Universiteit Leiden

Psychologie
Faculteit der Sociale Wetenschappen

Explaining differences in learning speed in the decision-making task of the International Brain laboratory by light cycle and training time consistency

Name: Félice Wulfse
Student number: s1782908
Date: 01/03/2023
Supervisor: Dr. A.E. Urai
Second reader: Dr. A. Ghosh
Word count: 8875
Cognitive Psychology
Thesis Msci Applied Cognitive Psychology

Layman's abstract

The International Brain Laboratory created a decision-making task to investigate learning in mice. During their experiments they found that while all mice were able to learn the task, some mice learned it more quickly than others. This study aimed to investigate the factors that influence the learning speed. We expected that mice that were trained in the dark would learn the task faster than mice that were trained in light conditions, because mice are generally active during the night. On top of that, we expected that learning speed could be influenced by whether a mouse was trained on the same time or on different times each training. The results suggested that the combination of these two variables may have an impact on how fast mice learn the task, but further research is needed to confirm this conclusion.

Abstract

The International Brain Laboratory created a standardized and reproducible decision-making task set up, and trained over hundred mice across seven laboratories in three countries. They discovered that almost all mice are trainable and performance increases in a similar pattern, but learning speed was variable across mice and across laboratories. This study aimed to investigate the differences in learning speed between mice and laboratories and explore the impact of the light cycle and training time consistency on these differences. Our hypothesis was that mice trained in an inverted light cycle (dark during the day and light during the night) would have higher learning speed than those trained in a non-inverted light cycle (light during the day and dark during the night) due to their nocturnal nature. Additionally, we believed that training time consistency would also influence learning speed. The results from our mixture analysis indicated that there may be an interaction effect between the light cycle and training time consistency, but this appears to vary depending on how you quantify learning speed and training time consistency. Further research should focus on the quantifications of the variables mentioned above and additional analyses.

Keywords: International Brain Laboratory, circadian rhythm, decision-making task, learning speed, light cycle, training time consistency.

Explaining differences in learning speed in the decision-making task of the International Brain Laboratory by light cycle and training time consistency

The International Brain Laboratory is a collaboration that aims to increase reproducibility in systems neuroscience. They have created a standardized and reproducible decision-making task set up, and trained over hundred mice across seven laboratories in three countries. They discovered that nearly all mice are trainable and performance increases in a similar pattern, but learning speed was variable across mice and across laboratories (The International Brain Laboratory et al., 2021). In this study we aim to explain those differences by two types of variables. On the one hand, there are laboratory-specific variables that might influence the learning speed. Examples are the type of food the mice eat, the building they live in or the use of an inverted versus a non-inverted light cycle. Then there are experimenter specific variables, as in some labs multiple experimenters are working together. These are for example the time the experimenters run their experiments. By exploring the impact of these variables on learning speed, we aim to further understand the role of the circadian rhythm on learning speed.

Effect of circadian rhythm on learning

In humans, the homeostatic sleep pressure and the circadian pacemaker together define the circadian rhythm. The homeostatic sleep pressure promotes sleep and causes fatigue based on the duration of wakefulness, resulting in decreased cognitive alertness. During sleep, the homeostatic sleep pressure decreases. On the other hand, the circadian pacemaker peaks in the early morning and is based on the time-of-day instead of the sleep-wake state (Schmidt et al., 2007). This shift in day-night rhythm affects behaviour in most organism (Daan, 2000) and impacts the performance of mammals in learning tasks (Fisk et al., 2018; Valentinuzzi et al., 2004). Human neurobehavioral efficiency may vary during the day because of the increasing homeostatic sleep pressure, the state of the circadian pacemaker, or the non-linear interaction between these two factors (Schmidt et al., 2007). This can affect various cognitive domains including attention, working memory, inhibition control and decision making (Li et al., 2020) and may explain some of the variation in learning speed observed in the mice in the decision-making task of The International Brain Laboratory et al. (2021). Baddeley et al. (1970) found that memory is affected by the time-of-day in humans in a simple memory task, with better immediate recall in the morning and no effect on a repeated item task. However, Elghoul et al. (2014) found that 9- to 10-year-old boys performed better on a psychomotor task (dart throwing) during the evening than during the morning (Elghoul et al., 2014). Similarly, Blake et al. (1967) found improved performance on five cognitive tasks during the day, when corrected for a practise effect using a Latin square design. In this study, a large group of participants performed eight different cognitive tasks, such as card sorting and calculations, from 8 am until 9 am. On top of that, Li et al. (2020) studied the effect of time-of-day on a risky decision-making task. Participants took more risks and had lower inhibitory control in the afternoon than in the morning (Li et al., 2020).

Mice compared to human

The International Brain Laboratory et al. (2021), among many other studies, used mice to model humans (The International Brain Laboratory et al., 2021); Guo et al., 2014; Radetsky et al., 2013). The mouse brain has become a widely used model for studying mammalian neural circuits because it provides the ability to gather both neural activity and behaviour data, which is particularly functional to behavioural research (Guo et al., 2014). Besides, the mouse brain resembles the brain of humans and is therefore a valid tool to investigate some aspects of human neural function (Nakajima et al., 2021). However, before we can extend the results of this study to humans, it is important to consider the differences in circadian rhythm between mice and humans. Mice are nocturnal, meaning their active phase is during the night (LeGates & Altimus, 2011; Ripperger et al., 2011; Roedel et al., 2006), whereas humans generally live during the day (Radetsky, Rea, Bierman & Figueiro, 2013). Despite this difference, diurnal and nocturnal species have a similar circadian rhythm of approximately 24 hours, which can be equally disrupted by changes shifts in light-dark patterns (Radetsky, Rea, Bierman & Figueiro, 2013). Thus, experimenters that use rodents to model humans must determine the light-dark cycle to be used in their experiments.

Using a non-inverted light cycle, i.e., light during the day and dark during the night, does not align with the nocturnal rhythm of mice. Roedel et al. (2006) investigated the cognitive state of DBA mice (a widely used inbred strain that is valuable in a large number of research areas) in the light phase and found a behavioural inhibition and cognitive disruption, with poorer performance in the task compared to the dark phase. Therefore, it is more appropriate to test mice during their active phase and using an inverted light cycle, with dark during the day and light during the night, potentially offers a way to handle this (Beeler & Prendergast, 2006). In the experiments of the International Brain Laboratory et al. (2021) an inverted light cycle was used in three of seven labs (Angelakilab, Wittenlab, Churchlandlab). The other four labs (Cortex, Danlab, Mainenlab, Zadorlab) used a non-inverted light cycle. This allows for the examination of the effect of the light cycle on learning speed in the decision-making task.

Training time consistency

The studies previously discussed have primarily focused on the impact of time-of-day on overall performance in cognitive tasks. However, in the current study, which aims to examine the effect of training time consistency on learning speed, the direct application of these findings is limited. This is because learning occurs gradually over time and not within a single session. We assume that training at a consistent time-of-day each session may lead to increased learning speed when the testing moment is at the same time-of-day, as training and testing conditions are kept similar (Holloway & Wansley, 1973). However, it is difficult to predict whether we can indeed expect a preference for consistent training times in this study. Furthermore, there are large individual differences between humans in the effect of time-of-day on the different cognitive domains due to morning and evening preferences (Hines, 2002; Schmidt et al., 2007; Kerkhof, 1985). This variability was confirmed by Ingram et al. (2016) in a study of morning and evening people using a decision-making task. They found that the effect of the time-of-day could be explained by phase differences in oscillating clock genes, which explains the individual differences in decision-making (Ingram et al., 2016). We expect to find the same inter-individual variability among the mice in this study.

The psychological and physiological fluctuations during the day influence our neurobehavioral efficiency in various ways. Given that learning in the decision-making task happens over sessions, we investigated the consistency in training time, as well as the effect of the light cycle used. All these variables can have an effect on the training and learning speed of the mice. There will be lots of unexplained variance, because we cannot have knowledge of e.g., an individual animal's time preference or their interaction with the experimenter. This study aims to determine whether some of the variability in learning speed across mice and labs can be attributed to the different light cycles used and different training time consistency.

Methods

The purpose of this study is to examine the impact of light cycle and training time consistency on learning speed in the experiments of the International Brain Laboratory et al. (2021). The methodology of the research involves the description of the decision-making task, the quantification of learning speed and training time consistency and the utilisation of appropriate techniques to assess the effects.

Decision making task

We shortly describe the phases of the standardized and reproducible measurement of the decision-making task from the International Brain laboratory. An elaborate description of the task can be found in the paper of The International Brain Laboratory et al. (2021). All mice underwent surgery in which a head bar was attached for head-fixation. After surgery, the subjects had a recovery phase. When recovery was over, the mice were set on water control (Urai et al., 2021) and they got habituated to the experiment environment. They started learning the basic task, in which they had to move a wheel to shift the stimulus (a Gabor patch presented on either the left or the right side of the screen) to the centre. In the basic task the probability of the stimulus appearing on the left or on the right was equal (0.5/0.5). When they were trained on the basic task, they learned the full task in which the probability of the stimulus appearing on the left or the right switched between 0.2/0.8 and 0.8/0.2. In both the basic and the full task, the contrast changes ranging from -100% to 100% over the trials. Less contrast makes the task more difficult, as it is less visible. Session duration was not fixed but determined by the experimenter. It depended on the performance, but was always stopped after 90 minutes. There were large differences in the number of trials per sessions. Therefore, some mice could have had more trials per session and learn the basic task in fewer sessions (The International Brain Laboratory et al., 2021).

Design

The data is collected by the seven labs across three countries. For this study, we further investigate the data. The data was downloaded through https://anne-urai.github.io/lab_wiki/IBLdata_behav.html.

Participants

For this study the data of 124 mice was analysed. The mice were trained across seven laboratories in total. In all of the experiments of the International Brain Laboratory et al. (2021), 197 mice were trained across nine laboratories. We had to exclude two labs because they used a mix of the type of light cycle, which goes beyond this study.

Measures

The seven laboratories have collected enormous amounts of data during the studies. We used two data frames for this study: session information (including the subject id, session start time, the performance in that session and the training status) and subject information (including the subject id, the laboratory name and the light cycle). A list of all the variables included in these two data frames can be found in Appendix 1.

Procedure

The goal of the research is to investigate what determines the differences in learning speed. We investigated the differences across the mice (their training time consistency/ their learning speed) and the differences in learning speed between laboratories (inverted versus non-inverted lights cycles). I performed a blinded analysis on the effect of laboratories using an inverted versus non-inverted light cycle. We did not want our hypothesis to influence the results. In the data frame, we created a variable (*Light_cycle)* that indicated whether a lab used an inverted or a non-inverted light cycle, and we had randomly attributed one of both conditions to the laboratories. After we agreed on the analyses, I received the true conditions for each lab.

Code availability

All analyses are done in the integrated development environment Spyder using Python programming language. The code is available at: https://github.com/Felicewulfse/int_brain_lab/blob/main/Analysis_script.py.

Ethics

The procedures and experiments of the International Brain Laboratory are carried out in accordance with the laws and approval of the countries and involved institutions. The data is available and can be accessed via their website (The International Brain Laboratory et al., 2021).

Quantifying learning speed and training time consistency

Learning speed

Before we could investigate what causes the differences in learning speed between the labs, we had to quantify learning speed. We developed four definitions which are outlined below:

1. Learning speed 1 = $\frac{performance_{last} - performance_{first}}{num\_sessions}$
2. Learning speed 2 = $\frac{performance_{max} - performance_{min}}{num\_sessions}$
3. Sessions to trained = *num_sessions until the subject is trained*
4. Sessions to eighty = *num_sessions until 80% of performance is reached*

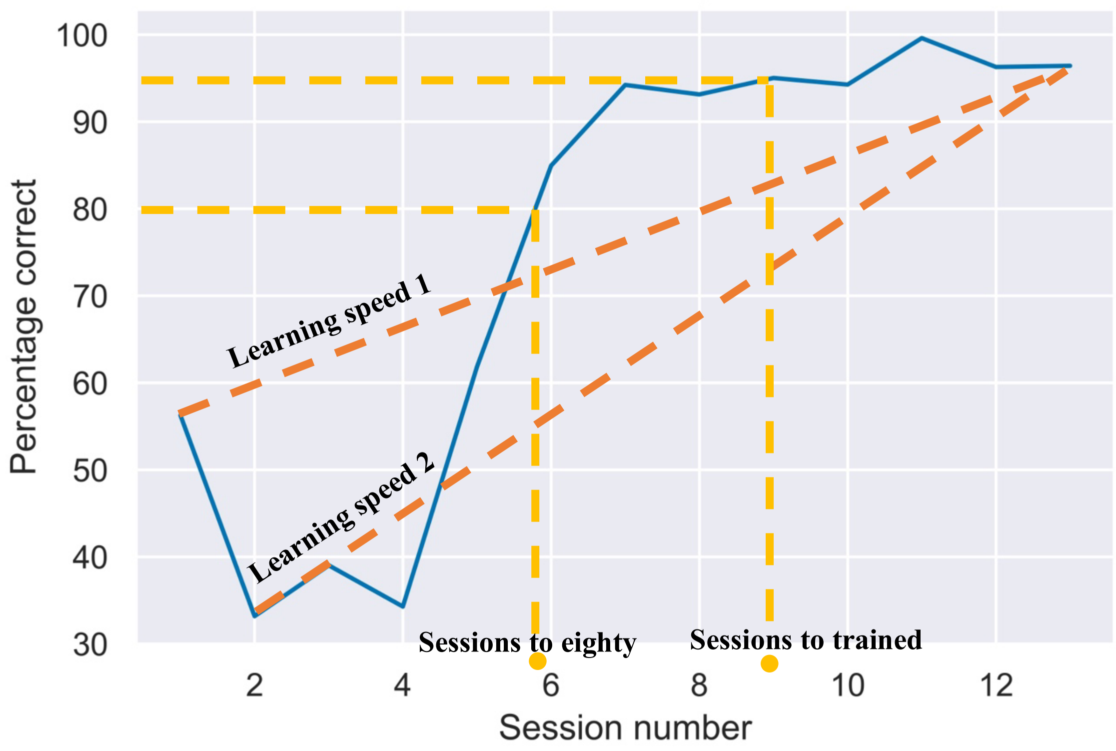
The first definition, learning speed 1, was computed as the performance change from the first to the last session divided by the number of sessions. This metric represents the average performance increase per session, and therefore might be a good indicator of learning speed. The second definition, learning speed 2, is again the performance change divided by the number of sessions, but now the performance change is computed as the lowest performance per subject subtracted from highest performance per subject. This measure serves as an alternative to learning speed 1, which may not accurately reflect the performance change if the mouse's last session performance was not its best. For both learning speed 1 and learning speed 2, a higher value indicates a higher learning speed. However, these definitions both rely on number of sessions, which varied across laboratories and could be due to lab-specific or experimenter-specific factors, rather than the mice's abilities. The total number of sessions is usually based on when the subjects achieved proficiency in the basic task, but in some mice the training continued to obtain even higher performance (IBL et al., 2021).

To address this limitation, two additional definitions for learning speed were computed that are not dependent on the total number of sessions. The International Brain Laboratory et al., (2021) used the number of training days as a measure of learning speed. In this study we use something similar: the number of sessions until a subject is trained. A subject was considered trained when the level of 'basic task proficiency' was reached. The criteria for this proficiency are present in appendix 1 part D of The International Brain Laboratory et al. (2021). The final definition for learning speed, sessions to eighty, defines the number of sessions until 80% of the maximum performance per subject is reached. For both sessions to trained and sessions to eighty a lower number of sessions is related to higher learning speed.

It is important to note that for learning speed 1 and learning speed 2 a higher value is related to a higher learning speed, while for sessions to trained and sessions to eighty a lower value is related to a higher learning speed. In Figure 1, we presented the four different learning speed variables for Example mouse 2.

Figure 1

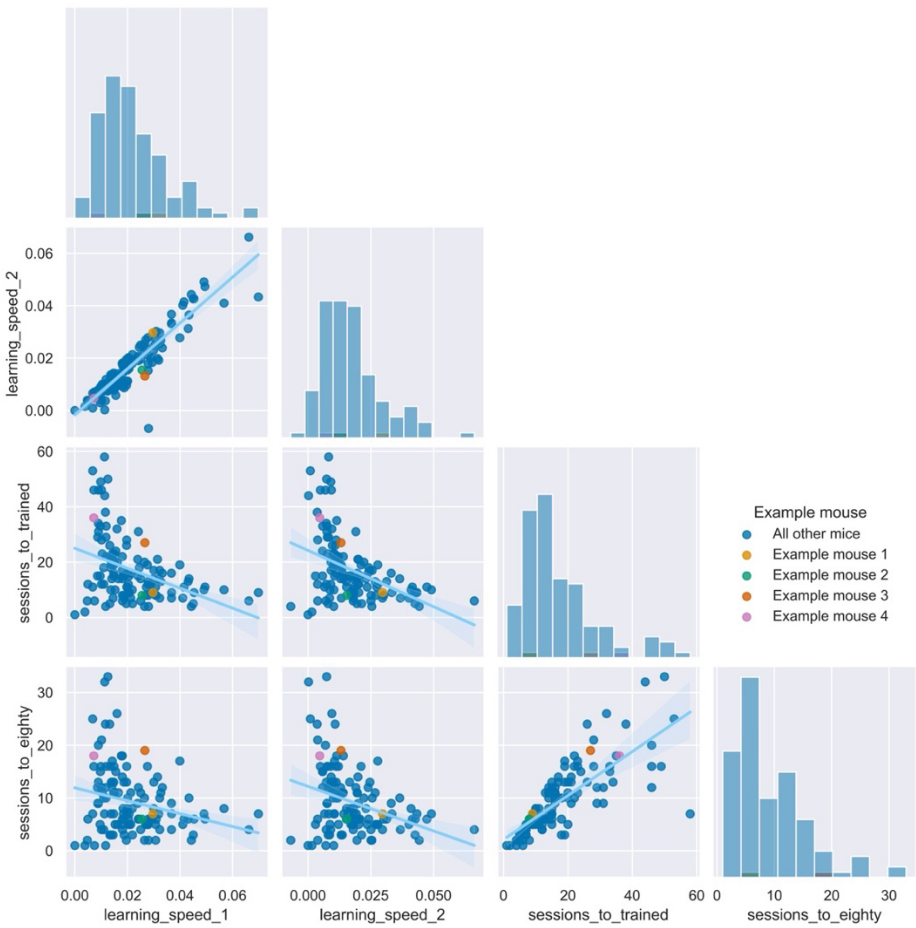
The different learning speed variables for Example mouse 2



To determine the correlation between the four different definitions of learning speed, we plotted a pair plot of these definitions in Figure 2. The results indicate that learning speed 1 and learning speed 2 are strongly correlated with each other ($r = 0.92$), while sessions to trained and sessions to eighty also show high correlation with each other ($r = 0.81$). However, learning speed 1 and learning speed 2 do not correlate with sessions to trained and sessions to eighty ($r = -0.46$, $r = -0.34$, $r = -0.47$, $r = -0.39$, respectively). These negative correlations are logical, as a larger number of sessions to reach basic task proficiency would result in lower learning speed. However, these negative correlations are not particularly strong, so caution should be taken when drawing conclusions from these definitions of learning speed. In Figure 3, the different learning speed variables for each lab are presented. As anticipated from the lack of strong correlations between the four variables, a similar pattern of high- and low-performing labs is not present across the different learning speed variables. However, there is some similarity between learning speed 1 and learning speed 2, and also between sessions to trained and sessions to eighty.

Figure 2

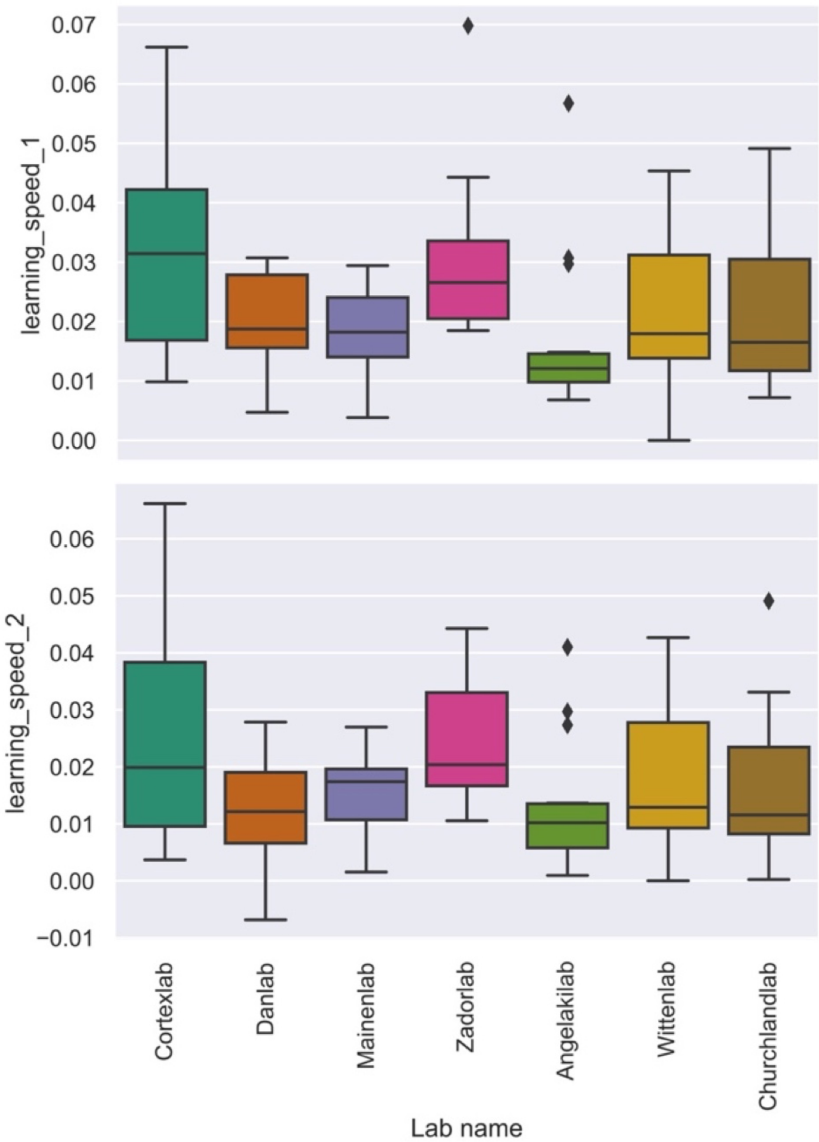
Correlations between the four learning speed variables

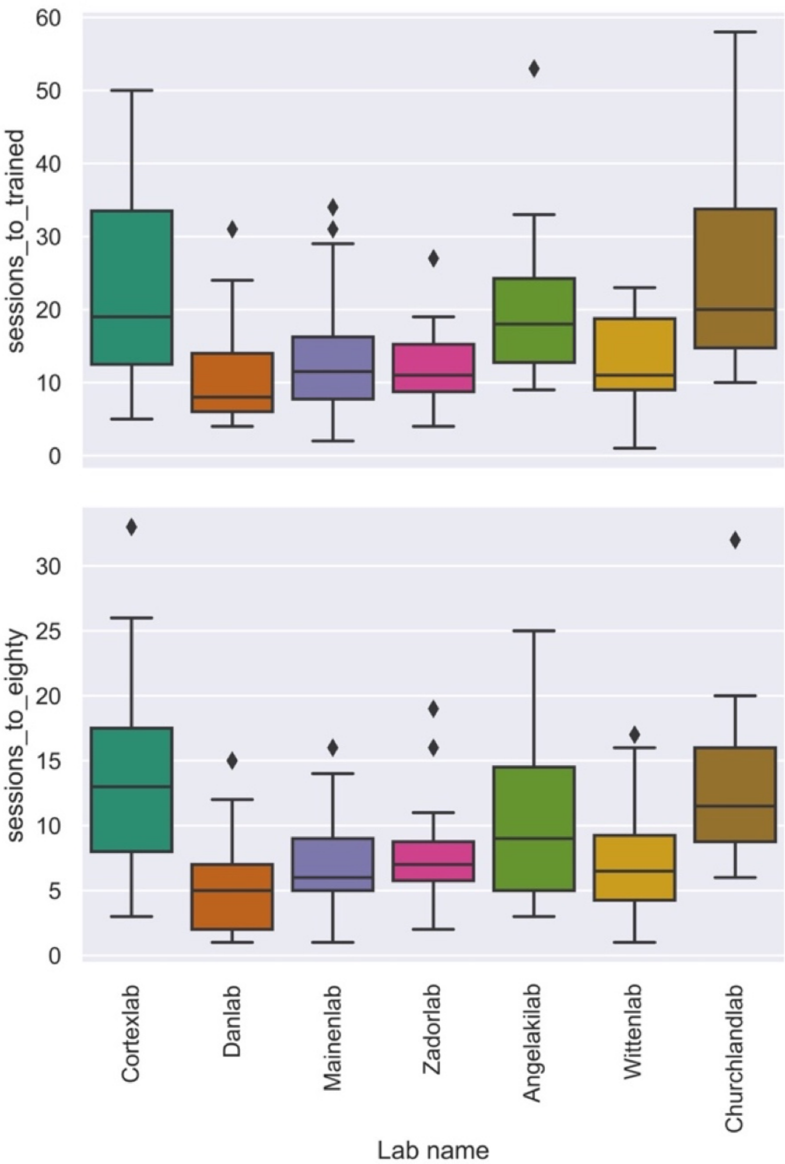


Note. The panels on the diagonal show histograms of the four learning speed variables. Each of the remaining panels show scores of the 124 mice on two of the four learning speed variables. Regression lines are given and they show the relationship between the two variables. The scores on the four learning speed variables for the example mice are plotted in colour to visualise how individual mice score.

Figure 3

Different learning speed variables across labs





Training time consistency

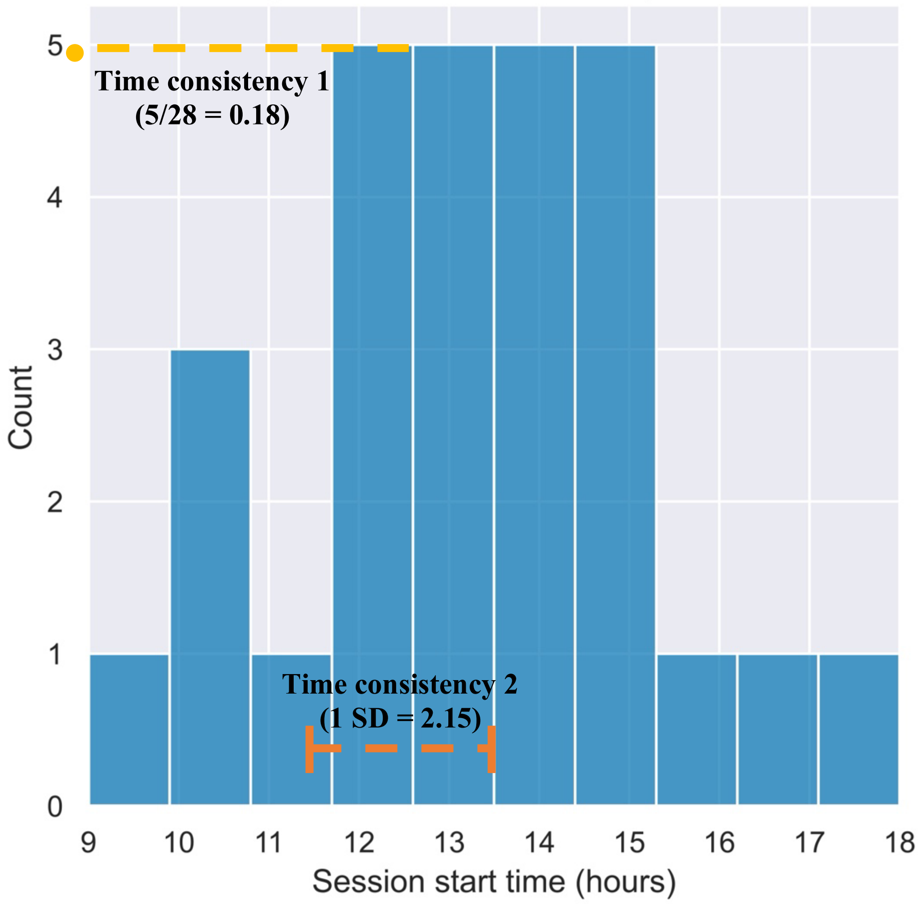
In order to test whether training time consistency affects learning speed, we computed two definitions for training time consistency, which are described below:

1. Time consistency 1: $\frac{occurences\ of\ training\ on\ preferred\ time}{num\_sessions}$
2. Time consistency 2: *standard deviation of time distributions*

The training times were rounded to the nearest hour for the ease of analysis. This was necessary as the raw data featured an excessive number of different session start times, which would have made it difficult to study their effects. The first definition, time consistency 1, is calculated as the number of times the subject is trained on the most frequent hour, divided by the total number of sessions. A higher value is related to a higher consistency. If for example a subject was trained 10 times at 11:00 AM, and twice at 7:00 PM, the time consistency would be 0.83 (10/12). However, if the same subject was trained 10 times at 11:00 AM, and once at 10:00 AM and once at 12:00 PM, the time consistency would still be 0.83, despite the difference in training times. One could argue that the subject in the second example has been trained more consistently. To address this issue, a second definition, time consistency 2, was created, which measures the standard deviation of the training time distributions. In Figure 4, we presented the two different time consistency variables for Example mouse 3. Then, the correlation between the two variables is shown in a pair plot in Figure 5. As expected, there is a negative correlation between the two definitions (r = -0.59), because for time consistency 1 a higher value is related to more consistency while for time consistency 2 a lower value is related to more consistency. Finally, we presented the average time consistency across labs in Figure 6.

Figure 4

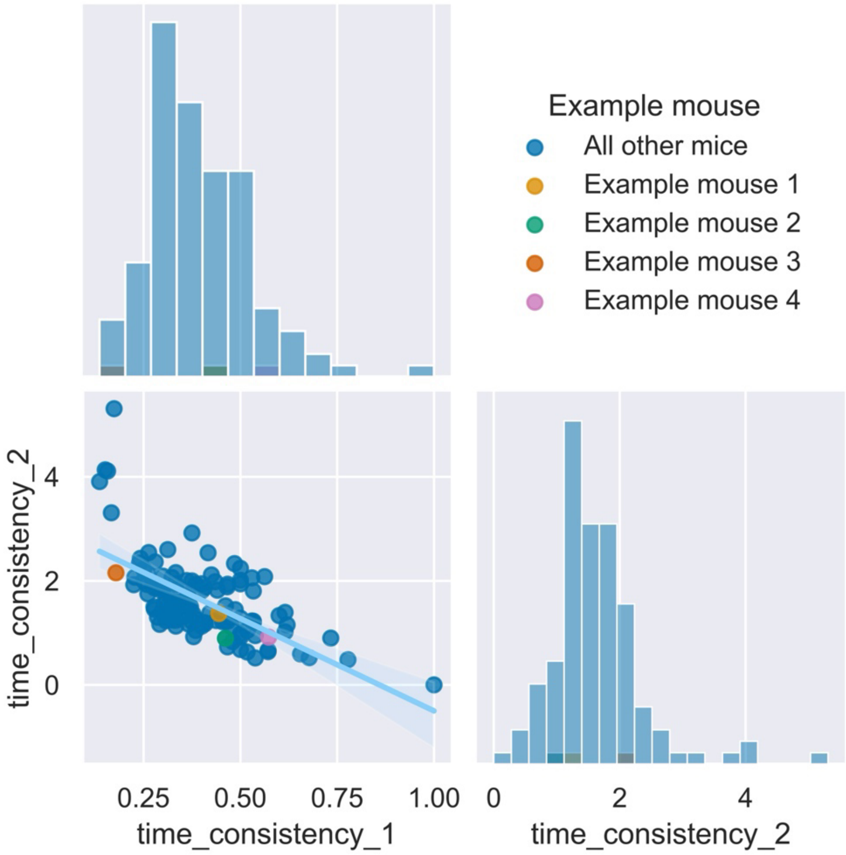
The two time consistency variables for Example mouse 3



Note. Time consistency 1 is calculated as the maximum number of sessions performed at a specific hour, divided by the total number of sessions. In this case for Example mouse 3, the total number of sessions is 28, and the maximum number of sessions performed at the same time is 5. We divide 5 by 28 which means the time consistency is equal to 0.18. Time consistency 2 on the other hand, is one standard deviation, which is 2.15 in this case.

Figure 5

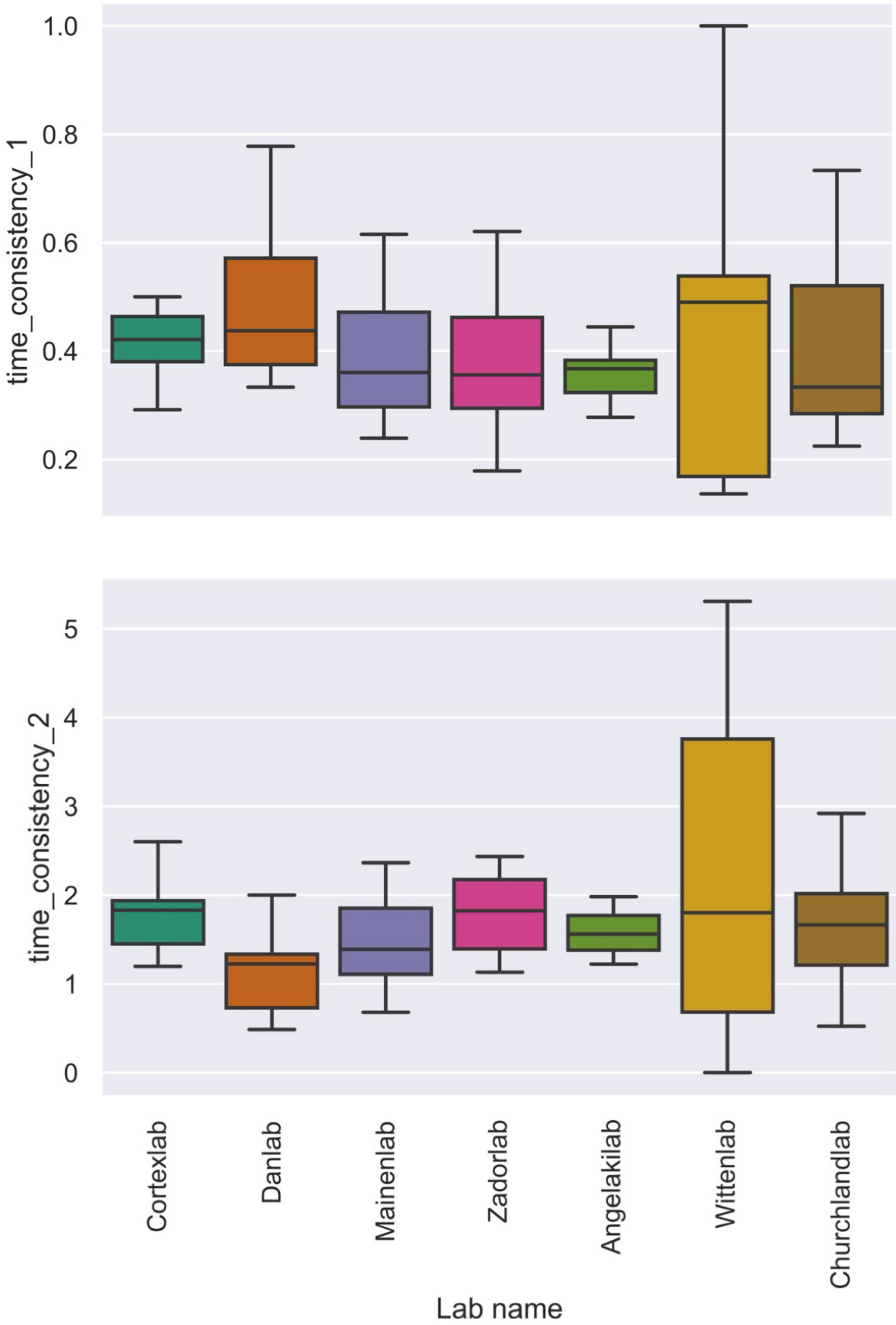
Correlations between the two time consistency variables



Note. The panels on the diagonal show histograms of the two time consistency variables. The remaining panel shows the scores of the 124 mice on the two time consistency variables. A regression line is given which shows the relationship between the two variables. The scores on the time consistency variables for the example mice are plotted in colour to visualise how individual mice score.

Figure 6

Time consistency variables per lab



Hypotheses

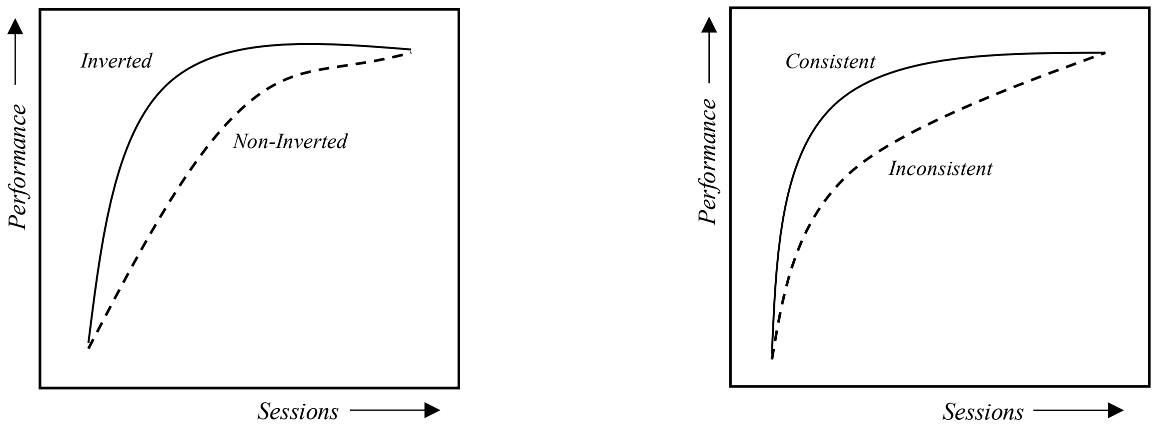
We proposed two hypotheses for this study which are presented in Figure 7. Firstly, we believed that mice that are trained in an inverted light cycle would have higher learning speed compared to mice that are trained in a non-inverted light cycle, as mice are nocturnal (Ripperger et al., 2011; Roedel et al., 2006). Secondly, we hypothesized that training time consistency would have an impact on the learning speed of mice. The scientific literature suggests that behaviour is influenced by the time of day (Daan, 2000; Fisk et al., 2018; Roedel et al., 2006; Valentinuzzi, et al., 2004), and a consistent training schedule seems to increase performance (Stroebel, 1976).

HI: The light cycle (Inverted vs. Non-inverted) influences the learning speed in mice

HII: Training time consistency affects the learning speed in mice

Figure 7

Hypotheses I and II



Testing

Kruskall-Wallis test

To examine the differences in learning speed among the seven labs, we applied a Kruskall-Wallis test. This statistical test is appropriate for testing differences among more than two groups when the data does not follow a normal distribution (Kruskal & Wallis, 1952). We used the kruskal() function from the 'scipy.stats' Python package for this analysis (Jones, Oliphant, Peterson et al., 2001).

Independent samples t-test

To determine the effect of inverted vs. non-inverted light cycle on learning speed differences between labs, we employed an independent samples t-test. This statistical test is used to compare the means of two independent groups (Gosset, 1908). The implementation was done using the stats.ttest_ind() function from the 'scipy.stats' Python package (Jones, Oliphant, Peterson et al., 2001).

Mixture models

To study the impact of the light cycle and the training time consistency, we utilized linear mixture models to account for the dependence of our observations within a specific laboratory. These models allow us to estimate both between-group and within-group variations. In linear mixture models, the difference between observations could be due to either random measurement error or differences in the underlying population (Brauer & Curtin, 2018). A mixture model, which combines fixed effects and random effects models, was used as it accommodates both types of effects. We decided to add only random intercepts to the models as adding random effects to all effects requires a lot of data per laboratory. The analysis was conducted using the 'pymer4' Python package, which is built on the 'lme4' R package. The 'lme4' package was created by Bates, Mächler, Bolker and Walker (2015) and 'pymer4' was created by Jolly (2018). The following formulas were used for the analyses: *Lmer(learning_speed ~ light_cycle + (1 | lab_name), Lmer (learning_speed ~ time_consistency + (1|lab_name)* and *Lmer(learning_speed ~ light_cycle + time_consistency + light_cycle * time_consistency + (1 | lab_name).*

Abbreviations used in the study

To enhance readability, we will utilize abbreviations for certain variables in the remaining part of the study. The four learning speed variables will be referred to as LS1, LS2, LS3, LS4, which represent learning speed 1, learning speed 2, sessions to trained and session to eighty, respectively. The non-inverted light cycle will be abbreviated as NIV light cycle, and the inverted light cycle will be denoted as IV light cycle. Finally, time consistency 1 and time consistency 2 will be referred to as TC1 and TC2, respectively.

Results

Performance in general

The training duration varied between individual mice and across different laboratories. The most rapid learner among the mice in this study achieved proficiency in the basic task within the first session, while the slowest learner took 58 sessions to reach this level of proficiency. The average performance of the 124 mice is illustrated in Figure 8, where it can be seen that the largest increase in performance occurs during the initial sessions. A similar pattern of performance increase can be observed in different laboratories, as shown in Figure 9, which presents the performance across sessions for each of the seven laboratories. Five labs trained between 12 and 15 mice, while two labs tested nearly double the number of mice (Table 1). Figure 9 highlights the differences that exist between laboratories, as previously concluded by the International Brain Laboratory et al. (2021). For instance, Danlab demonstrates a higher rate of average performance increase during the initial sessions compared to Churchlandlab (i.e., the graph for Danlab is steeper), but the average performance in Danlab decreases after the 10^{th} session, whereas the average performance in Churchladlab begins to increase after the 6^{th} session. It is important to note that the average performance per laboratory should be interpreted with caution as it may not accurately represent the performance of individual mice. Although, a Kruskall-Wallis test revealed that there are significant differences between the laboratories for LS1, LS3 and LS4 ($H = 16.96, p < .01, H = 27.78, p < .001, H = 28.25, p < .001$). There were no significant differences in LS2 ($H = 12.35, p = .05$). Finally, Figure 10 presents the number of sessions for mice across different laboratories. We can observe some variation in the number of sessions per lab. For example, the maximum number of sessions at Zadorlab is 30, while Cortexlab had a maximum of 75 sessions. The average number of sessions was found to be 23.7.

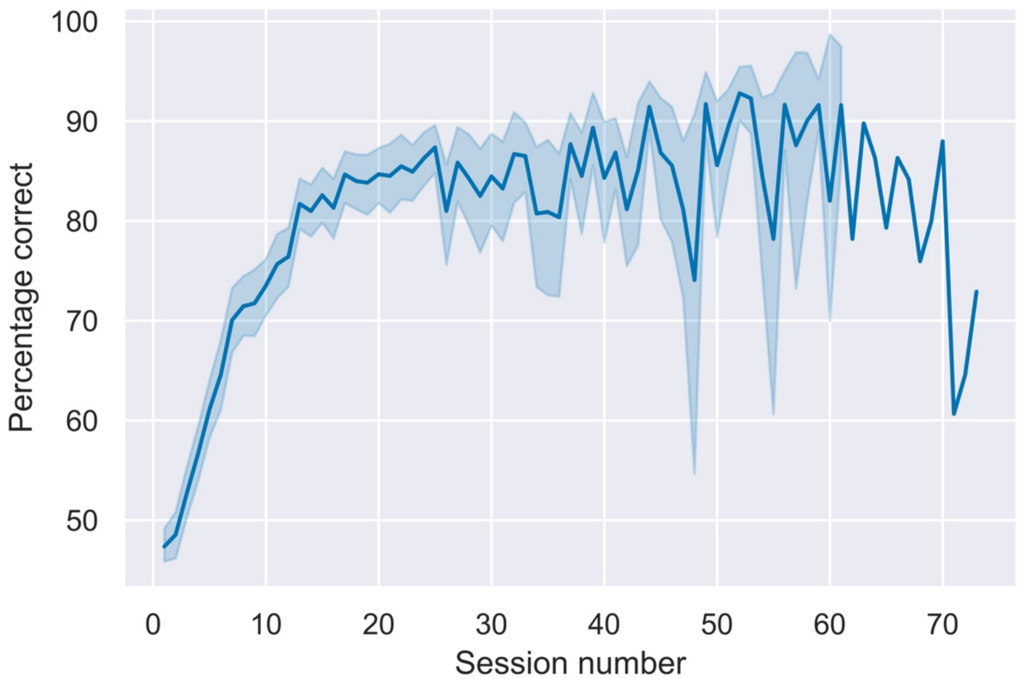
Table 1

Number of subjects per lab

Lab_name	Number of subjects
Cortexlab	15
Danlab	13
Mainenlab	32
Zadorlab	12
Angelakilab	14
Wittenlab	14
Churchlandlab	24
Total	124

Figure 8

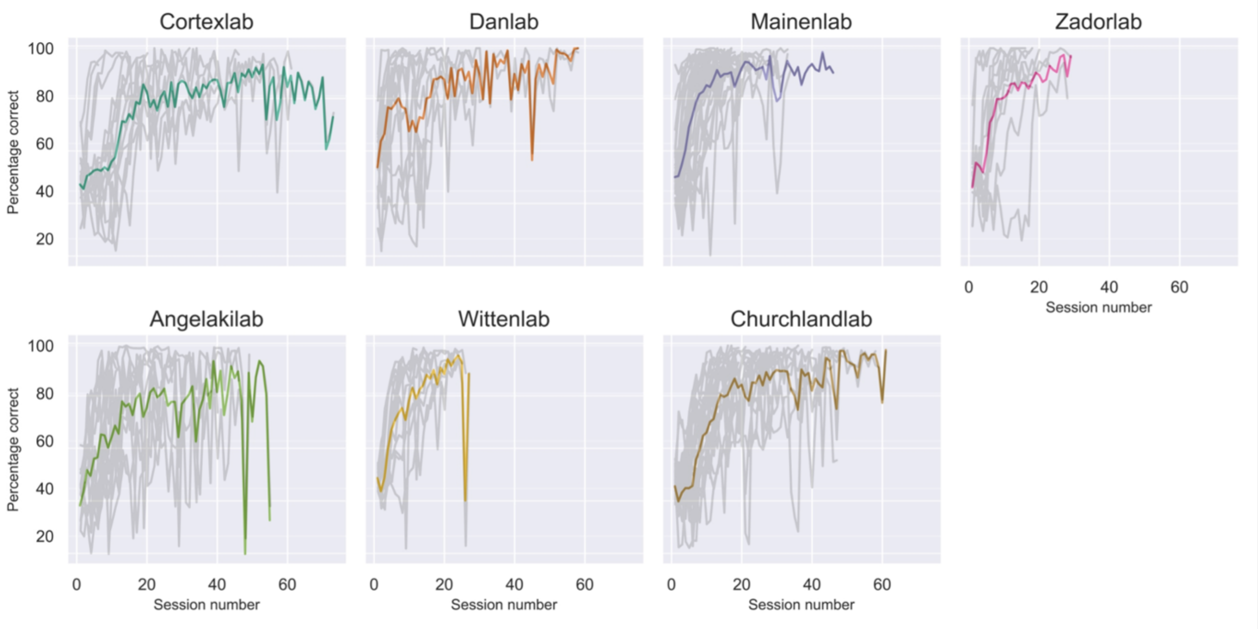
Average performance across all labs



Note. The error bar in this plot shows the 95% confidence interval.

Figure 9

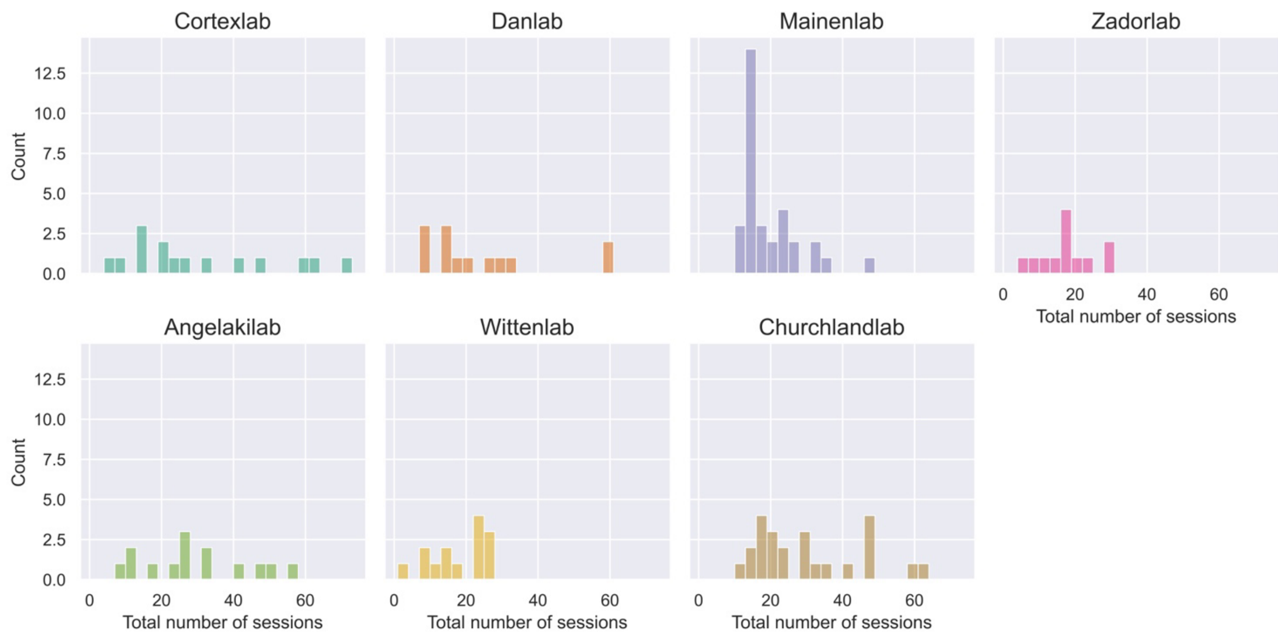
Performance over all labs



Note. Each panel displays one lab. The average performance per session is plotted for each lab. Additionally, the individual mice are plotted per lab.

Figure 10

Total number of sessions per mouse across labs



Example mice

To facilitate a clearer understanding of the performance across sessions, four representative mice have been selected and their details such as laboratory and sex, are given in Table 2. These example mice were trained in different laboratories, have a different sex and were trained for varying numbers of sessions. Figure 11 shows the performance of the example mice, demonstrating the substantial individual differences in learning speed between mice. For instance, example mouse one and two were trained for a relatively short number of sessions (9 and 13 sessions, respectively), while example mouse three and four were trained for a greater number of sessions (28 and 61 sessions, respectively). Furthermore, the performance of some mice, such as example mouse one, increases significantly during the initial sessions, while for other, as example mouse three, the performance improvement only becomes apparent after 18 sessions. Table 3 provides additional information about the learning speed and the training time consistency for the example mice. Moreover, the distribution of the training time for the example mice is shown in Figure 12, with the y-axis indicating the start time of the sessions in hours. It is important to note that these values have been rounded to the nearest hour. There is a substantial variation in the training time between mice, with example mouse four begin trained predominantly in the morning (around 9 and 10 AM), while example mouse three displays a greater degree of variability in training time. The following sections contain a more in-dept analysis of learning speed and training time consistency.

Table 2

General information of the example mice

Example mouse	Lab name	Sex	Subject birth date	Light cycle	Total no sessions
1	Angelakilab	M	2018-12-18	IV	9
2	Mainenlab	M	2018-10-23	NIV	13
3	Zadorlab	F	2019-09-17	NIV	28
4	Churchlandlab	M	2019-08-06	IV	61

Table 3

Learning speed and training time consistency for example mice

Example mouse	LS1	LS2	LS3	LS4	TC1	TC2
1	0.03	0.03	9.0	7.0	0.44	1.37
2	0.03	0.02	8.0	6.0	0.46	0.89
3	0.03	0.01	27.0	19.0	0.18	2.15
4	0.01	0.00	36.0	18.0	0.57	0.91

Figure 11

Performance of the example mice

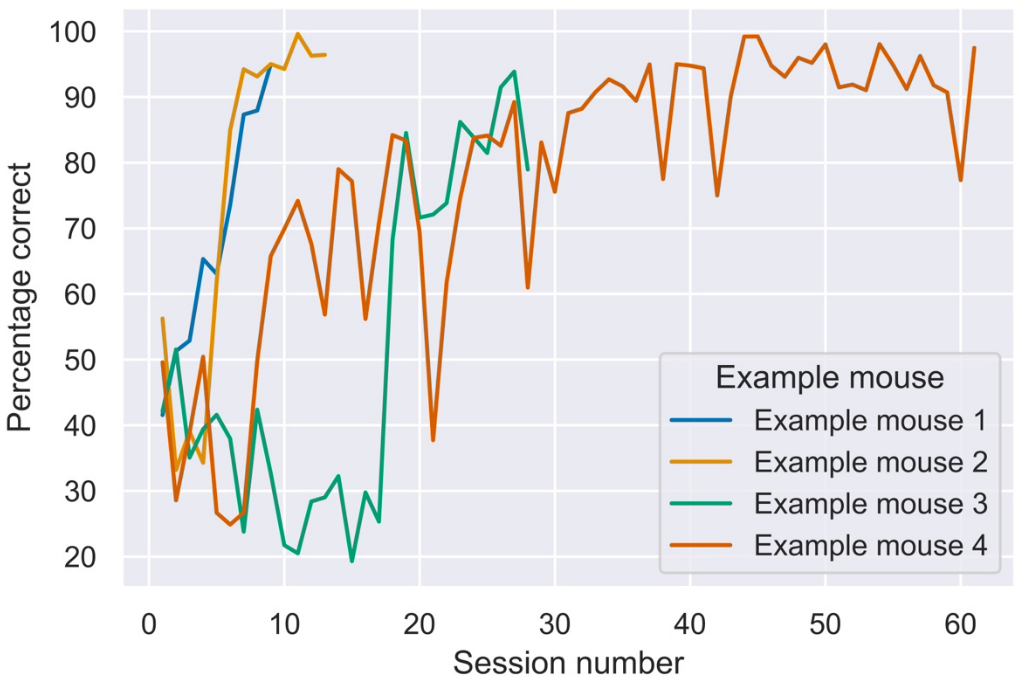
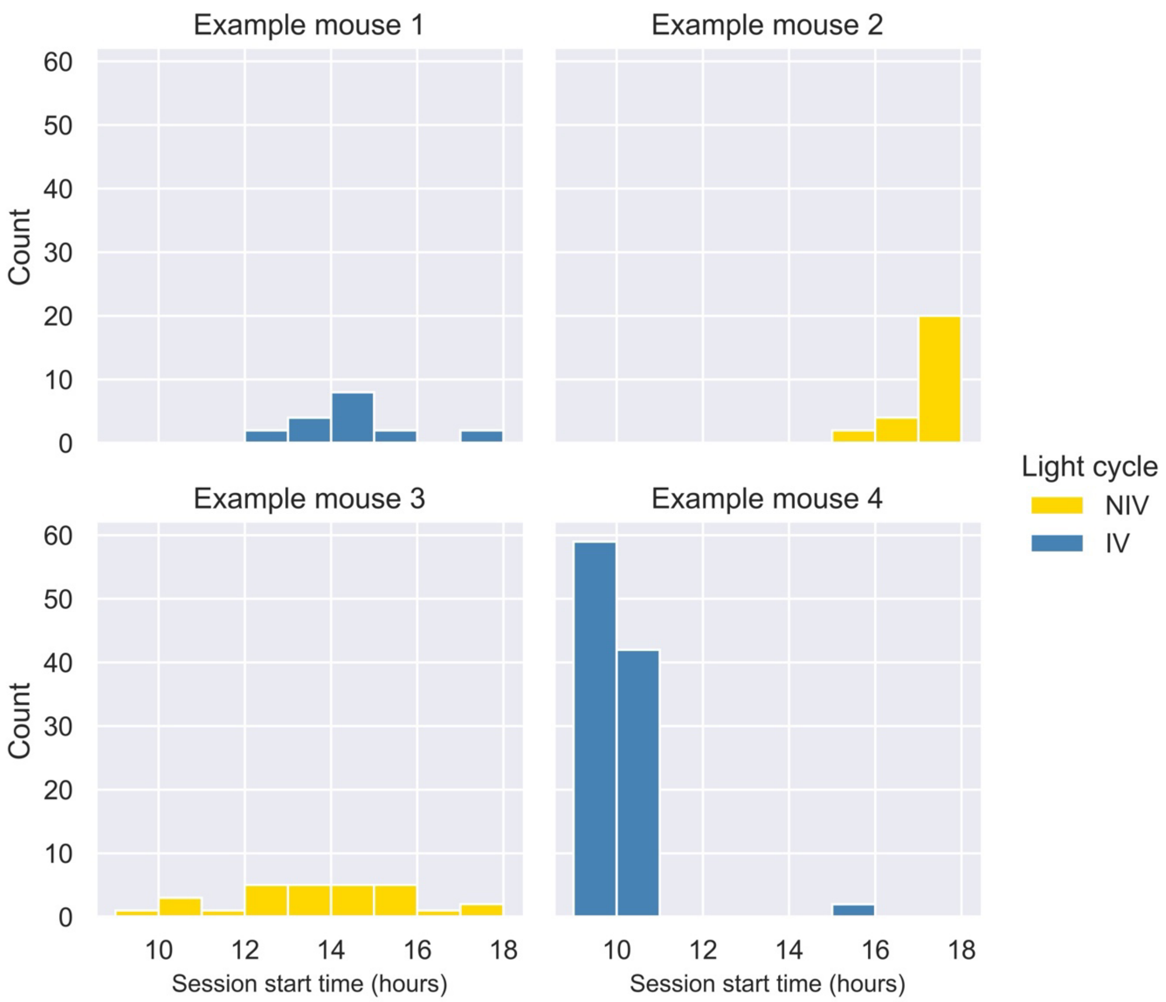


Figure 12

Training time distribution for the example mice

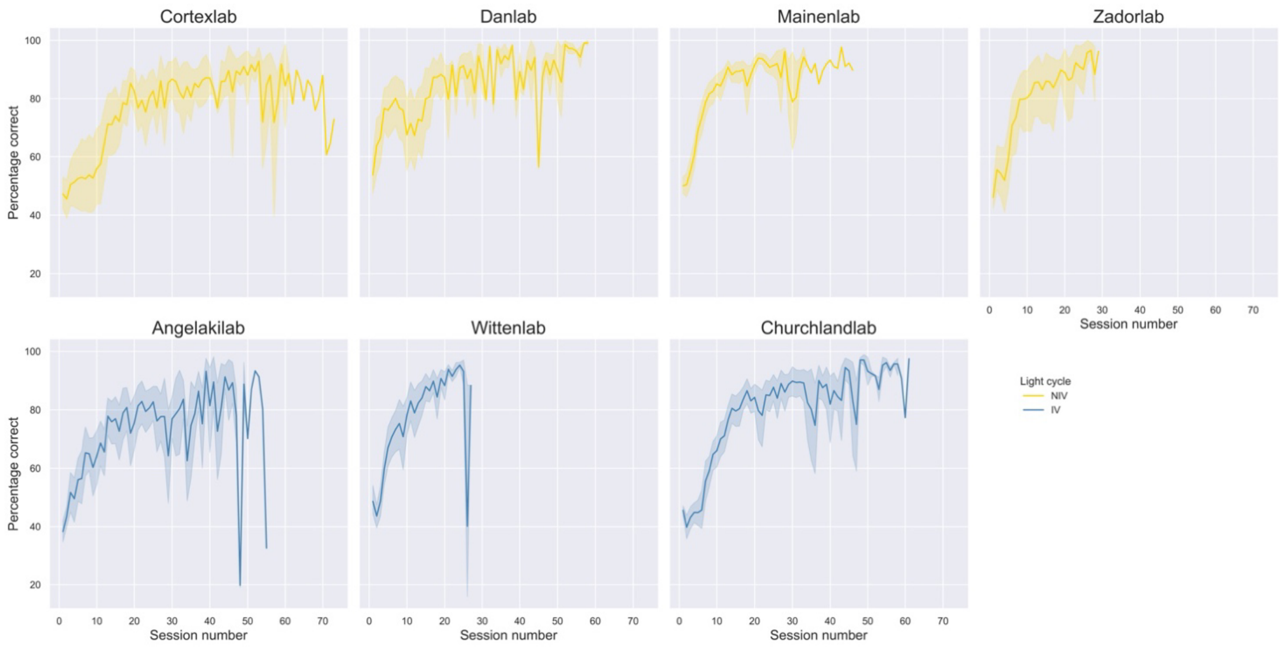


Light cycle

The lighting conditions during the experiments have the potential to influence the learning speed of the mice. This study analysed the performance of mice trained under either a non-inverted (NIV) or an inverted (IV) light cycle. Four laboratories used a NIV light cycle, which provided light during the day and darkness at night, while the remaining three laboratories used an IV light cycle, with darkness during the day and light at night. These three labs conducted their experiments in the dark, with the expectation that the shift in light-dark cycle would result in the mice being in their active phase. The performance of both light cycle groups presented in both Figure 13 and Figure 14. While there appears to be a slight difference in the learning speed between the two light cycle groups, it is difficult to draw definitive conclusions from this data. The results seem to suggest that mice that are trained in an NIV light cycle have a higher learning speed, which contradicts the hypothesis that nocturnal animals would perform better under an IV light cycle with darkness during the day.

Figure 13

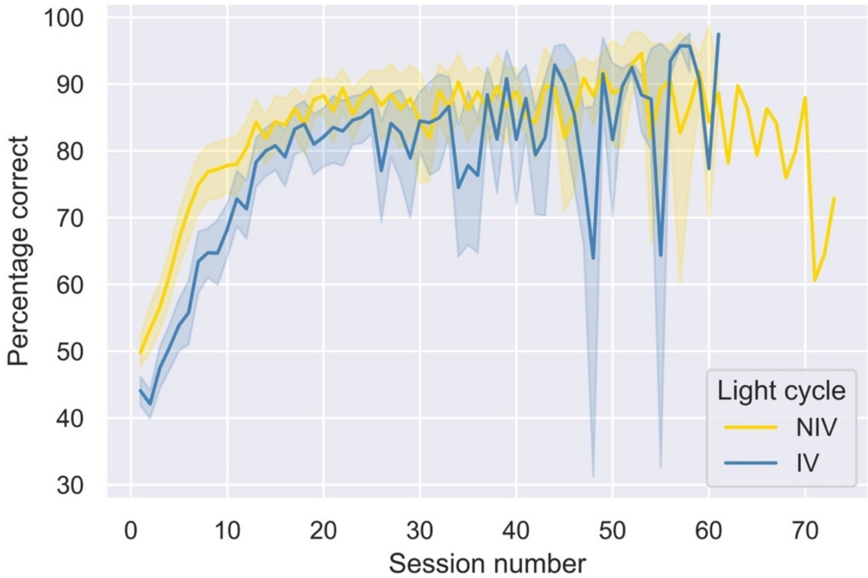
Performance across labs and their light cycle



Note. The error bars in this plot show the 95% confidence interval.

Figure 14

Performance for labs with a NIV or an IV light cycle

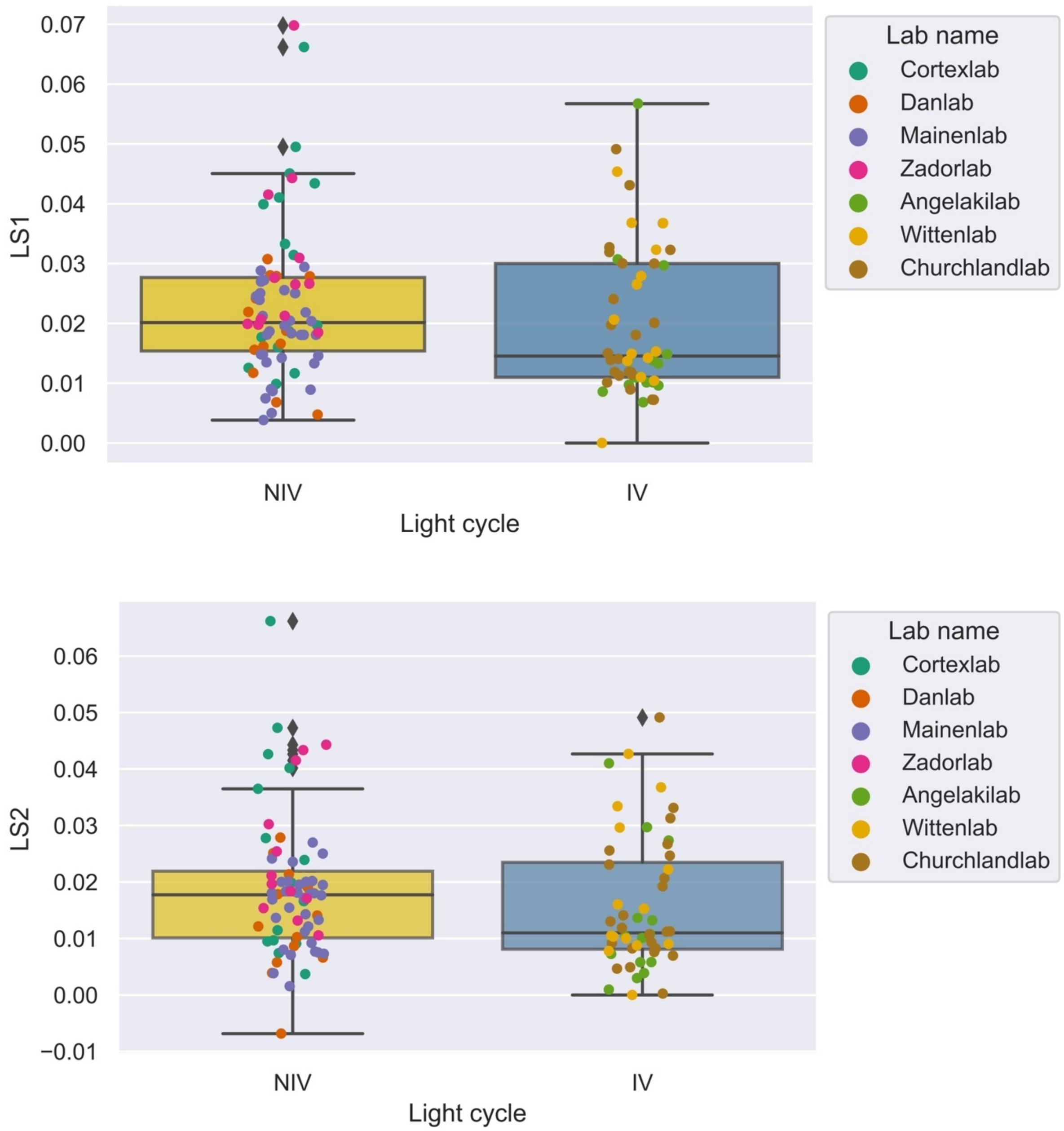


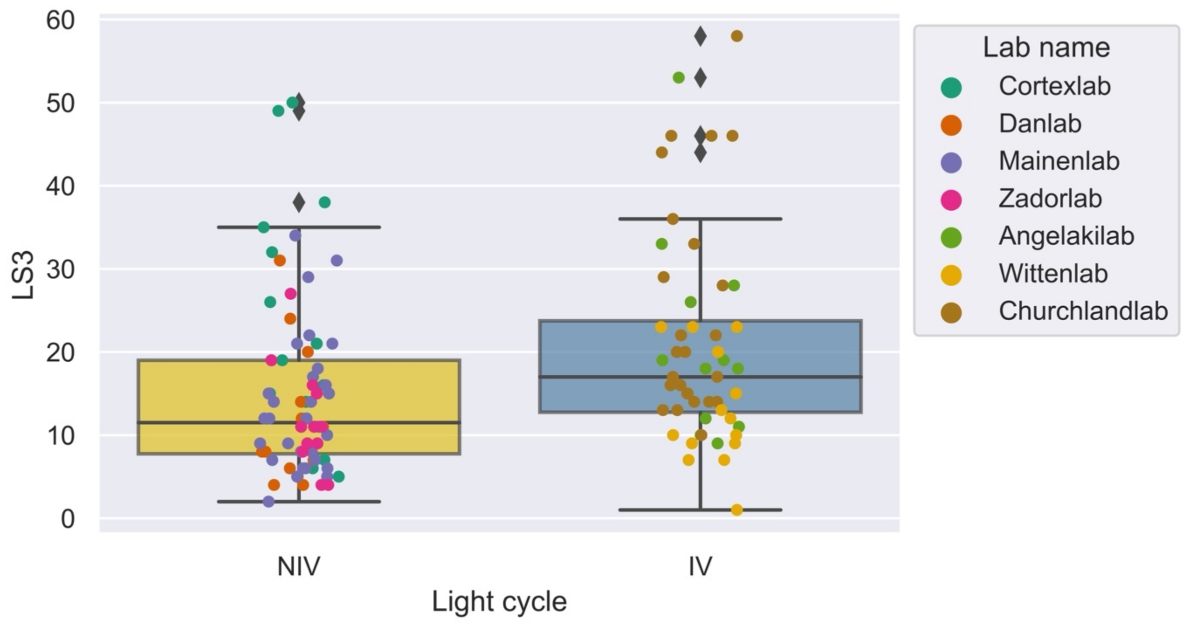
Note. The error bars in this plot show the 95% confidence interval.

In order to analyse the impart of the light cycle on the learning speed of the mice, we performed t-tests to investigate differences in LS1-LS4. The results of these tests are presented in Table 4, which displays the average scores on LS1-LS4 of mice that were trained in an IV or NIV light cycle. Although LS1 and LS2 were found to be slightly lower for labs that utilized an IV light cycle, these differences were not considered to be significant. Conversely, both LS3 and LS4 were found to be slightly higher for labs that utilized an IV light cycle and these differences are significant. It is important to note that a higher value for LS3 and LS4 relates to slower learning. On top of that, the means for each light cycle group (NIV/ IV) may be influenced by the performance of mice in a specific laboratory. For instance, if the mice in Mainenlab performed significantly better than those in the other labs, this would result in a higher mean for the entire NIV light cycle group. These differences may not be specifically attributable to the light cycle, but instead could be lab-specific. To address this, the learning speed variables were analysed per laboratory and displayed in Figure 16.

Figure 15

Learning speed across labs with a NIV light cycle and an IV light cycle





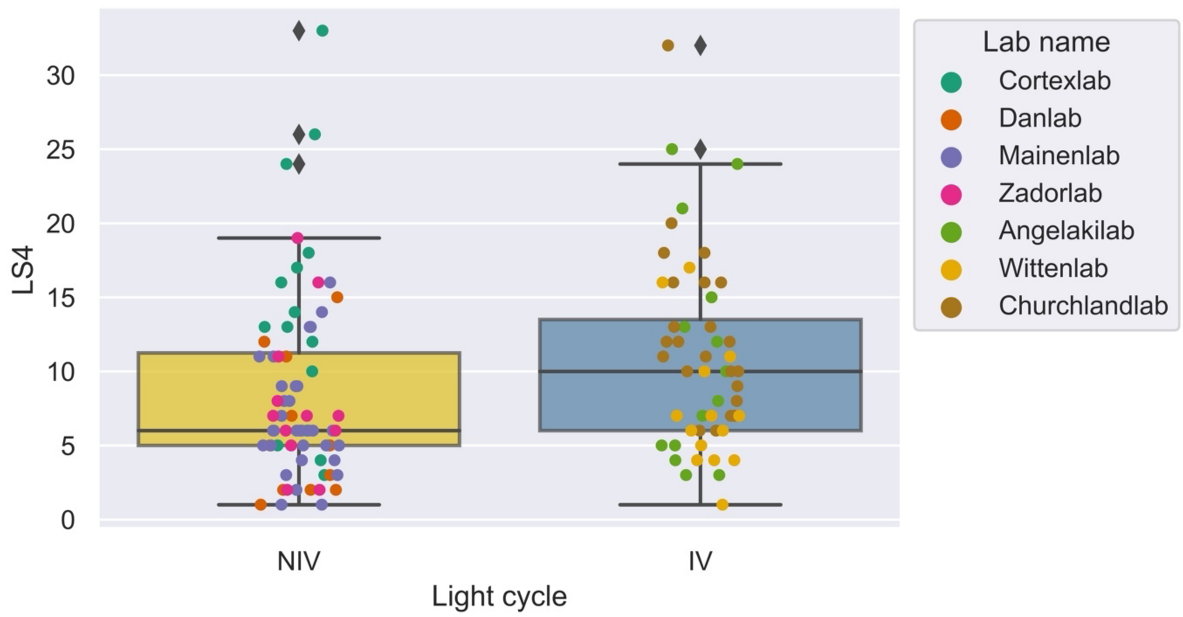


Table 4

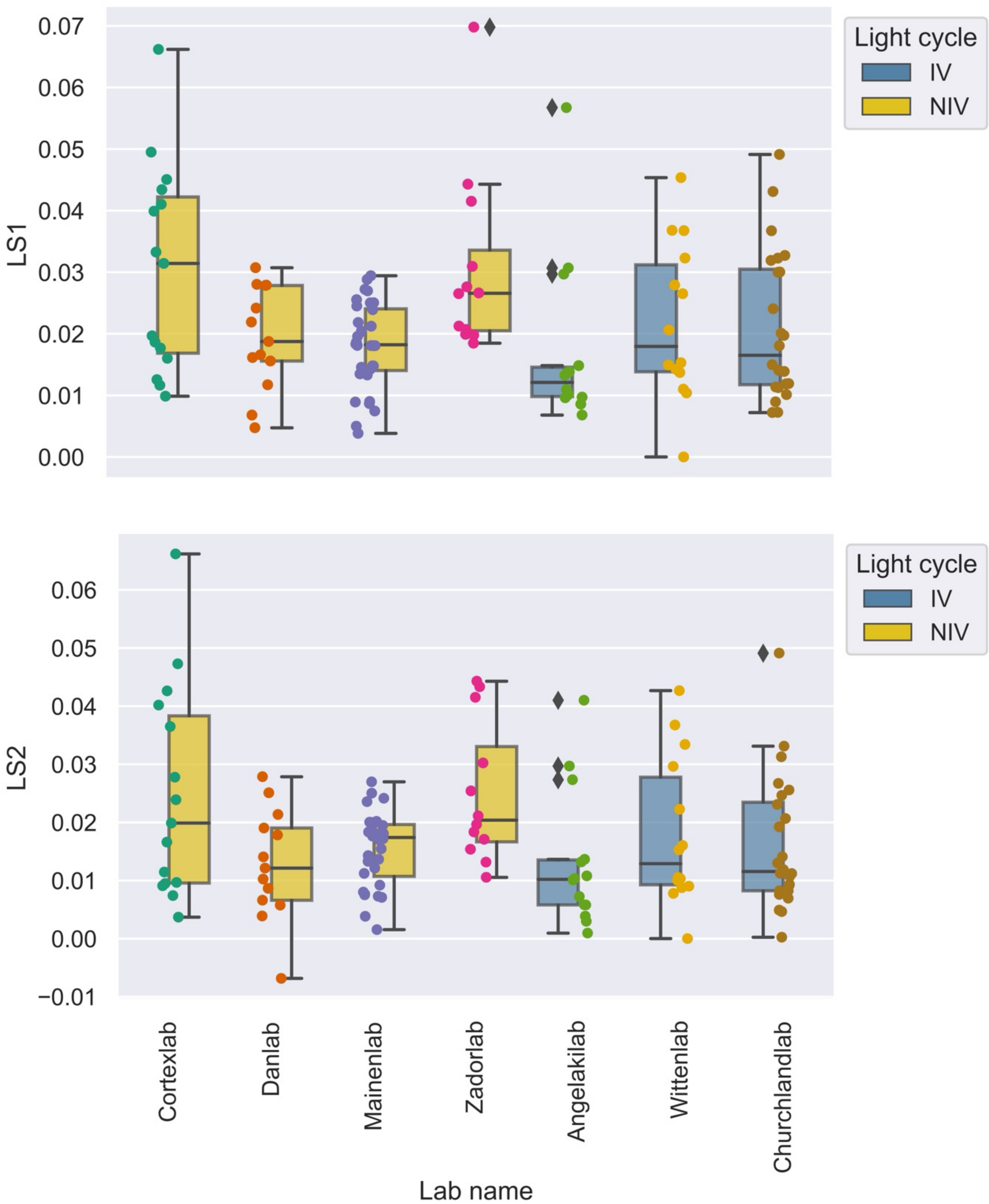
Learning speed variables for non-inverted light cycle and inverted light cycle

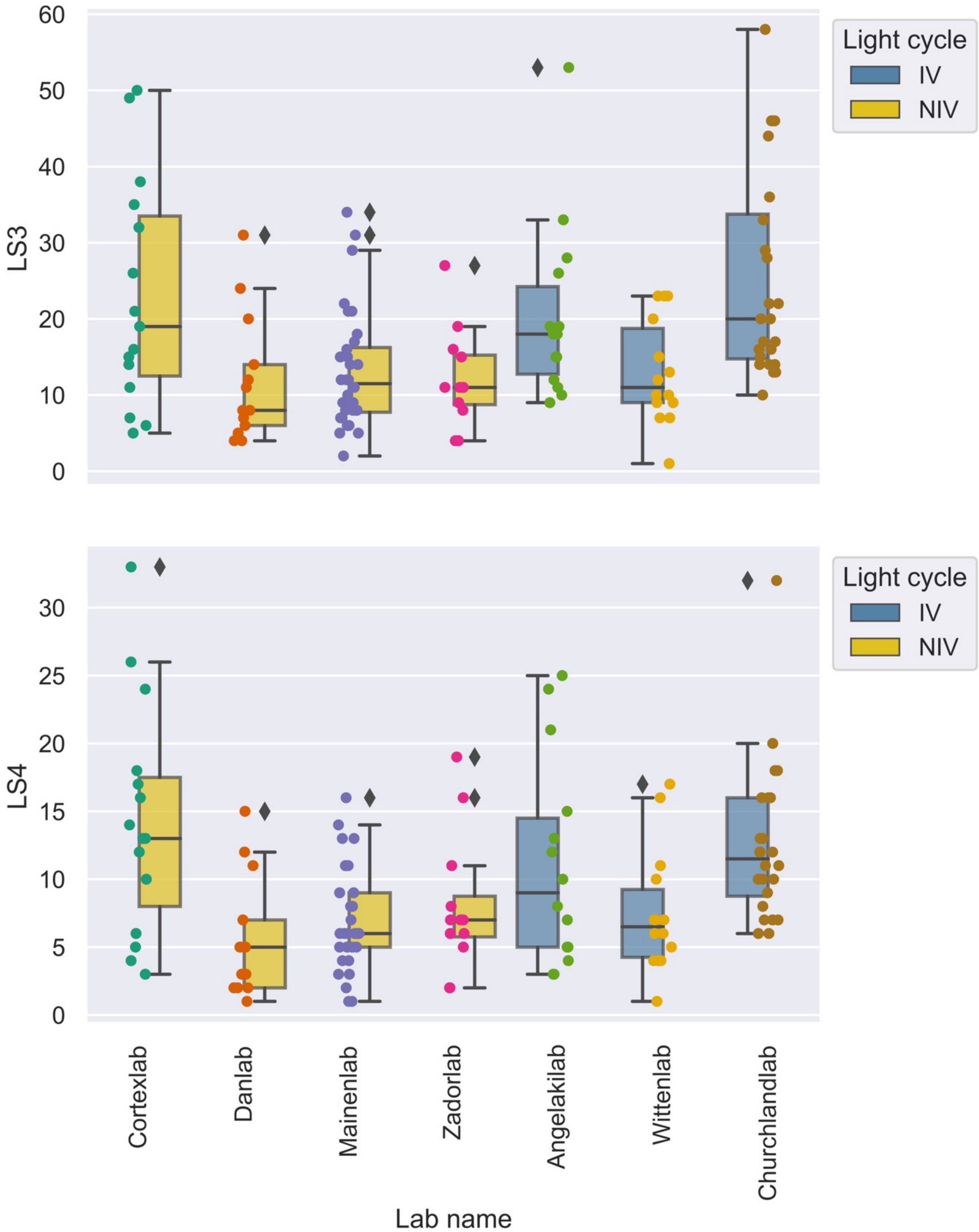
Learning speed variable	NIV light cycle		IV light cycle			
	M	*SD*	*M*	*SD*	*t*	*p* value
LS1	0.0	0.0	0.0	0.0	1.23	0.222
LS2	0.0	0.0	0.0	0.0	1.25	0.215
LS3	14.7	10.3	20.7	12.6	-2.90**	0.004
LS4	8.4	6.1	10.8	6.3	-2.14*	0.035

*$p < .05$. **$p < .01$. ***$p < .001$

Figure 16

Light cycle and learning speed per lab




LS3
LS4
Light cycle
IV
NIV
60
50
40
30
20
10
0
25
15
5
Cortexlab
Danlab
Mainenlab
Zadorlab
Angelakilab
Wittenlab
Churchlandlab
Lab name

By examining Figure 16, it appears that Cortexlab may be an outlier in the NIV light cycle group for LS3 and LS4. Its values are higher than those of the other three labs (Danlab, Mainenlab and Zadorlab). However, this is not necessarily the case for LS1 and LS2. To determine the reason for Cortexlab's higher scores, we analysed the number of mice and sessions in Table 1 and Figure 9. Cortexlab has 15 subjects, which is similar to Danlab (13 subjects) and Zadorlab (12 subjects) but not to Mainenlab (32 subjects). Additionally, some mice in Cortexlab have been trained for more sessions than those in Danlab, Mainenlab and Zadorlab. This also explains higher values for LS3 and LS4. The total number of sessions is typically determined by the subjects' proficiency in the basic task, but in some cases, training continued to improve performance (The International Brain Laboratory et al., 2021). Hence, caution should be taken when interpreting LS1 and LS2, as they may be smaller than what is representative when mice trained for more sessions than necessary. However, as LS1 and LS2 are high it is not a concern for this matter. The higher scores for LS3 and LS4 may be due to the larger variability in Cortexlab's scores compared to the other labs in the NIV light cycle group. When examining the IV light cycle group, Angelakilab stands out. It has the same number of subjects as Wittenlab (14 subjects), but not as many as Churchlandlab (24 subjects). The mice in Angelakilab score low on LS1 and LS2. Figure 10 shows that some mice in Angelakilab have been trained for more sessions than those in Wittenlab, but not necessarily more than those in Churchlandlab. Churchlandlab has a lot of variability in all learning speed variables.

We cannot identify a reason for the differences in scores between laboratories. However, the differences in learning speed between laboratories that used a NIV compared to an IV light cycle might be due to other lab-specific variables, which is why we performed a mixture analysis. The results are presented in Table 5, which shows that none of the learning speed variables are significant. This allows us to reject our hypothesis that mice that are trained in IV light cycle have higher learning speed than mice in a NIV light cycle.

Table 5

Results of mixture model for light cycle

Learning speed variable	Coefficient	Coefficient value	*t*	*p* value
LS1	Light cycle	-0.00	-1.00	0.364
LS2	Light cycle	-0.00	-0.92	0.401
LS3	Light cycle	4.84	1.11	0.317
LS4	Light cycle	1.76	0.68	0.526

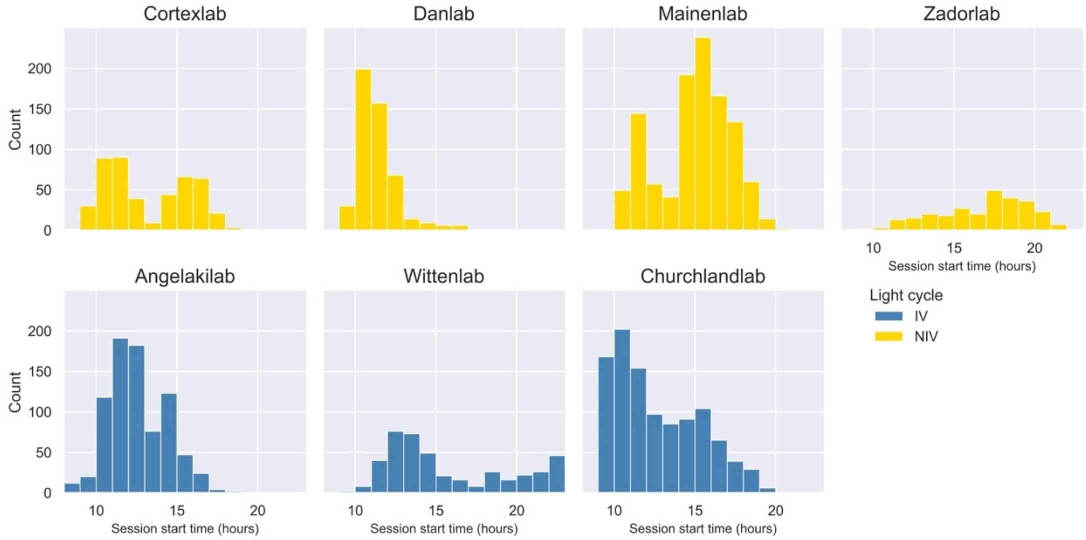
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Training time consistency

The training time in the experiments varies from lab to lab and also from session to session. The inconsistency in training time could impact the learning speed. Hence, we further explore these differences and the training time distribution for each lab is shown in Figure 17. It can be observed that there is quite some variation. For instance, Danlab mostly trains the animals in the morning, while Mainenlab often trains in the afternoon.

Figure 17

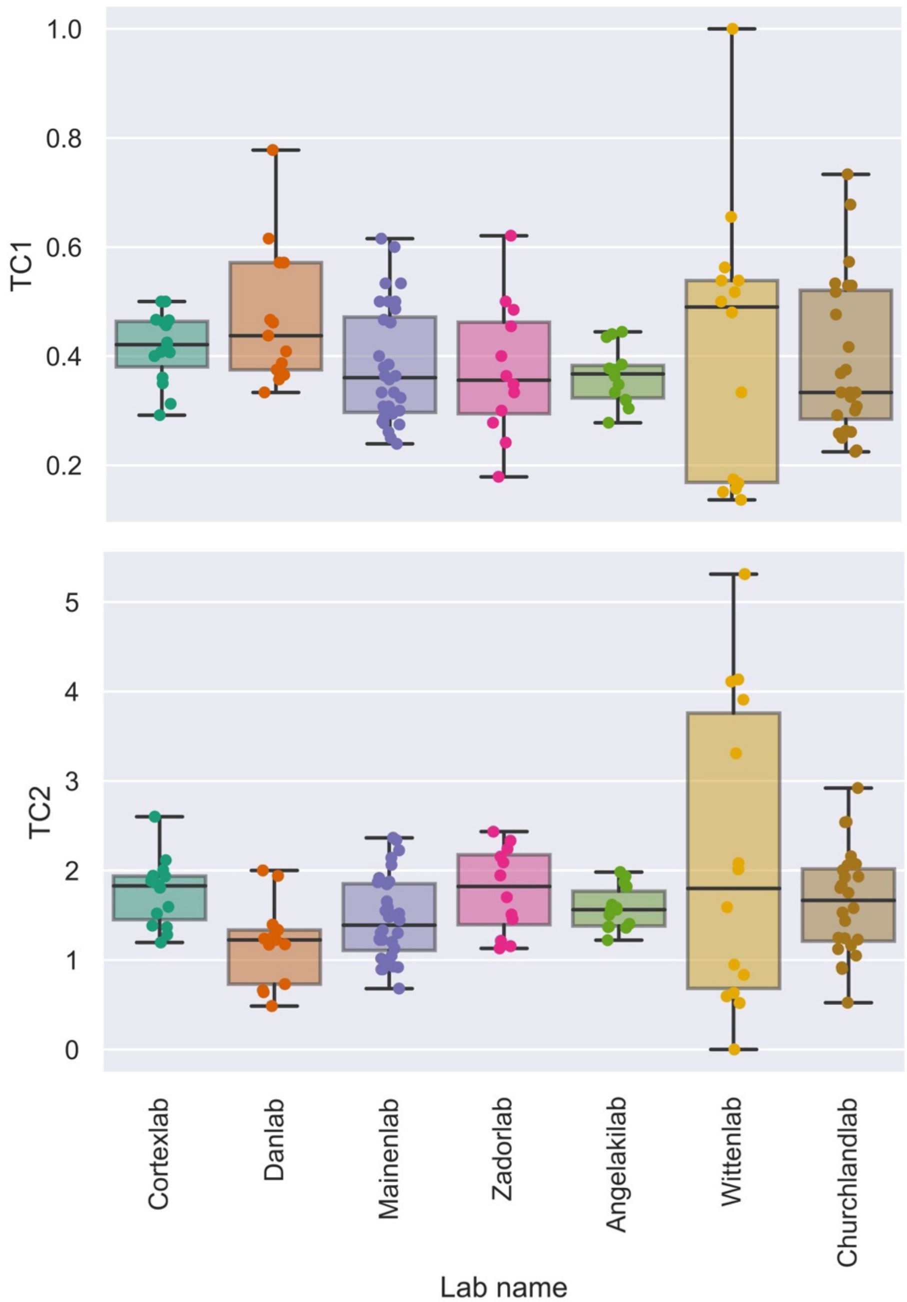
Training time distributions per lab with light cycle



In Figure 18 we presented TC1 and TC2 for all labs and to interpret this figure it is important to repeat that for TC1 a higher value is related to more consistency while for TC2 a higher value is related to less consistency. Again, we observe differences between labs. For example, there is little variance in Angelakilab, while there is quite some variance in Wittenlab. When examining the training time distributions shown in Figure 17, it appears that TC1 and TC2 at each lab make sense. For instance, Danlab has high scores for TC1 (indicating high consistency) and low scores for TC2 (indicating high consistency), which aligns with the fact that their experiments are mostly performed in the morning.

Figure 18

TC1 and TC2 per lab

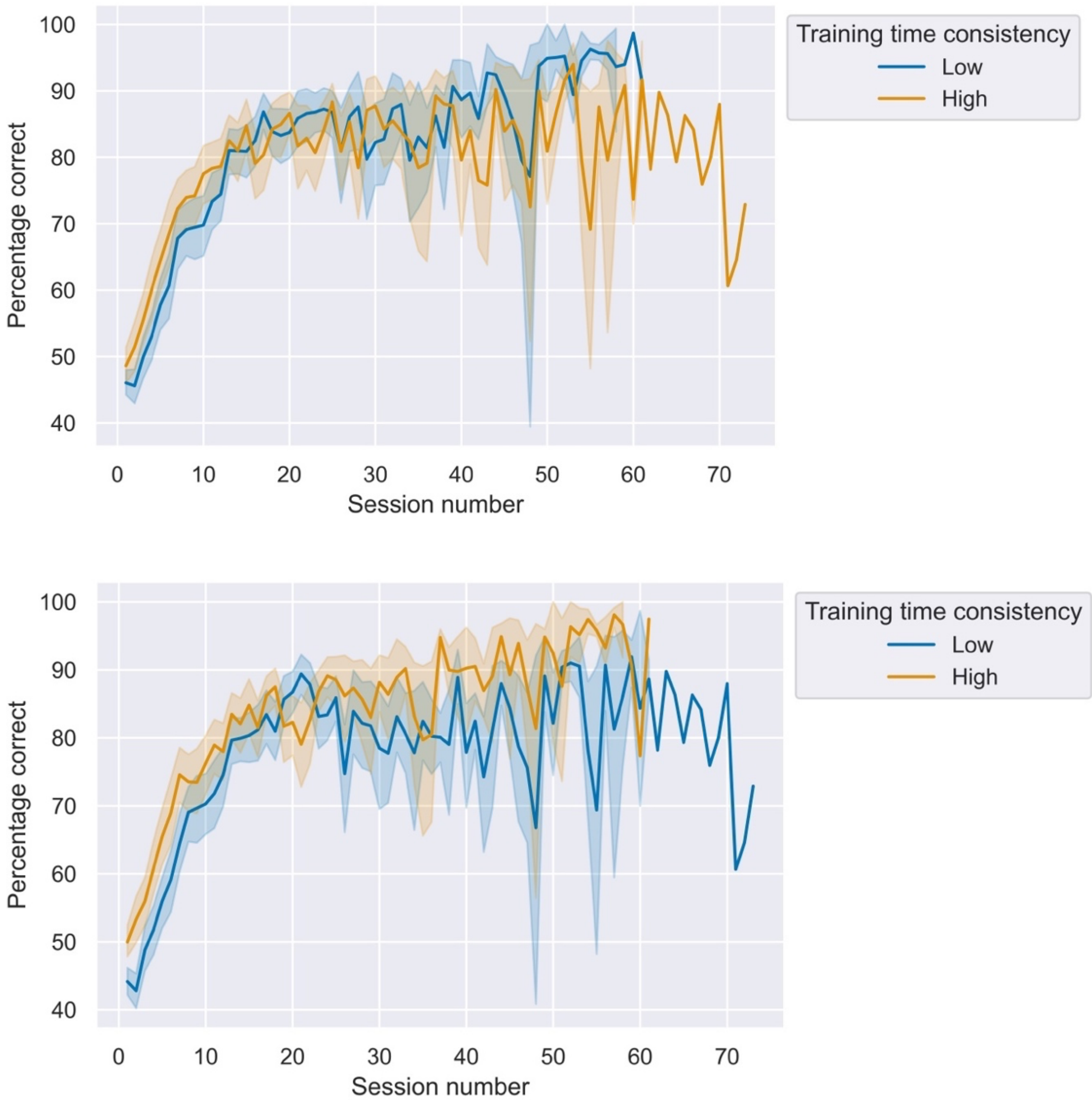


To better understand the impact of training time consistency on learning speed, we have categorised the mice into either low training time consistency or high training time consistency. The median values for TC1 and TC2 were used to distinguish between the two levels of consistency. For TC1, values below 0.38 are considered low and values above 0.38

are considered high. For TC2, the opposite holds true, as low values (below 0.50) indicate high consistency and high values (above 0.50) indicate low consistency. The results are present in Figure 19, which shows the differences in learning speed based on consistency levels for all labs. It seems that higher consistency leads to higher learning speed, although the differences are small and further testing is required before reaching a definite conclusion.

Figure 19

Peformance per sessions with different (categorical) time consistency for all labs



Note. The error bars in this plot show the 95% confidence interval.

For both TC1 and TC2, we conducted t-tests to compare the means of the four learning speed variables between low and high training time consistency groups. The results for TC1 showed that none of the learning speed variables were significantly different (Table 6). For TC2, however, the means of learning speed between low consistency and high consistency groups differed significantly for LS3 and LS4 (Table 7). In both LS3 and LS4, higher consistency in training time resulted in lower means. This suggests that consistent training times lead to faster learning. However, as this difference was only observed in two out of eight tests, we should be cautious when drawing conclusions. It is also important to note that other lab-specific variables might have influenced the results. To address this, we used mixture models to control for laboratory specific variation.

Table 6

TC1, t-tests for low vs high consistency

Learning speed variable	Low consistency		High consistency			
	M	*SD*	*M*	*SD*	*t*	*p* value
LS1	0.0	0.0	0.0	0.0	-1.63	0.106
LS2	0.0	0.0	0.0	0.0	-1.70	0.092
LS3	18.9	11.0	15.5	12.2	1.65	0.102
LS4	10.3	6.5	8.4	6.0	1.78	0.078

**p* < .05. ***p* < .01. ****p* < .001

Table 7

TC2, t-tests for low vs high consistency

Learning speed variable	Low consistency		High consistency			
	M	*SD*	*M*	*SD*	*t*	*p* value
LS1	0.0	0.0	0.0	0.0	-1.61	0.109
LS2	0.0	0.0	0.0	0.0	-0.99	0.322
LS3	19.7	12.4	14.8	10.4	-2.40*	0.018
LS4	10.7	6.7	8.0	5.6	-2.52*	0.013

**p* < .05. ***p* < .01. ****p* < .001

The outcomes of the mixture models for TC1 are presented in Table 8. This shows that none of the learning speed variables were found to be significant. The outcome of the mixture analysis for TC2 are presented in Table 9, and again, none of the learning speed variables showed significance. After controlling for laboratory specific variation, it appears that differences in learning speed could not be attributed to training time consistency. Therefore, we can conclude that training time consistency does not have impact on learning speed in this study.

Table 8

Mixture models for time consistency 1 (continuous)

Learning speed variables	Coefficient	Coefficient value	*t*	*p* value
LS1	TC1	0.01	0.67	0.503
LS2	TC1	0.01	0.65	0.518
LS3	TC1	-13.72	-1.92.	0.057
LS4	TC1	-7.05	-1.83.	0.069

$*p < .05$. $**p < .01$. $***p < .001$

Table 9

Mixture models for time consistency 2 (continuous)

Learning speed variables	Coefficient	Coefficient value	*t*	*p* value
LS1	TC2	0.00	0.06	0.95
LS2	TC2	-0.00	-0.34	0.731
LS3	TC2	1.12	0.84	0.405
LS4	TC2	0.73	1.02	0.311

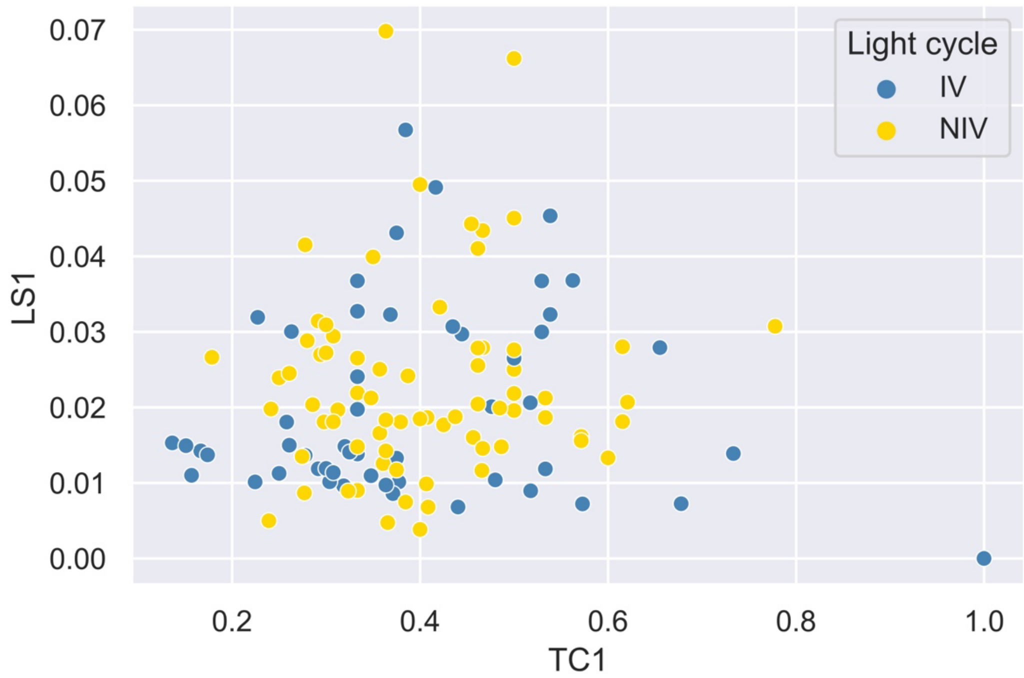
$*p < .05$. $**p < .01$. $***p < .001$

Light cycle and time consistency in one model

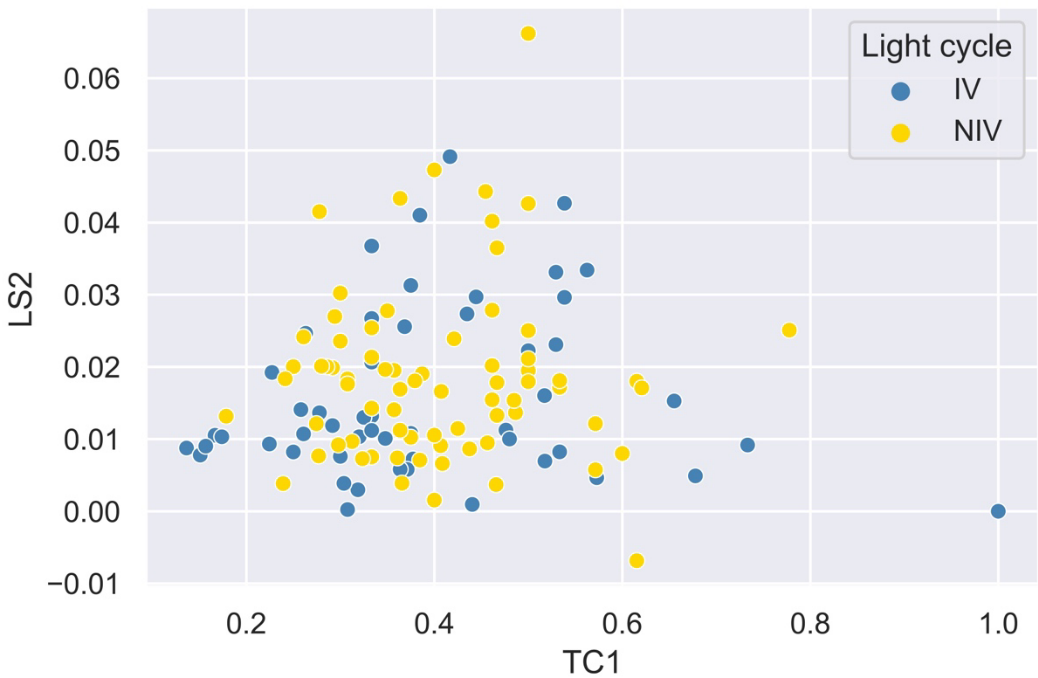
Additionally, we created mixture models that considered both the light cycle and training time consistency, as there could be an interaction between the two factors. For instance, the impact of training time consistency may only be evident when mice are trained in an IV light cycle. First, we investigate Figure 20, which shows the relationship between training time consistency and learning speed for all variables. Here, we also present the light cycle. From this figure we cannot conclude there is a pattern between the variables. The outcome of the analysis for TC1 and TC2 are presented in Table 10 and 11, respectively.

Figure 20

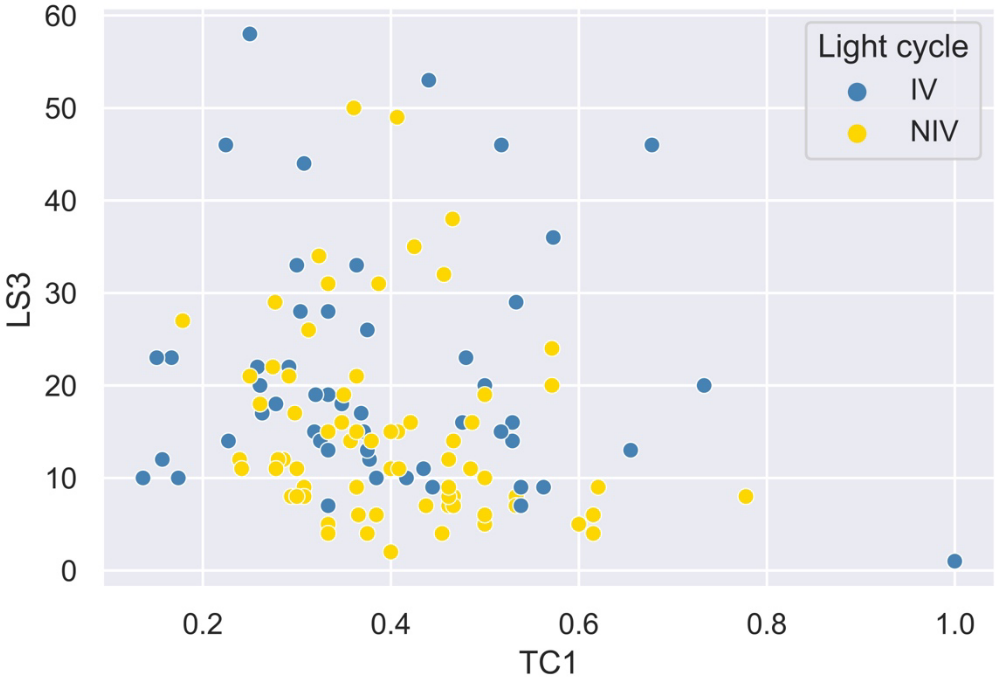
The effect of time consistency on learning speed (by light cycle)



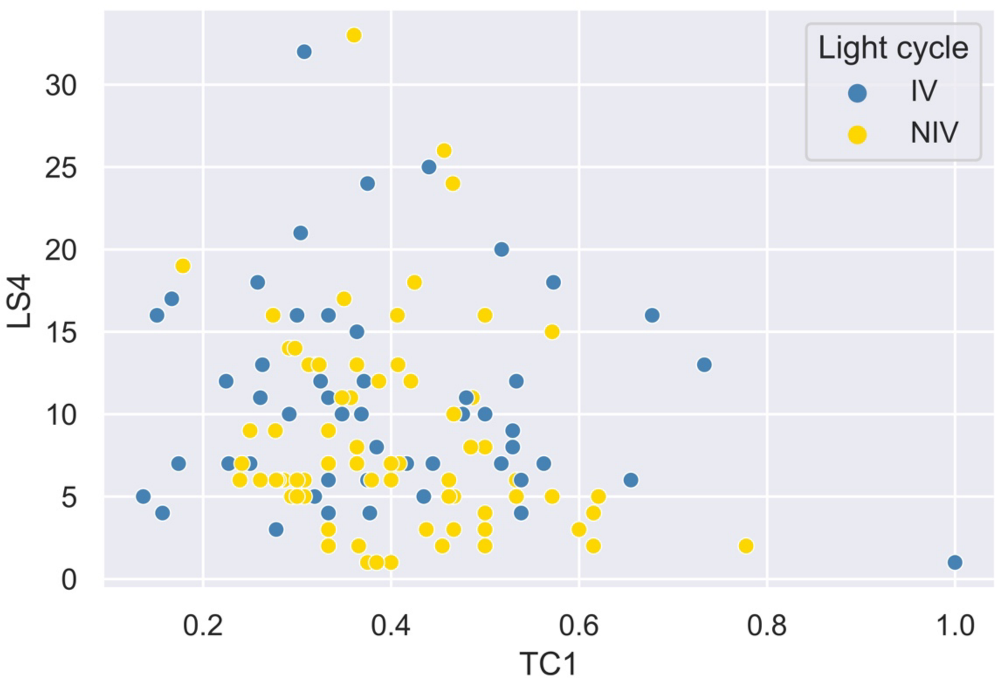
Note. There were no significant effects in this analysis



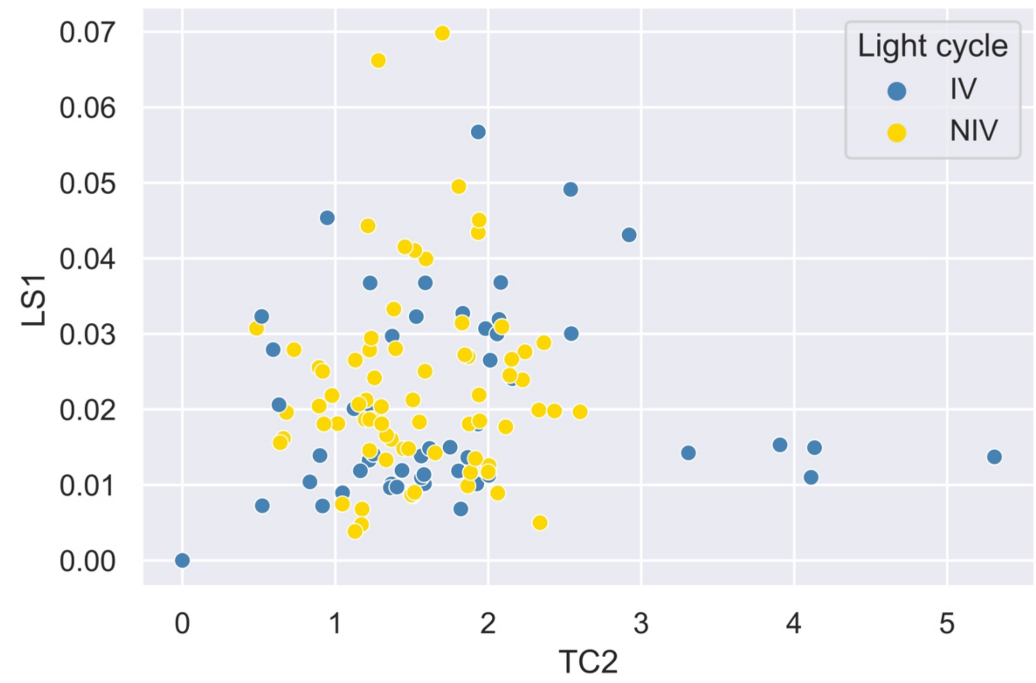
Note. There were no significant effects in this analysis



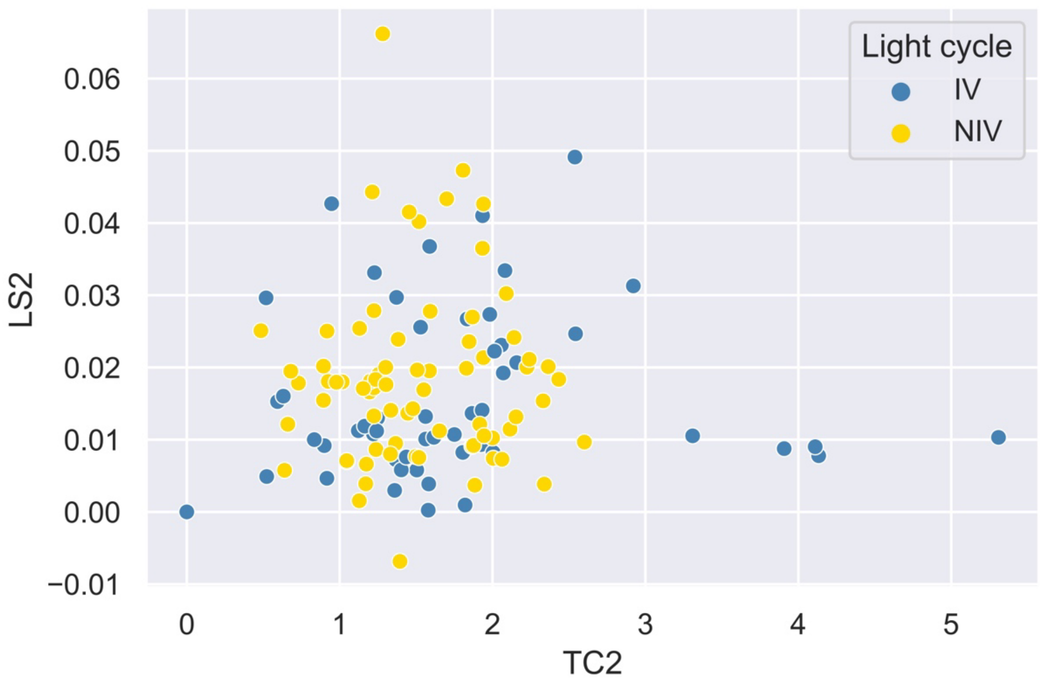
Note. The effect of TC1 on LS3 was significant in this analysis. The effect of the light cycle and the interaction effect between TC1 and the light cycle on learning speed were not significant.



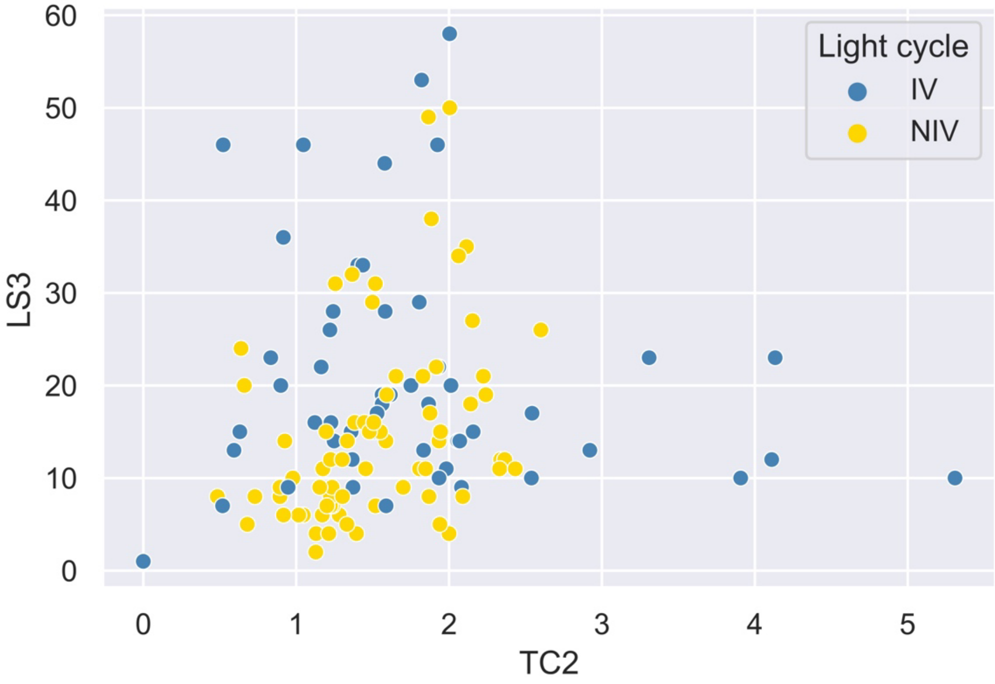
Note. There were no significant effects in this analysis



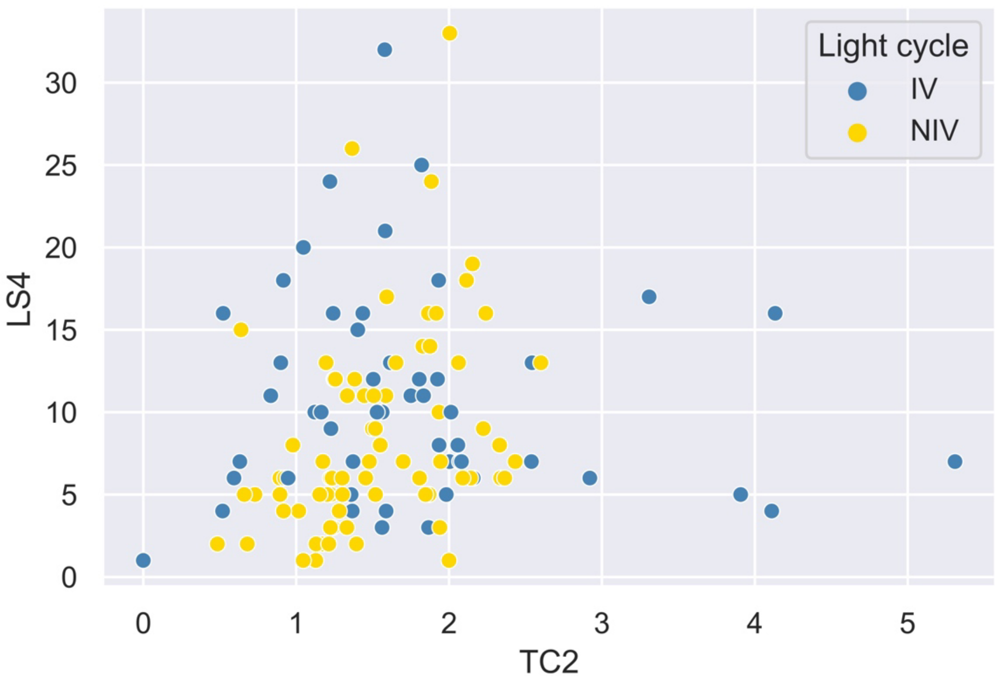
Note. There were no significant effects in this analysis



Note. There were no significant effects in this analysis



Note. The effects of TC2, light cycle and the interaction between the two were significant in this analysis



Note. The effects of the light cycle and the interaction effect between TC2 and the light cycle were significant in this analysis. The effect of TC2 on learning speed was not significant.

The results of the mixture models for TC1 show that none of the direct or interaction effects were significant, except for the effect of TC1 on LS3 (Table 10). For TC2, a few more effects were found to be significant (Table 11). The model predicting LS3 showed significance for the light cycle, TC2 and the interaction between the light cycle and TC2. Additionally, the model predicting LS4 demonstrated significance for TC2 and the interaction between the light cycle and TC2. It is remarkable that LS1 and LS2 were not explained by any of the models in this study, and were also not the best measurements for learning speed. Furthermore, TC1 was only significant in one model (explaining LS3), while TC2 was significant in both LS3 and LS4. Although we should be cautious drawing conclusions, there appears to be some interaction effect between the light cycle and training time consistency.

Table 10

Mixture models for light cycle and time consistency 1

Learning speed variable	Coefficient	Coefficient value	*t*	*p* value
LS1	Light cycle	-0.00	-0.10	0.917
	TC1	0.01	0.78	0.438
	Light cycle * TC1	-0.01	-0.49	0.625
LS2	Light cycle	-0.00	-0.37	0.712
	TC1	0.01	0.45	0.651
	Light cycle * TC1	-0.00	-0.09	0.926
LS3	Light cycle	-2.96	-0.4	0.691
	TC1	-25.21	-2.17*	0.032
	Light cycle * TC1	18.72	1.27	0.205
LS4	Light cycle	-1.84	-0.45	0.657
	TC1	-12.35	-1.97.	0.052
	Light cycle * TC1	8.61	1.08	0.281

$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$

Table 11

Mixture models for light cycle and time consistency 2

Learning speed variable	Coefficient	Coefficient value	*t*	*p* value
LS1	Light cycle	-0.01	-1.37	0.182
	TC2	-0.00	-0.83	0.407
	Light cycle * TC2	0.00	1.01	0.316
LS2	Light cycle	-0.01	-1.42	0.169
	TC2	-0.00	-1.22	0.225
	Light cycle * TC2	0.00	1.20	0.234
LS3	Light cycle	16.46	2.57*	0.017
	TC2	6.68	2.45*	0.016
	Light cycle * TC2	-7.42	-2.38*	0.019
LS4	Light cycle	7.19	2.03.	0.054
	TC2	3.38	2.28*	0.024
	Light cycle * TC2	-3.49	-2.06*	0.041

*$p < .05$. **$p < .01$. ***$p < .001$

Discussion

This study aimed to explore the reasons behind the variations in learning speed among the mice and laboratories in the decision-making task experiments performed by the International Brain Laboratory et al. (2021). We evaluated the effects of a light cycle and training time consistency on the learning speed of mice. First, quantifications for learning speed and training time consistency were created, which resulted in four definitions for learning speed and two for training time consistency. After that, mixture analyses were performed to test the effects of light cycle and training time consistency. Although no significant effect was found when light cycle and training time consistency were analysed separately, the mixture models incorporating both variables and an interaction effect showed some significant results. These results, however, were not found in models predicting LS1 or LS2, which may not be the best measures of learning speed. Instead, LS3 and LS4, which are not dependent on the total number of sessions, might provide a more accurate representation of learning speed.

Although the study found some interaction effects between light cycle and training time consistency, these effects were not consistently found in all analyses. This could be due to the definitions of the variables. Therefore, further research should focus on finding the most suitable quantification of learning speed and training time consistency. One possible approach is to evaluate the performance of mice in the first five sessions, as the International Brain Laboratory et al. (2021) found that early performance can predict later performance to some extent. Another limitation of the current study was that, in the mixture analysis, only random intercepts were added to correct for lab-specific variables. It would be beneficial to assign a random effect to all effects, but this would require a significant amount of data, and it is unclear whether this approach would converge. Additionally, investigating the impact of time of day on learning performance could be interesting, as this factor may interact with both light cycle and training time consistency. It may also be worth exploring whether neural sensitivity to rewards could explain some variance in the current study, as previous research has shown that this factor fluctuates during the day. In the study of Byrne et al. (2017) neural reward activation followed a quadratic time-of-day effect, with higher activation at 10.00 AM and 7.00 PM compared to 2.00 PM in a validated reward task with healthy young men. The effect of time of day on learning was dependent of the neural sensitivity to rewards. In the decision-making tasks of the International Brain Laboratory (2021) rewards play a large role in the learning of mice. When the mice respond correct, they receive a reward (water with sugar) (The International Brain Laboratory et al. 2021), which stimulates that behaviour. If the results of Byrne et al. (2017) can be applied to the mice in the study of the International Brain Laboratory et al. (2021), it may explain some of the variance in learning speed.

The results of this study will contribute to a deeper understanding of the circadian rhythm and its impact on decision-making tasks in mice. If robust effects are found in future research, it may be worthwhile to investigate whether similar effects could be observed in humans. This knowledge could be valuable in fields like education and training, where it could lead to increased productivity. The study also supports the efforts of the International Brain Laboratory et al. (2021) to tackle the reproducibility crisis in psychology and provides insight into the differences found across the seven laboratories that participated in the International Brain Laboratory's study.

References

Baddeley, A., Hatter, J. Scott, D. & Snashall, A. (1970). Memory and time of day. *Quarterly Journal of Experimental Psychology, 22*(4), 605-609.

Bains, R. S., Wells, S., Sillito, R. R., Armstrong, J. D., Cater, H. L., Banks, G., & Nolan, P. M. (2018). Assessing mouse behaviour throughout the light/dark cycle using automated in-cage analysis tools. *Journal of neuroscience methods*, *300*, 37–47.

Bates, D., Maechler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Brauer, M., & Curtin, J. (2018). Linear Mixed-Effects Models and the Analysis of Nonindependent Data: A Unified Framework to Analyze Categorical and Continuous Independent Variables that Vary Within-Subjects and/or Within-Items. *Psychological methods, 23*(3), 389 – 411.

Beeler, J. & Prendergast, B. (2006). Low amplitude entrainment of mice and the impact of circadian phase on behavior tests. *Physiology & Behavior, 87*(5), 870-880.

Blake, M. (1969). Time of Day Effects on Performance in a Range of Tasks. *Psychonomic Science, 9(*6), 349-350.

Byrne, J., Hughes, M., Rossell, S., Johnson, S. & Murray, G. (2017). Time of Day Differences in Neural Reward Functioning in Healthy Young Men. *Journal of Neuroscience, 37*(37), 8895-8900.

Chen, R., Seo, D., Bell, E., von Gall, C. & Lee, G. (2008). Strong Resetting of the Mammalian Clock by Constant Light Followed by Constant Darkness. *The journal of Neuroscience, 28*(46), 11839–11847.

Daan, S. (2000). Learning and circadian behavior. *Journal of biological rhythms*, *15*(4), 296-299.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). *Journal of statistical software, 67*(1), 1-48.

Elghoul, Y., Frikha, M., Abdelmlak, S., Chtourou, H., Dammak, K., Chamari, K. & Souissi, N. (2014). Time-of-day effect on dart-throwing performance and the perception of the difficulty of the task in 9–10 -year-old boys. *Biological Rhythm Research, 45*(4), 523-532.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika, 10(*4), 507-521.

Fisher, R. A. (1926). The arrangement of field experiments. *Journal of the Ministry of Agriculture of England, 33*, 503-513.

Fisk, A. S., Tam, S. K. E., Brown, L. A., Vyazovskiy, V. V., Bannerman, D. M., & Peirson, S. N. (2018). Light and cognition: Roles for circadian rhythms, sleep, and arousal. *Frontiers in Neurology*, 9,1–18.

Guo, Z., Hires, S., Li, N., O'Connor, D., Komiyama, T., et al. (2014). Procedures for Behavioral Experiments in Head-Fixed Mice. *PLoS ONE, 9*(2), 1-16.

Hines, C. (2002). Relationships among time of day, task performance, and motivation in parochial high school students. *Southern Illinois University at Carbondale ProQuest Dissertations Publishing, 3065358*.

Holloway, F.A. & Wansley, R.A. (1973). Multiple retention deficits after active and passive avoidance learning. *Behavioral Biology,* (9), 1-14.

Ingram, K., Ahmet A., Kwon, So., Woods, K., Escobar, S., Gordon, M., Smith, I., Bearden, N., Filipowicz, A. & Jain, K. (2016). Molecular insights into chronotype and time-of-day effects on decision- making. *Scientific reports, 6*(29392).

Jolly, (2018). Pymer4: Connecting R and Python for Linear Mixed Modeling. *Journal of Open Source Software, 3*(31), 862.

Jones, E., Oliphant, T., Peterson, P., et al. (2001). SciPy: Open source scientific tools for Python.

Jürgen A. Ripperger, J.A., Jud, C., & Albrecht, U. (2011). The daily rhythm of mice. *Febs Letters*(585), 1384–1392.

Kerkhof, G. (1985). Inter-individual differences in the human circadian system: a review. *Biological Psychology, 20*(2), 83-112.

Kruskal, J. B., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association, 47*(260), 583-621.

LeGates, T. & Altimus, C. (2011). Measuring Circadian and Acute Light Responses in Mice using wheel Running Activity. *Journal of Visualized Experiments,* (48).

Li, M., Mai, Z., Yang, J., Zhang, B. & Ma, N. (2020). Ideal Time of Day for Risky Decision Making: Evidence from the Balloon Analogue Risk Task. *Nature and Science of Sleep, 12*, 477-486.

Nakajima, R., Hagihara, H., & Miyakawa, T. (2021). Similarities of developmental gene expression changes in the brain between human and experimental animals: rhesus monkey, mouse, Zebrafish, and Drosophila. *Molecular brain, 14*(1), 135-135.

Radetsky, L., Rea, M., Bierman, A., & Figueiro, M. (2013). Circadian Disruption: comparing humans with mice. *Chronobiology International, 30*(8), 1066-1071.

Roedel, A., Storch, C., Holsboer, F., & Ohl, F. (2006). Effects of light or dark phase testing on behavioural and cognitive performance in DBA mice. *Laboratory Animals, 40*(4), 371–381.

Student (W. S. Gosset). (1908). *The probable error of a mean. Biometrika, 6*(1), 1-25.

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. *Proceedings of the 9th Python for Scientific Computing Conference (SciPy 2010),* 51-56.

Schmidt, C., Collette, F., Cajochen, C. & Peigneux, P. (2007). A time to think: Circadian rhythms in human cognition. *Cognitive Neuropsychology, 24*(7), 755-789.

Stroebel, C.F. (1967). Behavioral aspects of circadian rhythms. *Comparative Psychopathology: Animal and Human*, 158-172.

The international Brain Laboratory. (2020). IBL Protocol for Mice Training. *Behavior, Appendix 2*.

The International Brain Laboratory. (2021). Standardized and reproducible measurement of decision-making in mice. *eLife,* (10), 1-28.

Urai, A., Aguillon-Rodriguez, V., Laranjeira, I., Cazettes F., The International Brain Laboratory, Mainen, Z., Churchland, A. (2021). Citric Acid water as an alternative to water restriction for high-yield mouse behavior. *eNeuro, 8(1), 1-8.*

Valentinuzzi, V. S., Menna-Barreto, L., & Xavier, G. F. (2004). Effect of circadian phase on performance of rats in the Morris water maze task. *Journal of Biological Rhythms, 19*(4), 312–324.

Appendix 1

Session information:

- *Subject_uuid*: the subject id of the mouse
- *Session_start_time*: the date and time of the start of the session
- *User_name*: the username of the experimenter
- *N_trials*: the number of trials in that session
- *Performance_easy*: the performance in the task
- *Threshold*: the estimated contrast threshold of that mouse
- *Bias*: tendency of a mouse to consistently go for one choice
- *Lapse_low*: the likelihood of the mouse making a mistake with low contrast
- *Lapse_high*: the likelihood of the mouse making a mistake with high contrast
- *Training_status*: the training status of the mouse (in training/ trained_1a/ trained_1b)
- *Session_duration*: the duration of the session

Subject information:

- *Subject_uuid*: the subject id
- *Lab_name*: the name of the laboratory
- *Lab_light_cycle*: the light cycle (inverted/ non-inverted) [unknown at first, see procedure]
- *Subject_nick_name*: the nickname of the mouse
- *Sex:* the sex of the mouse
- *Subject_birth_date*: the birth date of the mouse
- *Time_zone*: the time zone in which the laboratory and the mouse were situated