PTvsPT and PTvsFM tasks (n=7, 4, Wilcoxon rank sum test, pval = 0.001, ranksum statistic : 23, only animals that reached critertion performance for PTvsFM task are included)*

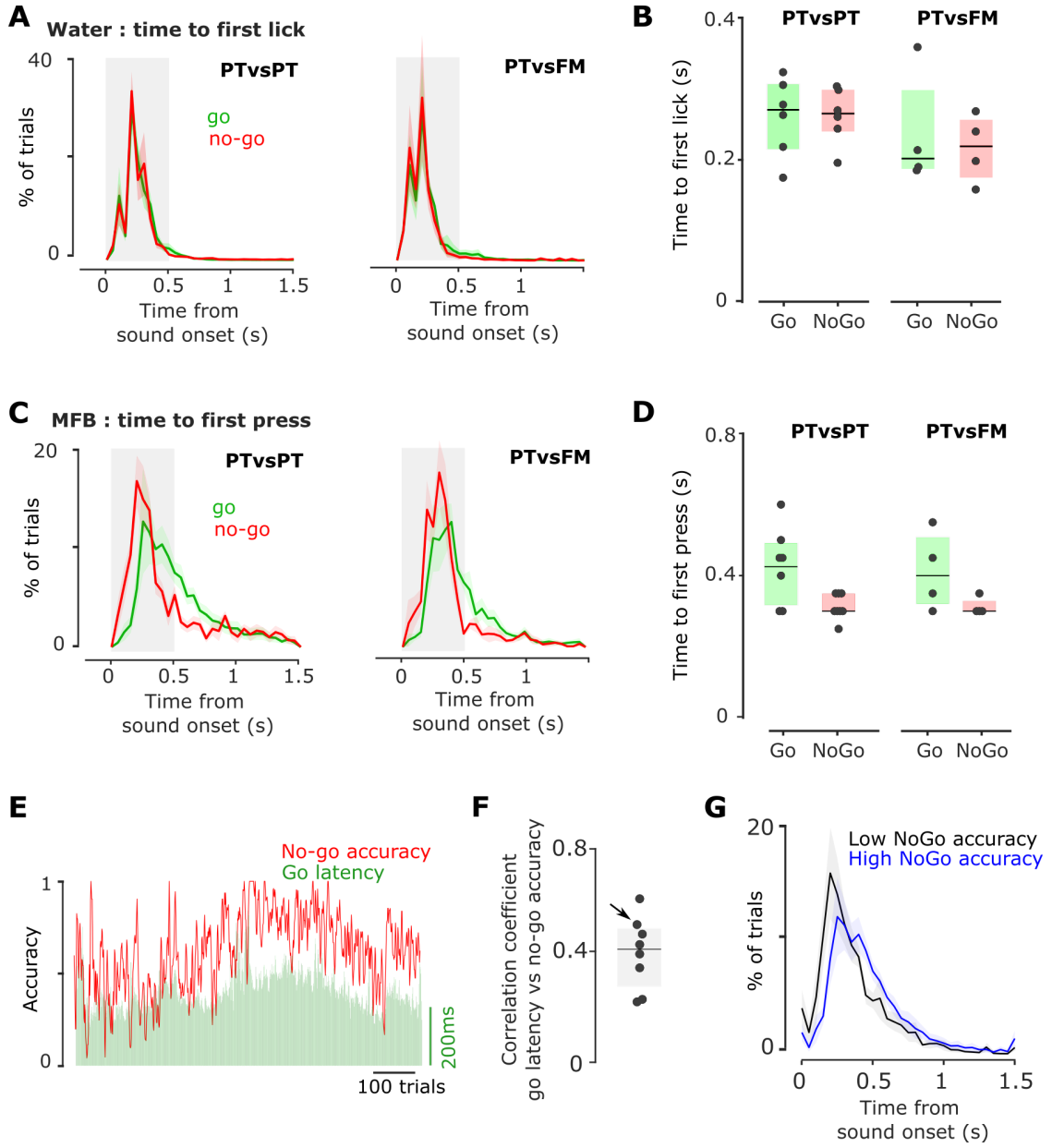


**Figure 7: Link between reaction time and task accuracy**

*A & C Histogram of reaction times (defined as time to first lick or lever press) for PTvsPT and PTvsFM tasks on correct go trials and incorrect no-go trials (MFB : n=7 & 4, Water : n=7 & 7, error bars are SEM).*

*B. D Reaction times for water and MFB groups (one mouse out of eight for PTvsPT was excluded as an outlier using the criteria of three scaled MAD from the median)*

*E. Concatenated blocks of trials for an example mouse during which go accuracy was consistently over 95% showing the correlation between fluctuations in no-go accuracy (red) and the go latency (green).*

*F. Correlation coefficient between no-go accuracy (red) and the go latency (green) for all mice. (n=8 mice, pearson correlation coefficient, all p-values<1E-4) Arrow indicates the animal shown in C.*

*G. Histogram of reaction times to Go stimulus for blocks with low NoGo accuracy (<60%, black) and high NoGo accuracy (>70%, blue) showing that, during blocks in which animals reliably pressed for NoGo trials, lever presses on Go trials occurred later.*

## Discussion

By implementing stimulation of the MFB as a reward in mice, we have shown in various auditory behavioural tasks that this approach provides a faster, more flexible and higher yield protocol for learning in comparison to water deprivation. Specifically, it requires fewer days of learning for a simple task and even provides faster learning and higher performance on a harder task. Moreover, this method allows to reliably obtain thousands trials per session without loss of motivation at up to 30 trials per minute (0.5s pre-trial period +1.5s trial duration). Additionally, MFB rewarded mice showed reduced impulsivity and signs of stress compared to their water deprived counterparts, which have previously been reported as having modified behaviour [19], [20]. Finally, studying reaction times, we observed that the behavioural dynamics in MFB rewarded tasks are similar to water deprivation.

A limitation to the present study is that water-deprived and MFB-rewarded mice were not performing exactly the same task. We elected to use a lever press and not licking for the MFB-rewarded group so as to avoid any need to water deprive the animals to initiate licking, thus demonstrating that the method can be used without any recourse to deprivation. This may impact learning rates. Indeed, lever pressing in itself requires motor learning [21] contrary to licking and could slow learning rates. Moreover, the water-deprived animals discriminated Go from NoGo stimulus by exceeding a threshold of the number of licks but in the MFB protocol any lever movement during NoGo trials counted as a mistake, thus requiring an all or none motor response. Both discrepancies limit our comparison between tasks, however based on the above we believe that it is likely to have led us to underestimate the gain in efficiency with the MFB protocol that therefore remains valid.

The strongest advantage of MFB stimulation is the lack of satiation, providing high and stable motivation throughout long sessions. Increased trials per session is a particular asset for protocols involving the study of neural activity since it allows to collect data from the same neurons whilst providing the freedom to examine neural responses to variations in stimuli or task contingency. It is therefore important to stress that MFB stimulation is entirely compatible with recordings of neural data as shown in [4] and [5]**.** Moreover, we observed no stimulation-induced agitation which is an important asset for calcium imaging. This will allow to more readily explore continuous variations in stimuli which can aid behavioural modelling [1] or higher dimensional tasks that are essential to probe neural coding [22].

MFB stimulation provides a novel methodology to train mice in conditions that are most conducive to their well-being, when compatible with the underlying scientific question. First, the approach refines current protocols by minimising the impact of the experiment on mice's daily activities since MFB mice are not deprived in any way and increases in learning speed mean that less time is spent engaged in the task. Second, our approach increases the reliability of learning and performance thus allowing for a reduction in the number of animals. Indeed, in our experiments, all MFB-rewarded mice learnt the harder pure-tone vs frequency modulated task above the criterion within a similar time window whereas water-deprived mice performances were more distributed and some mice never achieved criterion performance. Nonetheless, the major drawback of the MFB protocol is that it requires the implantation of an extra electrode during head-post surgery. In pure

behavioural experiments that do not require craniotomy, the potential stress of this additional procedure must be counterbalanced with the benefits.

A potential issue with direct stimulation of the reward system is that it is a less ethologically relevant motivation than water or food and therefore could distort both behavioural and neural activity. MFB stimulation evokes dopamine release in the Nucleus Accumbens [23] via activation of the Ventral Tegmental Area, key structures in reinforcement and motivation and 6-hydroxydopamine lesions of the MFB can impair instrumental behaviour [24]. Interestingly, during self-stimulation this dopamine release is swiftly reduced [25], [26], in a fashion reminiscent of water-reward inducing dopamine only before learning. The MFB also houses glutamatergic, monoaminergic and GABAergic fibres that extend to around 50 different brain areas that are both antidromically and orthodromically activated upon stimulation [27]. Therefore, a complex array of neural systems could participate in the effect of MFB stimulation and the extent to which this mimics the physiological reward pathways - that themselves are incompletely resolved [30] - remains to be dissected [28]. On the other hand, it is important to bear in mind that although ethologically relevant, water deprivation should not be considered as a “neutral state”. It induces metabolic [19], most obvious in weight loss, and behavioural, as we measured in the open field (Fig. 5), stress as well as brain-wide changes in neural activity [20]. In this study, we have shown that psychometric performance (Fig. 2) and response timing (Fig. 6,7) are very similar between water and MFB reward. This is consistent with observations that training but not peak performance differ depending on reward type (food vs water)[11] or lab to lab differences[8]. Therefore, for the majority of studies focused on sensory or motor behaviours and not motivation per se, MFB stimulation is unlikely to modify the results of interest.

A number of other studies have proposed refinements of the water deprivation protocol. Adding sucrose to water can increase trial number [29] and replacing the animal’s drinking water with ad libitum access to citric acid water allows for a less impactful but still efficient version of the deprivation protocol [9]. However, even with these improvements mice still undergo weight loss due to deprivation and, as reported in supplementary table 1 and shown in Fig. 3D, deprivation methods have not been reported to go beyond a thousand trials per session. Recently, Burgess et al. [7] demonstrated that optogenetic activation of dopaminergic neurons in the ventral tegmental area could be used as a very efficient reward system. This approach showed similar advantages to ours in terms of accelerating learning and enhancing trial yield. However, electrical MFB stimulation is more flexible and easier to implement. First, optogenetic stimulation requires the use of a specific CRE line making it incompatible with studies in WT or other CRE lines of interest to the researcher. Second, whereas MFB implantation is achieved with one brief surgery, the optogenetic approach requires injecting a viral vector and waiting for expression (2-3 weeks) before finally implanting the optic fibre in a second surgery.

Here we explored only a single, specific task amongst the many available. Further work would be required to expand the technique to different motor outputs and task rules. A useful extension would be to couple MFB with licking since this is a fast acquired behaviour and would allow to reuse previous protocols. This could be accomplished by initially motivating licking with a brief deprivation before switching to MFB-reward alone and returning to ad-libitum water delivery. It is likely that MFB will easily generalise given that it has already been shown to work well in freely moving animals both with nose-poke[4] and

with spatial conditioning [3], [5]. Finally, further optimisation of MFB stimulation parameters (delay, duration, etc) could lead to improved results.

In conclusion, MFB stimulation adds to the toolbox of methods available to train rodents depending on the demands of the scientific question. It is an easy-to-implement alternative to deprivation and the strong improvement in trial-yield that it provides should enable and accelerate novel experiments that require fine perceptual measurements, analysis of behavioural variability and linking single-trial behaviour to neural data.

## Materials and methods

### *Subjects and surgical procedures*

7 to 9-week old C57Bl6 male mice (Janvier Labs) were used. Mice were housed in an animal facility (08:00–20:00 light) in cages of 3 to 5 individuals with access to enrichment (running wheel, cotton bedding and wooden logs).

All experiments were performed in accordance with the French Ethical Committee (Direction Générale de la Recherche et de l'Innovation) and European legislation (2010/63/EU). Procedures were approved by the French Ministry of Education in Research after consultation with the ethical committee #59 (authorization number APAFIS#28757-2020122111048873).

Surgeries were performed in order to equip all mice with a metal head-plate for fixation. Implantations of a bipolar stimulation electrode (60-µm-diameter twisted stainless steel, PlasticsOne) in the medial forebrain bundle were performed for the MFB group. Mice were injected with buprenorphine (Vétergesic, 0,05-0,1 mg/kg) 30min prior to anaesthesia induction with isoflurane (3%) and then placed on a thermal blanket during the whole procedure with isoflurane maintained at 1-1.5%. The eyes were protected with Ocrygel (TVM Lab). Lidocaine was injected under the skin of the skull 5 minutes prior to incision.

The MFB electrode was implanted using stereotaxic coordinates (AP -1.4, ML +1.2, DV +4.8) which leads to an implantation site at the dorsal limit of the MFB (Fig. S1 for histological verification). It was then fixed along with the headplate to the skull using dental cement (Ortho-Jet, Lang).

After surgery, mice received a subcutaneous injection of sterile glucose and metacam (1mg/kg). Mice were subsequently housed for one week with metacam delivered via drinking water or dietgel (ClearH20) without any manipulation.

### *Behavioural setup*

Behavioural experiments were monitored and controlled using homemade software (Elphy, G. Sadoc, UNIC, France for water deprivation and Matlab homemade GUI for MFB) coupled to a National Instruments card (PCIe-6351). Sounds were amplified (SA1 Stereo power amp, Tucker-Davis Technologies) and delivered through high frequency loudspeakers (MF1-S, Tucker-Davis Technologies) in a pseudo-random sequence. Mice were placed on a 3D printed contention tube with two piezo-electric sensors inserted at the base to allow monitoring of movement.

*MFB stimulation:* MFB stimulation was delivered via a pulse train generator (PulsePal V2, Sanworks) that produced 2ms biphasic pulses at 50Hz for 100ms. The stimulation was delivered when the mouse pressed the lever, raising the opposite end and interrupting an infrared beam (ZHITING, infrared motor speed metre, 3.3V). The lever was weighted to fall back into its initial position as soon as the mouse removed its paw. It was designed using Autodesk Fusion 360 and produced with an Ultimaker-3 3D printer.

*Water deprivation:* Water delivery (5µl/reward) was controlled with a solenoid valve (LVM10R1-6B-1-Q, SMC). A voltage of 5V was applied through an electric circuit joining the lick tube and aluminium foil on which the mouse was sitting, so that lick events could be monitored by measuring the voltage through a series resistor in this circuit.

### *Behavioural procedures*

Fig. S2 provides an overview of the behavioural protocols.

*Sounds and task structure*: Mice were trained on four different tasks. In each case, sounds of 500ms were presented at 70dB with a sample rate of 192kHz. The identity of Go and NoGo sounds was counterbalanced across mice.
Frequency discrimination task (PTvsPT): Discrimination between a pure tone at high frequency (16kHz) and a pure tone at low frequency (6kHz). Frequency psychometrics (PTvsPT PS): Discrimination between a group of 8 high frequency pure tones (10kHz - 16kHz) and a group of 8 low frequency pure tones (6kHz - 9.5kHz) in which the 16 tones were equally spaced on a logarithmic scale. Pure tone vs frequency-modulated sound (PTvsFM): Discrimination between a pure tone at 6kHz and either a frequency sweep rising from 6 to 16kHz or a 2-step rise consisting of 250ms of 6kHz and 250ms of 16kHz.

*MFB stimulation:* Mice performed behaviour five days per week (Monday to Friday). During the entire behavioural training period, food and water were available *ad libitum*. Animal weight was monitored daily.
Habituation: During four days, mice were habituated to head fixation by progressively increasing the duration from 30min to 2 hours without reward or sound stimulation. Lever press shaping: During 1 or 2 days mice received systematic reward when they pressed the lever with a 2s minimal inter reward interval. In order to encourage lever pressing, the experimenter actively shaped the mouse's behaviour for 10-50 trials by placing its paw on the lever with a small plastic stick. During this initial phase, before calibration, stimulation started at 3V and was reduced by the experimenter if it elicited a motor reaction. Go training: Once animals were spontaneously pressing, go training was conducted. Go trials were presented with 80% probability, while the remaining trials were blank trials (no stimulus). A trial consisted of a random inter-trial interval (ITI) between 1 to 2 s to avoid prediction of stimulus appearance, a random 'no press' period between 3 and 5 s and a fixed response window of 1.5 s. Any lever press longer than 50ms was counted as a lever press. Lever press during the response window on a Go trial was scored as a 'hit' and triggered immediate MFB stimulation. No lever press was scored as a 'miss' and the next trial immediately followed. Calibration: Once reliable pressing to the sound was established (>80% hits for the Go stimulus), the voltage of the stimulation was systematically reduced from the voltage used during go training down to 0 in 0.5V increments in order to obtain a calibration curve. During all following experiments, the minimal voltage that elicited >90% pressing was used. We found that this voltage could be adjusted by 0.5V to either motivate (increase) the mouse during harder tasks or to slow unsolicited pressing (decrease) and that this could facilitate learning. Calibration was performed using the Go sound from the PTvsPT protocol since this was learnt first. Go/NoGo training: After Go training, the second sound (NoGo) was introduced. During presentation of the NoGo sound, the absence of lever press was scored as a 'correct rejection' (CR) and the next trial immediately followed. Any lever press during NoGo trials was scored as a 'false alarm' (FA), no stimulation was given, and the animal was punished with a random time-out period between 5 and 7 s. Each session contained 50% probability for each trial type. Psychometric measurement: For the frequency task, when mice reached 80% accuracy on the Go/NoGo training, the number of sounds increased to 16 (8 high frequency sounds and 8 low frequency sounds) in order to establish psychometric curves for frequency discrimination.
During all phases of the protocol, the session duration was 1h30-3h and was only limited by time constraints on the use of the apparatus.
Impact of inter-trial interval: We measured whether MFB rewarded mice could maintain their performance at high levels even with a short inter-trial interval. We performed in

succession on the same day 4 bouts of 200 Go/NoGo trials, decreasing the average total inter-trial interval (random ITI + random no press window) from an average of 5s down to 0.5s.

Maximal number of trials measurement: To provide a benchmark for the number of trials obtainable in a single session we performed on the PTvsPT task a 4 hour session with a 2s ITI. Mice were checked at regular intervals during the session for signs of stress and offered water after 2h.

Object preference assay: Before beginning training on the task, the efficacy of MFB stimulation can be rapidly tested using an object preference assay in order to eliminate mice with poor electrode placement. We performed this in freely moving mice in a neutral cage with two objects. After an initial 5 min habituation phase during which time spent at each object is timed, mice undergo 15 min conditioning during which when they approach one object they receive stimulation. During conditioning the preference is generally obvious by eye but can be quantified by timing the time spent at each object.

*Water deprivation :*

Mice performed behaviour five days per week (Monday to Friday). Water deprivation began after habituation to head fixation. Before starting the training procedure, mice were water restricted for two consecutive days. Afterwards, water restriction (0.8mL/day) was interleaved with a 16h *ad libitum* supply overnight every Friday. During the entire behavioural training period, food was available ad libitum. Animal weight was monitored daily.

Habituation : Prior to deprivation, during four days, mice were habituated to head fixation by progressively increasing the duration from 30min to 2 hours without reward or sound stimulation. Free lick: During the first day of training, mice received water by licking the lick port without any sound. Go training: After this period, Go training was conducted where Go trials were presented with 80% probability, while the remaining trials were blank trials (no stimulus). A trial consisted of a random inter-trial interval (ITI) between 3 to 0 s to avoid prediction of stimulus appearance, a random ‘no lick’ period between 1.5 and 3.5 s and a fixed response window of 1.5s. Licking during the response window on a Go trial above lick threshold (3-5 consecutive licks) was scored as a ‘hit’ and triggered immediate water delivery. Licking below the threshold was scored as a ‘miss’ and the next trial immediately followed. At the beginning of each session, 10 trials with ‘free rewards’ were given independently of the licking to motivate the mouse. Go/NoGo training : When animals reached more than 80% ‘hits’ for the Go stimulus, the second sound (NoGo) was introduced, and the lick count threshold was set between 5 and 8 licks (this parameter was adjusted at the beginning of training for each mouse depending on their behaviour). During presentation of the NoGo sound, licking below threshold was scored as a ‘correct rejection’ (CR) and the next trial immediately followed, licking above threshold on NoGo trials was scored as a ‘false alarm’ (FA), no water reward was given, and the animal was punished with a random time out period between 5 and 7 s. Each session contained 40% probability for the Go and NoGo stimuli and 20% blank trials. Psychometric measurement: For the frequency task, when mice reached 80% accuracy on the Go/NoGo training, the number of sounds increased to 16 (8 high frequency sounds and 8 low frequency sounds) in order to establish psychometric curves for frequency discrimination.

During all phases of the protocol, the session was continued until animals had received a maximum of 0.8mL or until they ceased licking in response to sound. This occurs within 300-600 trials. Throughout the protocol, we performed 320 trials per day except towards the end of training on Thursday or Friday when mice sometimes performed up to 640 trials.

Mice that did not attain 0.8mL during the task received the missing amount directly in their cage

*Open field*: Before the start of task learning and then again after 3 months of training, mice were placed at the centre of a brightly lit 50cmx50xm open field for three minutes. The surface and walls were cleaned between each mouse. The animal's position was tracked using an online homemade tracking program (Matlab, Mathworks).

*Home cage activity:* In order to measure mice's activity level in their homecage, we monitored the turning of a running wheel from 8 pm to 8 am, the dark period which was undisturbed by behavioural experiments. The measurement was made by counting infrared beam breaks due to the passage of an opaque section of the wheel. We counted as "active" any period of 5s during which the wheel made one full turn.

***Data analysis***

*Learning and performance analysis:* Learning curves were obtained by calculating the fraction of correct responses over blocks of 180 trials. Discrimination performance over one session was calculated as (hits + correct rejections)/total trials. The learning rate is defined as the number of training days to reach accuracy greater than or equal to 80%. Note that for all evaluations of learning time we only took into account valid trials left after excluding all periods during which the mouse failed to respond for more than 15 trials.

*Response time analysis:* Response times are defined as the time to first lever press / first lick. Discrimination time was defined as the earliest 10ms time bin at which a significant difference between lever pressing / licking signal for Go and NoGo trials was found. We used the non-parametric Wilcoxon rank-sum test to obtain the p-value for the null hypothesis that the Go and NoGo distributions were identical for each time bin and then applied the Benjamini-Hochberg correction for multiple testing over all time bins. In order to avoid biassing statistical significance due to differing number of trials, we performed multiple random subsampling of datasets to restrict to 100 trials and took the mean reaction time over all shuffles.

*Piezoelectric signal:* The acquired piezoelectric signal was filtered to remove 50Hz contamination. Changes in the animal's position within the tube can create baseline shifts in the signal. We therefore measured the amount of movement using the absolute value of the first-order temporal derivative. The absolute values of this signal cannot be compared between MFB-rewarded and water-deprived mice since water-deprived mice were placed on aluminium foil in order to detect licking and the deformations of the foil attenuate the values readout by the sensor.

*Statistics:* For each statistical analysis provided in the manuscript, the Kolmogorov–Smirnov normality test was first performed on the data. If the data failed to meet the normality criterion, statistics relied on non-parametric tests. We therefore represent the median and quartiles of data in boxplots in all figures, in accordance with the use of non-parametric tests. Ranksum and signed rank: we report the signed rank statistic if the number of replicates is too weak to provide the normal Z statistic

***Code and 3D printer files availability***

In order to facilitate other laboratories trying out the new methodology, we have made available on a Github page (https://github.com/AntoninVerdier/MFBasic) the 3D printer files for contention tube and lever, simplified Arduino code to implement a minimalist Go/NoGo task with MFB stimulation and a detailed description of all the hardware needed and how to set it up. For further assistance, please contact the corresponding author.

Acknowledgements
We thank Noémi Dominique and Léa Vinel for assistance with behavioural experiments.

Funding
This work was funded by the European Research Council (ERC CoG 770841), the Fondation pour l'Audition (FPA IDA02), S. Bagur was funded by the Fondation pour la Recherche Médicale (SPF202005011970, FRM). This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 964568 (Hearlight). We acknowledge the support of the Fondation pour l'Audition to the Institut de l'Audition.

Competing interests
None of the authors have any competing interests to declare

Author contributions
SB contributed the original idea of the project and designed research with AV and BB. SB performed MFB experiments and AV performed water-deprived experiments. SB and AV analysed the data and wrote the paper with comments from BB. BB provided funding and equipment for task implementation.

Bibliography


[1] M. Carandini and A. K. Churchland, ‘Probing perceptual decisions in rodents’, *Nat. Neurosci.*, vol. 16, no. 7, pp. 824–831, Jul. 2013, doi: 10.1038/nn.3410.

[2] W. A. Carlezon and E. H. Chartoff, ‘Intracranial self-stimulation (ICSS) in rodents to study the neurobiology of motivation’, *Nat. Protoc.*, vol. 2, no. 11, pp. 2987–2995, Nov. 2007, doi: 10.1038/nprot.2007.441.

[3] J. Naudé *et al.*, ‘Nicotinic receptors in the ventral tegmental area promote uncertainty-seeking’, *Nat. Neurosci.*, vol. 19, no. 3, pp. 471–478, Mar. 2016, doi: 10.1038/nn.4223.

[4] G. de Lavilléon, M. M. Lacroix, L. Rondi-Reig, and K. Benchenane, ‘Explicit memory creation during sleep demonstrates a causal role of place cells in navigation’, *Nat. Neurosci.*, vol. 18, no. 4, pp. 493–495, Apr. 2015, doi: 10.1038/nn.3970.

[5] T. Kobayashi, H. Nishijo, M. Fukuda, J. Bures, and T. Ono, ‘Task-Dependent Representations in Rat Hippocampal Place Neurons’, *J. Neurophysiol.*, vol. 78, no. 2, pp. 597–613, Aug. 1997, doi: 10.1152/jn.1997.78.2.597.

[6] S. Ceballo, Z. Piwkowska, J. Bourg, A. Daret, and B. Bathellier, ‘Targeted Cortical Manipulation of Auditory Perception’, *Neuron*, vol. 104, no. 6, pp. 1168-1179.e5, Dec. 2019, doi: 10.1016/j.neuron.2019.09.043.

[7] C. P. Burgess *et al.*, ‘High-Yield Methods for Accurate Two-Alternative Visual Psychophysics in Head-Fixed Mice’, *Cell Rep.*, vol. 20, no. 10, pp. 2513–2524, Sep. 2017, doi: 10.1016/j.celrep.2017.08.047.

[8] The International Brain Laboratory *et al.*, ‘Standardized and reproducible measurement of decision-making in mice’, Neuroscience, preprint, Jan. 2020. doi: 10.1101/2020.01.17.909838.

[9] A. E. Urai *et al.*, ‘Citric Acid Water as an Alternative to Water Restriction for High-Yield Mouse Behavior’, *eneuro*, vol. 8, no. 1, p. ENEURO.0230-20.2020, Jan. 2021, doi: 10.1523/ENEURO.0230-20.2020.

[10] C. L. Dent and A. R. Isles, ‘An Overview of Measuring Impulsive Behavior in Mice: Measuring Impulsive Behavior in Mice’, *Curr. Protoc. Mouse Biol.*, vol. 4, no. 2, pp. 35–45, Jun. 2014, doi: 10.1002/9780470942390.mo140015.

[11] P. M. Goltstein, S. Reinert, A. Glas, T. Bonhoeffer, and M. Hübener, ‘Food and water restriction lead to differential learning behaviors in a head-fixed two-choice visual discrimination task for mice’, *PLOS ONE*, vol. 13, no. 9, p. e0204066, Sep. 2018, doi: 10.1371/journal.pone.0204066.

[12] H. Shoji, K. Takao, S. Hattori, and T. Miyakawa, ‘Age-related changes in behavior in C57BL/6J mice from young adulthood to middle age’, *Mol. Brain*, vol. 9, no. 1, p. 11, Dec. 2016, doi: 10.1186/s13041-016-0191-9.

[13] L. Valek and I. Tegeder, ‘Failure of Diphtheria Toxin Model to Induce Parkinson-Like Behavior in Mice’, *Int. J. Mol. Sci.*, vol. 22, no. 17, p. 9496, Aug. 2021, doi: 10.3390/ijms22179496.

[14] K. Juczewski, J. A. Koussa, A. J. Kesner, J. O. Lee, and D. M. Lovinger, ‘Stress and behavioral correlates in the head-fixed method: stress measurements, habituation dynamics, locomotion, and motor-skill learning in mice’, *Sci. Rep.*, vol. 10, no. 1, p. 12245, Dec. 2020, doi: 10.1038/s41598-020-69132-6.

[15] A. Berditchevskaia, R. D. Cazé, and S. R. Schultz, ‘Performance in a GO/NOGO perceptual task reflects a balance between impulsive and instrumental components of behaviour’, *Sci. Rep.*, vol. 6, no. 1, p. 27389, Jun. 2016, doi: 10.1038/srep27389.

[16] Y. Niv, N. D. Daw, D. Joel, and P. Dayan, ‘Tonic dopamine: opportunity costs and the control of response vigor’, *Psychopharmacology (Berl.)*, vol. 191, no. 3, pp. 507–520, Mar. 2007, doi: 10.1007/s00213-006-0502-4.

[17] W.-K. You and S. P. Mysore, ‘Dynamics of visual perceptual decision-making in freely behaving mice’, Neuroscience, preprint, Feb. 2020. doi: 10.1101/2020.02.20.958652.

[18] Z. C. Ashwood *et al.*, ‘Mice alternate between discrete strategies during perceptual decision-making’, *Nat. Neurosci.*, vol. 25, no. 2, pp. 201–212, Feb. 2022, doi: 10.1038/s41593-021-01007-z.

[19] C. M. Bekkevold, K. L. Robertson, M. K. Reinhard, A. H. Battles, and N. E. Rowland, ‘Dehydration parameters and standards for laboratory mice’, *J. Am. Assoc. Lab. Anim. Sci. JAALAS*, vol. 52, no. 3, pp. 233–239, 2013.
[20] W. E. Allen *et al.*, ‘Thirst regulates motivated behavior through modulation of brainwide neural population dynamics’, *Science*, vol. 364, no. 6437, p. eaav3932, Apr. 2019, doi: 10.1126/science.aav3932.
[21] A. J. Peters, S. X. Chen, and T. Komiyama, ‘Emergence of reproducible spatiotemporal activity during motor learning’, *Nature*, vol. 510, no. 7504, pp. 263–267, Jun. 2014, doi: 10.1038/nature13235.
[22] P. Gao *et al.*, ‘A theory of multineuronal dimensionality, dynamics and measurement’, Neuroscience, preprint, Nov. 2017. doi: 10.1101/214262.
[23] A. C. Michael, M. Ikeda, and J. B. Justice, ‘Mechanisms contributing to the recovery of striatal releasable dopamine following MFB stimulation’, *Brain Res.*, vol. 421, no. 1–2, pp. 325–335, Sep. 1987, doi: 10.1016/0006-8993(87)91302-3.
[24] A. Smith, M. Amalric, G. Koob, and M. Zigmond, ‘Effect of Bilateral 6-Hydroxydopamine Lesions of the Medial Forebrain Bundle on Reaction Time’, *Neuropsychopharmacology*, vol. 26, no. 6, pp. 756–764, Jun. 2002, doi: 10.1016/S0893-133X(01)00420-1.
[25] M. R. Kilpatrick, M. B. Rooney, D. J. Michael, and R. M. Wightman, ‘Extracellular dopamine dynamics in rat caudate–putamen during experimenter-delivered and intracranial self-stimulation’, *Neuroscience*, vol. 96, no. 4, pp. 697–706, Mar. 2000, doi: 10.1016/S0306-4522(99)00578-3.
[26] P. A. Garris, M. Kilpatrick, M. A. Bunin, D. Michael, Q. D. Walker, and R. M. Wightman, ‘Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation’, *Nature*, vol. 398, no. 6722, pp. 67–69, Mar. 1999, doi: 10.1038/18019.
[27] R. Nieuwenhuys, L. M. G. Geeraedts, and J. G. Veening, ‘The medial forebrain bundle of the rat. I. General introduction’, *J. Comp. Neurol.*, vol. 206, no. 1, pp. 49–81, Mar. 1982, doi: 10.1002/cne.902060106.
[28] S. Ikemoto, ‘Brain reward circuitry beyond the mesolimbic dopamine system: A neurobiological theory’, *Neurosci. Biobehav. Rev.*, vol. 35, no. 2, pp. 129–150, Nov. 2010, doi: 10.1016/j.neubiorev.2010.02.001.
[29] Z. V. Guo *et al.*, ‘Procedures for Behavioral Experiments in Head-Fixed Mice’, *PLoS ONE*, vol. 9, no. 2, p. e88678, Feb. 2014, doi: 10.1371/journal.pone.0088678.
[30] J. D. Salamone and M. Correa, ‘The Mysterious Motivational Functions of Mesolimbic Dopamine’, *Neuron*, vol. 76, no. 3, pp. 470–485, Nov. 2012, doi: 10.1016/j.neuron.2012.10.021.

# Supplementary materials

| *Article* | *Reward* | *Action* | *Sensory modality* | *Task* | *Trial per session* |
|---|---|---|---|---|---|
| *Jaramillo et al, 2014* | *Water* | *Nose poke in freely moving mouse* | *Auditory* | *TWFC with change in stimulus / action contingency* | *median : 600 (300-1000)* |
| *Guo et al, 2014* | *Water with sucrose* | *Licking in head fixed mouse* | *Somatosensory* | *TWFC* | *400-600* |
| *et al, 2018* | *Soymilk (food deprivation)* | *Licking in head fixed mouse* | *Visual* | *TWFC* | *250* |
| *Urai et al, 2021* | *Citric acid water* | *Wheel movement in freely moving mouse* | *Visual* | *TWFC* | *750* |
| *et al, 2016* | *Water reward and air puff punishment* | *Licking in head fixed mouse* | *Visual* | *Go/NoGo* | *100-300* |
| *Poort et al, 2015* | *Water* | *Likcing in head fixed mouse* | *Visual* | *Go/NoGo* | *mean : 199 (31–342)* |
| *Burgess et al, 2017* | *Optogenetic stimulation of dopaminergic neurons in VTA* | *Wheel movement in head fixed mouse* | *Visual* | *TWFC* | *median : 800, likely up to 1000* |
| *The International Brain Laboratory et al, 2020* | *Water* | *Wheel movement in head fixed mouse* | *Visual* | *TWFC* | *median : 700, up to 1000* |
| *Birtalan et al, 2020* | *Water (fully automated)* | *Nose poke in freely moving mouse* | *Visual* | *5-choice serial reaction time task* | *mean : 753* |

**Supplementary table 1**

*Comparison of reward systems for training in the mouse with number of trials per session*

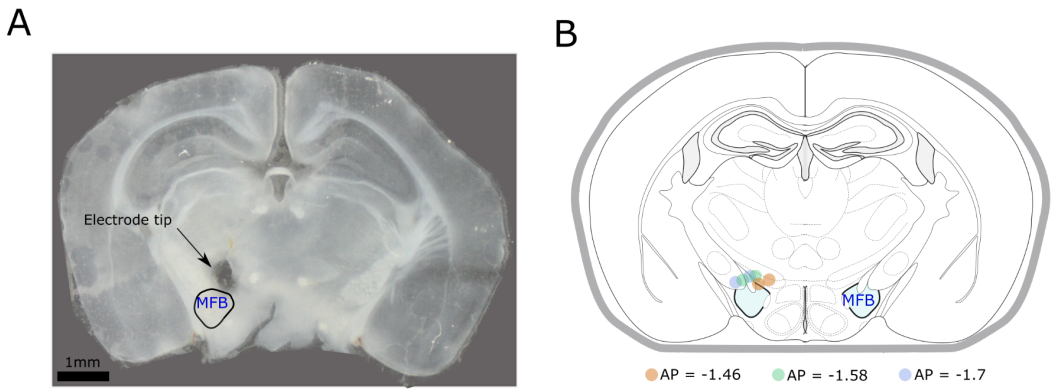


**Figure S1 MFB stimulation electrode localisation**

*A. Post-hoc localisation of electrode tip just above the MFB from an example mouse. B. Localisation of electrode tips from multiple mice.*

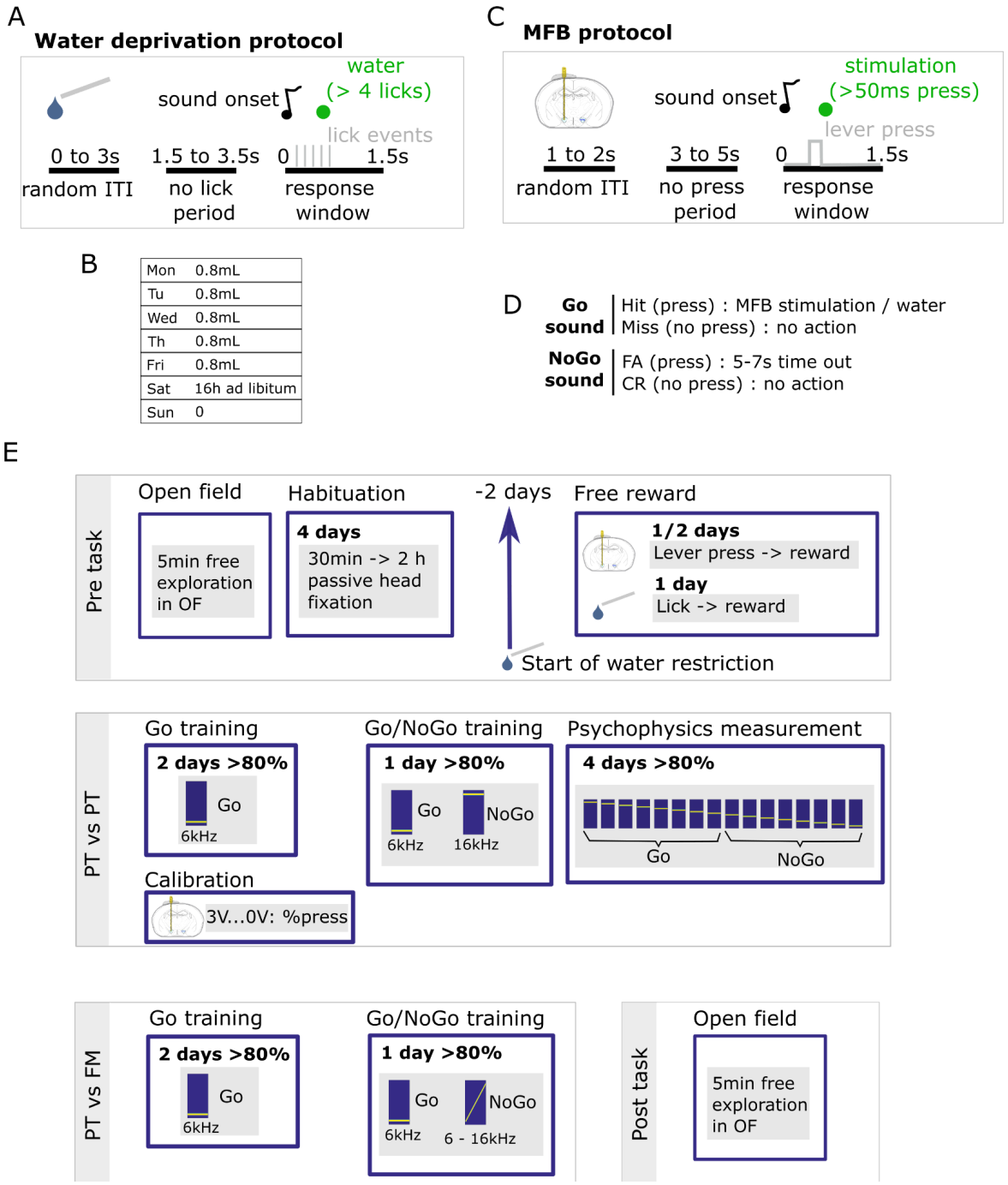


**Figure S2 Behavioural protocols**

*A. Timing of events in discrimination task for water deprived animals.*
*B. Water restriction regimen.*
*C. Timing of events in discrimination task for MFB rewarded animals.*
*D. Definition of trial types for both reward types.*
*E. Steps of the behavioural protocol outlining the succession of tasks. All information applies to both MFB-rewarded and water-deprived animals unless indicated by the corresponding pictogram*

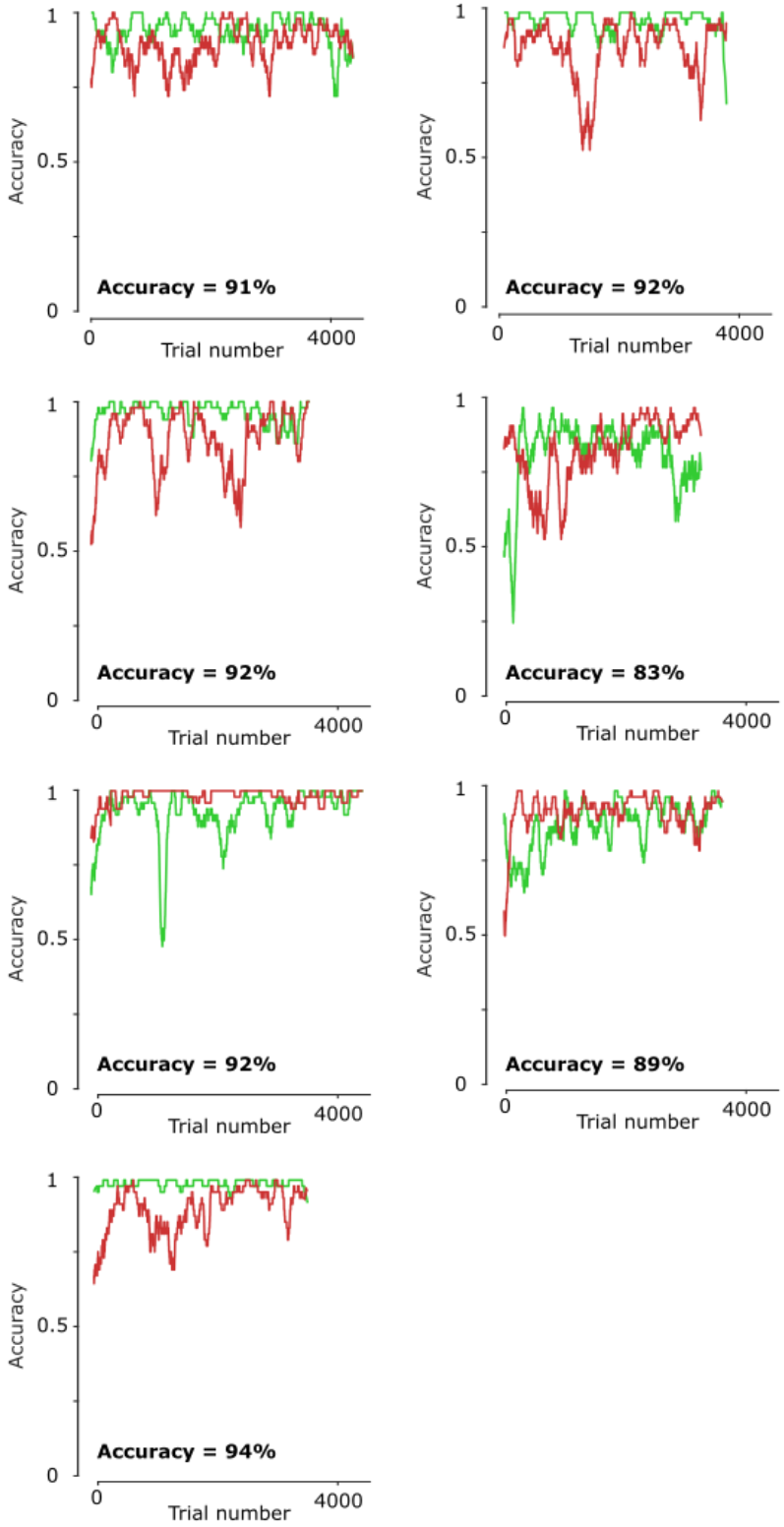


**Figure S3 Full sessions for all mice during 4h sessions with fast ITI**