# The Object Space Task reveals increased expression of cumulative memory in a mouse model of Kleefstra syndrome

Evelien H.S. Schut, Irene Navarro Lobato, Alejandra Alonso, Steven Smits, Mehdi Khamassi, Anumita Samanta, Moritz Negwer, Nael Nadif Kasri, Lisa Genzel

Neurobiology of Learning and Memory 173 (2020) 107265

Contents lists available at ScienceDirect

Neurobiology of Learning and Memory

journal homepage: www.elsevier.com/locate/ynlme

# The Object Space Task reveals increased expression of cumulative memory in a mouse model of Kleefstra syndrome

Evelien H.S. Schut[a,b], Alejandra Alonso[a], Steven Smits[a], Mehdi Khamassi[c], Anumita Samanta[a], Moritz Negwer[b], Nael Nadif Kasri[b], Irene Navarro Lobato[a], Lisa Genzel[a,b,*]

[a] Donders Institute for Brain Cognition and Behaviour, Radboud University, Nijmegen, Netherlands
[b] Department of Human Genetics and Department of Cognitive Neuroscience, Radboudumc, Donders Institute for Brain, Cognition and Behaviour, 6500 HB Nijmegen, Netherlands
[c] Institute of Intelligent Systems and Robotics, Sorbonne Université, CNRS, Paris, France

ARTICLE INFO

*Keywords:*
Cumulative memory
Semantic
Autism
Kleefstra Syndrom
*Ehmt1*$^{+/-}$

ABSTRACT

Kleefstra syndrome is a disorder caused by a mutation in the *EHMT1* gene characterized in humans by general developmental delay, mild to severe intellectual disability and autism. Here, we characterized cumulative memory in the *Ehmt1*$^{+/-}$ mouse model using the Object Space Task. We combined conventional behavioral analysis with automated analysis by deep-learning networks, a session-based computational learning model, and a trial-based classifier. *Ehmt1*$^{+/-}$ mice showed more anxiety-like features and generally explored objects less, but the difference decreased over time. Interestingly, when analyzing memory-specific exploration, *Ehmt1*$^{+/-}$ show increased expression of cumulative memory, but a deficit in a more simple, control memory condition. Using our automatic classifier to differentiate between genotypes, we found that cumulative memory features are better suited for classification than general exploration differences. Thus, detailed behavioral classification with the Object Space Task produced a more detailed behavioral phenotype of the *Ehmt1*$^{+/-}$ mouse model.

## 1. Introduction

Most if not all mental disorders are accompanied by memory deficits, with the quality of the deficit depending on the overlap between the underlying circuit needed for the respected memory type and the circuit affected by the disorder. For example, semantic memories depend on cortical structures such as the medial prefrontal cortex and anterior cingulate cortex, while episodic memories are thought to rely on intact hippocampal functioning (Genzel & Wixted, 2017; Moscovitch, Cabeza, Winocur, & Nadel, 2016; Squire, Genzel, Wixted, & Morris, 2015). Thus, larger deficits in episodic in contrast to semantic memories are expected in disorders affecting the hippocampus more than cortex. This also means that detailed characterization of the memory deficit can help predict, which circuits are affected by a disorder and lead future imaging or molecular investigations.

In animal models, it is harder to distinguish between semantic and episodic memories and in addition most commonly one-trial memory paradigms are used to assess memory deficits in disease. Such tasks, e.g. contextual fear conditioning or simple object in place memory, mainly test hippocampal processing (Squire et al., 2015). Further, the one-trial nature of these tasks make it impossible to distinguish if animals are using memory systems that would compare to human semantic or episodic memory. However, critical for human cognition are semantic memories representing our knowledge of the world accumulated over longer time periods not just one-trial memory events. Semantic or cumulative memory is not tested by one-trial tasks and thus rarely assessed in animal models of disease. We recently developed a new task – the Object Space Task – that addresses this deficit: across different conditions both simple memories (potentially episodic-like) as well as abstracted, cumulative memories (potentially semantic-like) are tested (Genzel et al., 2019). The task is based on a rodent's tendency of natural exploration of new objects in an open-field environment. In the key condition of this task – overlapping – spatial configurations with two objects are presented to the animal over multiple trials per day, for four consecutive days. This allows the animal to accumulate information over time in order to construct an abstracted, cumulative and perhaps semantic-like memory across training days. The additional advantage of this task is that it allows behavioral characterization beyond the memory measure, such as general movement patterns within an open field.

Monogenetic causes of neurodevelopmental disorders are a molecular entry point in understanding underlying mechanism and circuits

* Corresponding author at: Donders Institute for Brain Cognition and Behaviour, Radboud University, Nijmegen, Netherlands.
*E-mail address:* l.genzel@donders.ru.nl (L. Genzel).

https://doi.org/10.1016/j.nlm.2020.107265
Received 24 March 2020; Received in revised form 6 May 2020; Accepted 30 May 2020
Available online 09 June 2020

of cognition. Kleefstra syndrome is a neurodevelopmental disorder characterized in humans by general developmental delay, severe intellectual disability and autism (Kleefstra, 2005; Kleefstra et al., 2006; Kleefstra et al., 2012; Kleefstra, Schenck, Kramer, & van Bokhoven, 2014), caused in most cases by haploinsufficiency of the *EHMT1* gene (Euchromatic Histone Methyltransferase 1). Due to the moderate to severe intellectual disability and accompanied diminished language capabilities, memory performance has not been tested in human subjects so far (Kleefstra et al., 2012; Stewart & Kleefstra, 2007; Willemsen et al., 2011). Previous studies have shown that the heterozygous *Ehmt1* knock out mouse (*Ehmt1*$^{+/-}$) recapitulates the core features of Kleefstra syndrome and the mouse model has been used to asses effects on memory (Balemans et al., 2010; Balemans et al., 2013; Benevento et al., 2016; Iacono et al., 2018) and thus potential contribution of this pathway to the intellectual disability seen in humans. Convincing evidence for the role of EHMT1 in learning and memory comes from a study in Drosophila that identified EHMT as a key regulator of classic learning and memory genes and mutant flies that did not express EHMT showed deficits in non-associative learning and courtship memory (Kramer et al., 2011).

The protein encoded by *EHMT1* (EHMT1 or GLP (G9a-like protein) acts as a histone methlystransferase, i.e. an epigenetic regulator. EHMT1 catalyzes mono- and dimethylation of histone H3 at lysine 9 (H3K9me2) (Shinkai & Tachibana, 2011), thereby working as repressor of transcription. The mouse homolog has been shown to specifically regulate the expression of several activity-dependent genes in the hippocampus following fear conditioning (Gupta-Agarwal et al., 2012). Furthermore, it has been shown to be critically involved in homeostatic synaptic scaling *in vitro* and in the developing visual cortex *in vivo* (Benevento et al., 2016; Martens et al., 2016).

*Ehmt1*$^{+/-}$ mice recapitulate core features of the human phenotype: they show delayed postnatal development and facial and cranial abnormalities that correspond to the phenotype observed in human patients (Balemans et al., 2014). At a behavioral level, *Ehmt1*$^{+/-}$ mice display reduced exploration, increased anxiety when exposed to novel environments, and impaired social behavior (Balemans et al., 2010). In a fear conditioning task, *Ehmt1*$^{+/-}$ mice show increased freezing and decreased extinction, another indication for increased anxiety in this mouse model (Balemans et al., 2013). *Ehmt1*$^{+/-}$ mice perform similar to their wildtype controls in the Barnes-maze, a task that uses anxiety to the open field as a motivator, but show a deficit in simple object recognition or placement memory (Balemans et al., 2013). Interestingly, *Ehmt1*$^{+/-}$ mice outperform wildtypes in a touch-screen pattern separation task, a process that heavily depends on the dentate gyrus in the hippocampus (Benevento et al., 2017). Similar results on pattern separation have been demonstrated in human individuals with high-functioning autism (Plaisted, O'Riordan, & Baron-Cohen, 1998). In contrast, Kleefstra Syndrome in humans is classically associated with severe learning disabilities. Earlier studies have shown impaired hippocampal-dependent learning in *Ehmt1*$^{+/-}$ mice and impaired hippocampal physiology, including reduced excitatory connectivity between CA3 and CA1 (Balemans et al., 2013). While impairments in hippocampal-dependent memory may to some extent reflect the episodic memory impairments in humans, the increased anxiety-like behavior in *Ehmt1*$^{+/-}$ mice may have confounded performance in hippocampal memory paradigms in previous studies. In addition, to our knowledge, cumulative learning abilities in these animals have not yet been tested in those mice.

Thus, to further characterize both cumulative and simple memory processes in *Ehmt1*$^{+/-}$ mice, we assessed performance of these animals in the Object Space Task (Genzel et al., 2019). The task contains both an simpler control as well as the key cumulative memory condition. In addition to the conventional behavioral analysis, we modeled the learning behavior over training days as well as build a classifier for individual trial behaviors automatically extracted from the videos with in-house deep learning networks based on DeepLabCut (Mathis et al., 2018). *Ehmt1*$^{+/-}$ mice showed overall decreased exploration of the objects and more corner sitting but expressed increased cumulative memory compared to controls. In contrast, *Ehmt1*$^{+/-}$ mice showed reduced simple memory in our stable condition. Computational modelling revealed that the difference in cumulative memory stems from a change in memory expression and not learning rate. Finally, using the classifier we showed that behavioural measures from video analysis of individual trials allow the automatic identification of genotype. The classifier performed best in our cumulative memory condition using memory-specific features (discrimination index). Thus we could show that our Object Space Task allows for the in-depth characterization of innate behavior as well as contrast simple and cumulative memories in animal models of disease that can guide future circuit investigations.

## 2. Methods

### 2.1. Subjects

Male wildtype and Ehmt1$^{+/-}$ mice (WT n = 31, *Ehmt1*$^{+/-}$ n = 24, littermates, bred in-house), 12–16 weeks of age at the start of behavioral training were group housed with ad libitum access to food and water. Mice were maintained on a 12hr/12hr light/dark cycle and tested during the light period. In compliance with Dutch law and Institutional regulations, all animal procedures were approved by the Centrale Commissie Dierproeven (CCD) and conducted in accordance with the Experiments on Animal Act. Project number 2016-014 and protocol number 2016-014-015, both approved by the CCD and local animal welfare body (Radboudumc).

### 2.2. Behavioral training

Habituation and behavioral training has been described previously in (Genzel et al., 2019). Briefly, all animals were extensively handled before the start of habituation (https://www.genzellab.com/#/animal-handling/). Mice were habituated to a square arena (75 cm × 75 cm) for five sessions within five days. In the first session, mice were placed in the arena together with all cage mates for 30 min. In the remaining sessions, mice were placed in the arena individually for 10 min. In the final two sessions, two objects (made from Duplo blocks, not used in the main experiment) were placed in the arena and animals were allowed to explore.

Animals were trained on all three conditions: stable, overlapping and random. Condition sequence and locations were counterbalanced among animals and sessions, and the experimenter was blinded to the condition and genotype.

In the random condition – the negative control – the objects are in semi-random configurations across all 21 trials (randomness restricted by counterbalancing). Trial 21 (test) always used the two positions not used in trial 20, to avoid recency guiding exploration behaviour. In the stable condition the two objects always remain in the same positions (trial 1–20) until the final trial (trial 21, test) when one of the objects is moved to a new location. In the stable condition at test the animal can be guided by either the memory of the most recent event or an accumulation of all events, thus is the more simple, control memory paradigm. Finally, the key cumulative memory condition, overlapping, always has one object at the same position, while the other object is placed in one of the three other corners (stable location is counterbalanced across animals). Thus over time and trials, animals build up a cumulative memory resulting in a preference for the less often shown location, which can be observed in trials 2–20 as well as test (trial 21) (Genzel et al., 2019). In addition, at the test trial, the same configuration as the final training trial is used, thus controlling for recency and episodic-like memory at test.

At the beginning of each five-day session, cues were placed on the walls inside the box and at least one 3D cue was placed above one of the other walls. Cue distribution was intentionally non-symmetric. A

camera was placed above the box to record every trial and to allow for online scoring of exploration time with our in-house scoring program, the Object Scorer. In each condition, animals were allowed to explore two objects for 5 min with an inter-trial interval of roughly 30 min (3–5 animals running interleaved). Before the beginning of each sample trial, the box and the objects were thoroughly cleaned with 70% ethanol. Each sample trial consisted of a different pair of identical objects varying in height width, texture and material. Objects were never repeated during the training period of one condition (one session). The test trial, 24 h after the last sample trial, consisted of again two identical objects and animals were allowed to explore for 10 min, however only the initial 5 min were used for analysis. The Object Scorer software was used for online scoring and extraction of exploration times during all trials.

### 2.3. Statistical analysis

We measured three main variables: total exploration time of each object, count of visits to each object and average exploration bout duration. Due to technical reasons, not all animals could be included in the count and bout duration analysis.

General behavioral differences during training were tested with a repeated measure ANOVA with day and session as factors for each variable separately. Additionally, we repeated the analysis but with condition instead of session (orthogonal thus not compatible in one analysis), which showed no significant effects.

To test for memory for all three measures the discrimination index was used calculated as the difference in time exploring the novel object location and stable location divided by the total exploration time. This results in a score ranging from $-1$ (preference for the stable location) to $+1$ (preference for the less stable object location). A score of 0 indicates no preference for either object location. To test for the presence of memory, the discrimination index was test with one-sample *t*-test to chance. In case of non-normality of the data Wilcoxon Signed Rank Test was used.

### 2.4. Model

The same computational model as in (Genzel et al., 2019) was used (see article for more detailed methods). In short, the model learns place-object associations and then translates this memory into an exploratory behavior: the objects that were stably found at the same location have a very low uncertainty and are thus less attractive during exploration than objects found at changing locations (high uncertainty in place-object association). The model employs two different parameters: a learning rate $\alpha$, which determines the speed of memory accumulation; an inverse temperature $\beta$, which determines the strength and sign of memory expression during exploratory behavior.

A low learning rate $\alpha$ (i.e., close to 0) means that the model will need numerous repetitions of the same observation (i.e., in the Object Space Task, many trials observing the same place-object association) to properly memorize it. In contrast, a high learning rate $\alpha$ (i.e., close to 1) means that the model quickly memorizes new observations at the expense of old observations, which are more quickly forgotten. As a consequence, with a low learning rate the exploratory behavior generated by the model will mostly reflect remote memories but not recent ones (cumulative or semantic-like memory). Conversely, with a high learning rate, exploratory behavior in the model will mostly reflect recent memories but not remote ones (recency or episodic-like memory).

Finally, an inverse temperature $\beta$ close to zero means that the model does not strongly translate memories into object preferences for exploration, thus showing little object preference. In contrast, a high inverse temperature will mean that the model's exploratory behavior is strongly driven by differences in relative uncertainty between place-object associations. A high positive inverse temperature will result in neophilic behavior: the model spends more time exploring objects associated with high uncertainty (i.e., novelty or constantly changing location); a high negative inverse temperature will result in neophobic behavior: the model spends more time exploring objects with low uncertainty (stable/familiar objects).

The model was fitted to each mouse's trial-by-trial behavior using a maximum likelihood procedure described in (Genzel et al., 2019). In brief, this model fitting process found the best parameter values for each subject that best explain the relative proportion of time spent exploring each object at each trial.

All model equations are described in (Genzel et al., 2019).

### 2.5. Automatic behavioral analysis

To classify genotype of mice in single trials, two main models were designed. The first model is a video action scoring classifier, upon which the general behavioural descriptions (i.e. features) were based. The second model is a genotype classifier that uses these general behaviours to predict WT/KO based on a single trial in the Object Space Task. For each model, the dataset was split into a 90% training and 10% validation set.

The model that scores mouse behaviours (Object Exploration, Wall Exploration, Corner Sitting) in a video is two-fold, in that one model classifies Object Exploration, whereas the other model classifies Wall Exploration and Corner Sitting. The first model is an inflated deep inception neural network based on Carreira and Zisserman (2017) human kinetic action recognition network. Transfer learning was applied to a restructured version of the neural network and re-trained on videos of mice performing the Object Space Task, labelled by humans for object exploration. Next, the second model uses DeepLabCut (Mathis et al., 2018) to extract limb configuration of mice over a single video. In addition to limb configuration, head direction was calculated as ears pointing to the nose. The limb configuration and head direction were then used in combination with knowledge about the square arena, such as location of walls and corners. Corner Sitting was defined as mean limbs location being near the corner, and Wall Exploration was defined as mean limb configuration being near a wall and head direction towards that wall. The dataset of 910,237 frames was randomly split into a 90% training and 10% test set.

The model that classifies genotype based on single trials, uses the behaviours extracted from the automatic behavioural scoring. To elucidate, behavioural summaries over the 5 min trials were calculated based upon the time-series of actions (see Table 1). These features were then used as input variables for a Random Forest and XGBoost classifier, with genotype as a target variable. The classifiers were trained to optimize the AUROC (Area Under the Receiver Operating Curve), for each subset of data (all trials, stable trials, overlapping trials, random trials). Significance of the AUROC was then tested under the permutation distribution. To assess which behavioural features were driving most of each model in their predictions, feature importance were calculated was drop-column importance. The top features per classifier were taken for additional analysis, to further investigate how behaviours differ between wildtype and Ehmt1$^{+/-}$ mice in single trials per condition. The mean difference was tested under the permutation distribution for each top feature, with genotype as a between-subject factor. The dataset for training the classifier consisted of 1738 unique trials (i.e. rows), each containing 45 behavioural descriptions (i.e. columns). For each mode, the dataset was randomly split into a 90% training and 10% validation set. To train the model, only the training set was used. Since classifier models such as XGBoost and Random Forest depend heavily on their hyper-parameters (e.g. depth of trees), a best model for the training data may be found using grid search with k-fold cross validation as the evaluation approach. Hence, the models were optimized using 10-fold cross validation on the 90% training set. Subsequently, the final model was tested against a 10% validation test that was not seen in any of the training cases.

**Table 1**
Model Features: Features extracted from the sequence data per trial. In total 45 features were calculated (above features for each object (i) and min (n) respectivly thus 13 feature description result 45 specific features).

| Feature | Description |
|---|---|
| First object | Whether the first object explored was either object 1 or object 2 |
| First object latency | Latency to first exploration of any object |
| $Stay_i$ | Likelihood of exploring object i next, after having last explored object i. Where i is either 1 or 2. |
| SteadyState1($SS_1$) | The probability of exploring object 1 after many explorations |
| Perseverance Index | Index ranging from −1 to 1 represent the tendency of switching between objects during exploration as −1, the tendency to reexplore the same object during exploration as 1, and no tendency as 0 |
| n_transitions | Total number of transitions made between objects during the whole trial. Per minute i, the total number of explorations of any object up until that minute |
| mini_n_explore | Per minute i, the total number of explorations of object j up until that minute |
| mini_n_explore_$object_j$ mini_time | Per minute i, the total time any object was explored up until that minute |
| mini_$object_j$_time | Per minute i and object j, the total time that object was explored up until that minute |
| bout_time | Mean time of explorations of any object |
| bout_time_$obj_j$ | Mean time of explorations of object j |
| $_{mini}$_DI | Per minute i, the discrimination index (DI) up until that time representing a preference for exploring object 2 as −1, a preference for exploring object 1 as 1, and no preference as 0 |

## 3. Results

To assess cumulative and the simple, control memory in the *Ehmt1*$^{+/-}$ mouse model, we used the Object Space Task with the conditions stable, overlapping and random (Genzel et al., 2019). In this study, mice (WT n = 31, *Ehmt1*$^{+/-}$ n = 24, eight weeks, male) were first habituated (one week) and then trained (three weeks, each week one session of 20 training and one test trials, five trials a day) in the task (Fig. 1). During the training period the animals went through the three conditions in a counterbalanced order.

### 3.1. General exploration differences: decreased exploration in *Ehmt1*$^{+/-}$ mice

Initially, we focussed our analysis on differences in explorative behaviour and compared total exploration time, total count of object visits, and average exploration bout length (scored manually, experimenters blinded to condition and genotype). *Ehmt1*$^{+/-}$ mice show decreased total exploration time (p = 0.004, detailed statistics in figure legend), however this difference decreased over time (each session is one week, session X gene interaction p = 0.001, Fig. 2). The difference in overall exploration time is explained by a lower number of visits to each object (p = 0.003), while average length of each individual exploration bout did not differ (p = 0.223). Interestingly, over time (both within a week as well as across weeks) bout length increased independent of genotype (p = 0.027), while the number of visits decreased over time in WT. Thus the convergence of total exploration time over sessions is explained by sustained number of visits with increasing bout length in *Ehmt1*$^{+/-}$ and decreased number of visits with increasing bout length in WT. Total exploration time, count of explorations and bout length did not differ significantly across training conditions for either genotype.

To further asses what the animals were doing when they were not exploring objects, we extracted different behaviors in an automatic way (object exploration, wall/cue exploration, corner sitting) from a randomly selected sub-set of the video data (1764 of 3465 trials) with in-house deep-learning networks based on DeepLabCut (Mathis et al., 2018).

The first, in-house model used Kinetic Action Recognition to extrapolate when the mice explored objects. The model was trained on in-lab available action labeled Object Exploration video data. The second model used pose estimation from DeepLabCut to extract the actions wall/cue exploration and corner sitting (because no labeled data are available for these actions). In the end, the predicted actions of both models are frame-wise concatenated for each video trial. The code implementation for this section can be found at https://github.com/Iglohut/autoscore_3d. Automatic behavioral classification confirmed the difference in object exploration time (p = 0.001) and revealed that

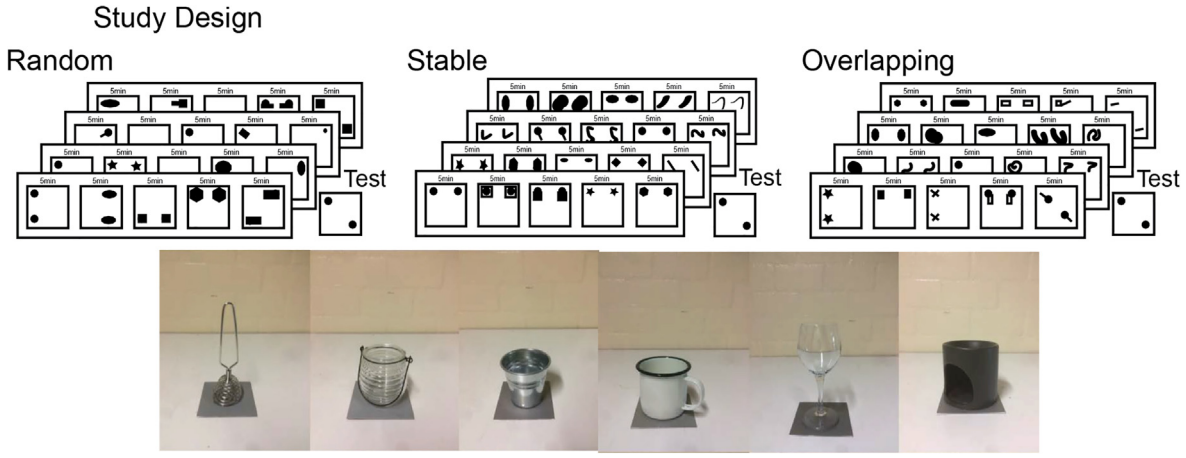


**Fig. 1.** Study Design: Across three weeks animals are trained in three conditions in a counterbalanced order with Mo-Thu training five trials/day (each 5 min) and Fri test (one trial 5 min). The three conditions differ in the underlying statistical regularities of object placement. Random: semi-random (constricted to equal number of occurance across the week) placement; Stable: during training the same two locations, one object moved at test; Overlapping: one location always contains an object, the second object is in one of the other three corners. Final training trial and test trial have the same configuration to control for recenency and episodic-like memory. The stable and moved location identities were counterbalanced across animals to control for general location preferences. Below examples of objects used.

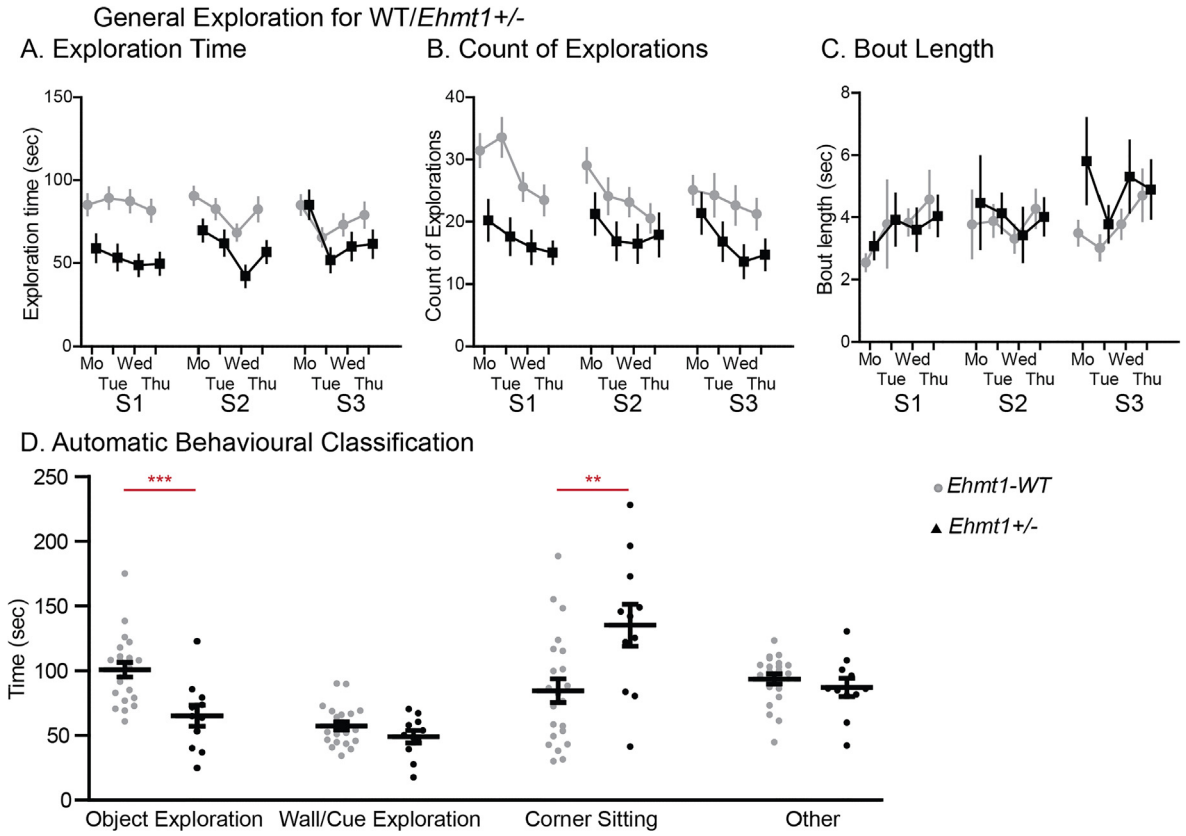


**Fig. 2.** General Exploration for WT/ *Ehmt1*$^{+/-}$ Shown is A. exploration time B. count of exploration bouts and C. average bout length across the three weeks/sessions (S1, S3, S3) of training. Each day averaged across five trials. *Ehmt1*$^{+/-}$ show decreased exploration time due to decreased number of object visits, but this difference decreases over the three week training period. Exploration Time sessionXgene $F_{2,104} = 7.3$ $p = 0.001$ (linear constrast $p = 0.002$), day $F_{3.2,208} = 17.3$ $p < 0.001$ (linear contrast $p < 0.001$), sessionXday $F_{5.5,416} = 4.6$ $p = 0.001$, gene $F_{1,52} = 9.0$ $p = 0.004$; Count session $F_{2,82} = 8.2$ $p = 0.001$ (linear contrast $p = 0.001$), sessionXgene $F_{2,82} = 5.6$ $p = 0.005$ (linear contrast $p = 0.019$), day $F_{2.5,104} = 26.1$ $p < 0.001$ (linear contrast $p < 0.001$), dayXgene $F_{3,123} = 4.0$ $p = 0.009$ (quadratic contrast $p = 0.026$), sessionXday $F_{4.5,185} = 2.2$ $p = 0.048$, sessionXdayXgene $F_{6,246} = 2.5$ $p = 0.022$, gene $F_{1,41} = 9.7$ $p = 0.003$; Bout Length session $F_{2,82} = 3.8$ $p = 0.027$ (linear contrast $p = 0.026$), day $F_{2.5,123} = 3.4$ $p = 0.025$ (linear contrast $p = 0.021$), dayXgene $F_{3,123} = 3.0$ $p = 0.031$ (linear contrast $p = 0.008$), sessionXday $F_{4.5,183} = 4.0$ $p = 0.003$, gene $F_{1,41} = 1.5$ $p = 0.223$. D. shows the different behaviours extracted from the automatic classification of a subset of videos. *Ehmt1*$^{+/-}$ spend less time exploring objects but more time sitting in the corners Behavior $F_{3,93} = 13.9$ $p < 0.001$, BehaviorXgene $F_{3,93} = 8.3$ $p < 0.001$, object exploration $p = 0.001$, corner sitting $p = 0.006$). Data shown as mean and SEM.

*Ehmt1*$^{+/-}$ animals spent more time sitting in corners ($p = 0.006$, Fig. 2D).

In sum, the initial difference in exploration time due to decreased object visits and more corner sitting in *Ehmt1*$^{+/-}$ *mice*, could indicate increased anxiety-like features (reluctance to leave corner for an object visit). This decrease in exploration and increase in sitting has been shown before (Balemans et al., 2010), but we could add that the anxiety decreases with time and habituation over the weeks of training.

### 3.2. Memory specific differences: decreased simple and increased cumulative memory in *Ehmt1*$^{+/-}$ mice

After characterizing general exploration features, we assessed memory performance by calculating the discrimination index for each exploration time, count of explorations and bout length (DI = [moved-stable]/[moved + stable]). In our simple memory condition (stable) only WT were above chance at test (in the discrimination indicies for exploration time, count, bout length) and not *Ehmt1*$^{+/-}$ mice (Fig. 3A). As expected in our negative control condition (random) neither genotype was above chance at test (Fig. 3B). There was no difference in these conditions during the training trials (1–20, see Figure Supplementary Fig. 1).

In our cumulative memory condition – overlapping – we would expect a positive discrimination index during training (trials 1–20) as well as test (trial 21). In contrast to stable with static objects and random, there is an aligned pattern across animals in overlapping already during training and the animals develop a preference for the less-stable location (resulting in a positive DI, Fig. 4 A). Across all trials (repeated measure ANOVA) there was a significant effect of trial ($p = 0.003$) as well as an effect of genotype ($p = 0.03$) but no interaction ($p = 0.47$). At training both WT and *Ehmt1*$^{+/-}$ showed a discrimination index above chance expressed across all three variables (exploration time, count, bout length) but *Ehmt1*$^{+/-}$ showed slightly higher discrimination index values especially during training (only significant $p < 0.05$ for exploration time) indicating increased memory expression (Fig. 4). At test both WT and *Ehmt1*$^{+/-}$ showed above-chance discrimination index in the overlapping condition, but due to high variability only significantly so ($p = 0.027$) in the count of explorations measure (Fig. 4 C). To ensure that this memory difference was not due to a difference in total exploration time that may represent differences in anxiety levels, we also checked performance in those animals that had this condition in their third week of training when there is no difference in explorative behaviour between genotype (S3, Figure Supplementary Fig. 2). Also here, *Ehmt1*$^{+/-}$ showed an increased discrimination index in comparison to WT.

Thus both WT and *Ehmt1*$^{+/-}$ have cumulative memory expression, with *Ehmt1*$^{+/-}$ showing a slightly stronger effect (overlapping condition). In contrast, only WT express simple memory at test (stable condition). This decrease in simple memory in stable is reminicent of previous findings in a one-trial object recognition and location

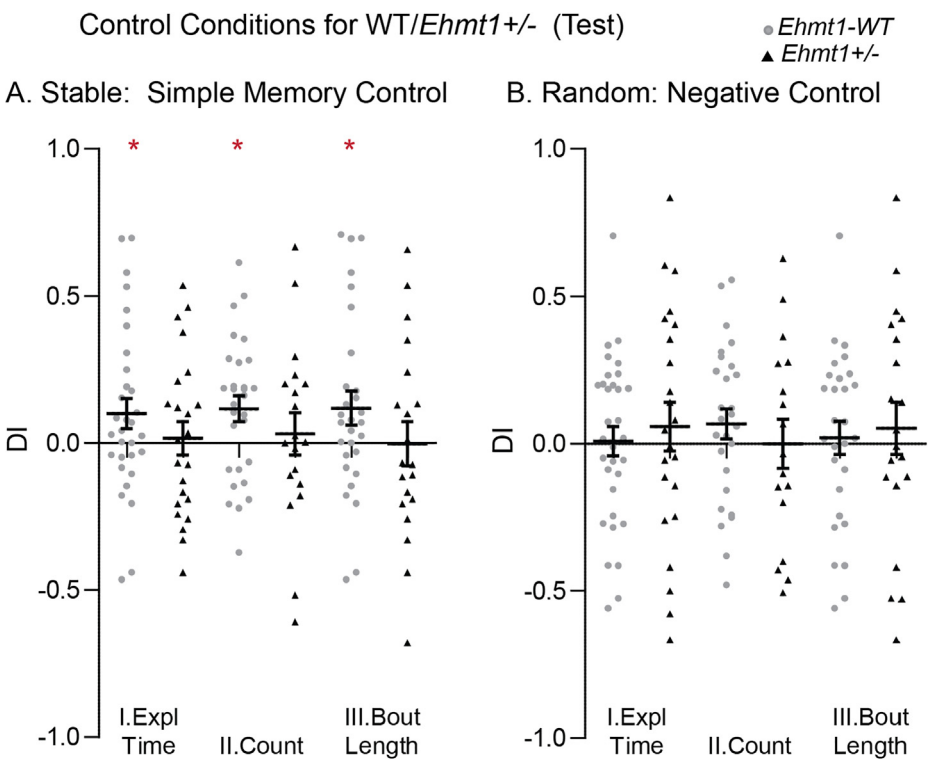


**Fig. 3.** Simple Memory (stable and random conditions): Shown are the discrimination indicies at test (trial 21) calculated for I. exploration time II. count of exploration bouts and II. average bout length for both stable and random. Only WT and not *Ehmt1*$^{+/-}$ performed above chance level in our simple memory test (stable). *T*-Test to chance in stable: *Exploration Time* WT wilcoxon rank test p = 0.03; *Count* WT *t*-test $t_{28}$ = 2.6 p = 0.01; *Bout Length* WT *t*-test $t_{28}$ = 2.1 p = 0.05. Data shown as mean and SEM.

paradigm in these mice (Balemans et al., 2013).

### 3.3. Modelling learning: *Ehmt1*$^{+/-}$ mice show same learning rate as WT but increased cumulative memory expression

The decreased overall exploration seen in the *Ehmt1*$^{+/-}$ mice could have confounded the difference seen in the discrimination index in the overlapping condition. Thus, to further disentangle these effects and to characterize the build-up of a memory trace and its expression, we applied a computational model (Genzel et al., 2019) that progressively learns place-object associations and makes decisions about which proportion of time to spend exploring each object in order to minimize uncertainty about these place-object associations. The model employs two parameters: a learning rate α which determines the balance between recent and remote memories; an inverse temperature β, which determines the balance between neophilic (preference for more novel/uncertain object location, positive values) and neophobic (aversion for more novel/uncertain object location, negative values) exploratory behaviors, and thus measures memory expression. Values around 0 indicate that the behaviour is not driven by memory. We fitted the model separately for each animal and each condition (stable, random, overlapping), in order to observe potential differences in the optimized parameters between conditions. Thus together α and β let us disentangle if animals actually have a better memory and/or learn faster (α) or just express their memory more independent of memory strength (β).

As with the conventional discrimination index, β values took more extreme either positive of negative values in *Ehmt1*$^{+/-}$ animals, which is why we took the absolute value of β as the key parameter for the following analyses: we want to characterize the strength of memory expression independent of neophobic or neophilic tendencies (for number of animals with either positive or negative discrimination index or beta value see Figure Supplementary Fig. 3, there was no significant difference for either type of split between tendency and genotype). Interestingly, *Ehmt1*$^{+/-}$ animals showed significantly higher absolute β values but only in the overlapping condition (Kruskal-Wallis, Chi2 = 4.09, p = 0.043, Fig. 5A). In contrast, learning rate α did not differ between genotype for any condition (Fig. 5B).

Next, we reran the model only within a day (bins of five trials) to characterize the development over the week of training. Memory expression (absolute β) decreased over the week (Fig. 5C), which may explain the decreased memory expression at test seen in the discrimination index of total exploration time (Fig. 4). However, the learning rate remained constant (Fig. 5D).

To conclude, our model shows that WT and *Ehmt1*$^{+/-}$ learned at a similar rate. Consequentially, differences in discrimination index are based on increased memory expression in *Ehmt1*$^{+/-}$. However, this difference is specific to cumulative memories that show underlying statistical regularities as tested in our overlapping condition.

### 3.4. Automatic behavioural scoring and classifier for WT/ Ehmt1$^{+/-}$

On a per-session level including all 21 trials, we have shown that Ehmt1$^{+/-}$ mice explore objects less and show increased memory expression, which is specific to our cumulative memory condition. Next, we analyzed discrete behaviors on a trial-by-trial level: We trained two classifiers on multiple behavioral features extracted from the video data automatically with deep-learning methods as explained in the Methods section. We included both general exploration features (e.g. min explore time) as well as memory specific features (e.g. discrimination index, total 45 features, see Table 1). For each model, the dataset was split into a 90% training and 10% validation set.

Extracted features were fed into two classifiers: Random Forest and XGBoost. To test whether the AUROC (Area Under the Receiver Operating Curve, ROC) performance metric of each model was above the expected value of chance, both models were tested under the permutation distribution (see Methods section). Performance of each model for each condition (all pooled, overlapping, stable, random) are depicted in Fig. 6A. All models except the ones in the random condition have an AUROC > 0.61 that is significant (∀i:pi < 0.001) under the permutation distribution. With the classifier performing at a similar level when only including the data of the overlapping condition in itself as using all data from all conditions, despite the data-set being only 1/3 of the size of all pooled.

Furthermore, inspection of each significant model's ROC-curve

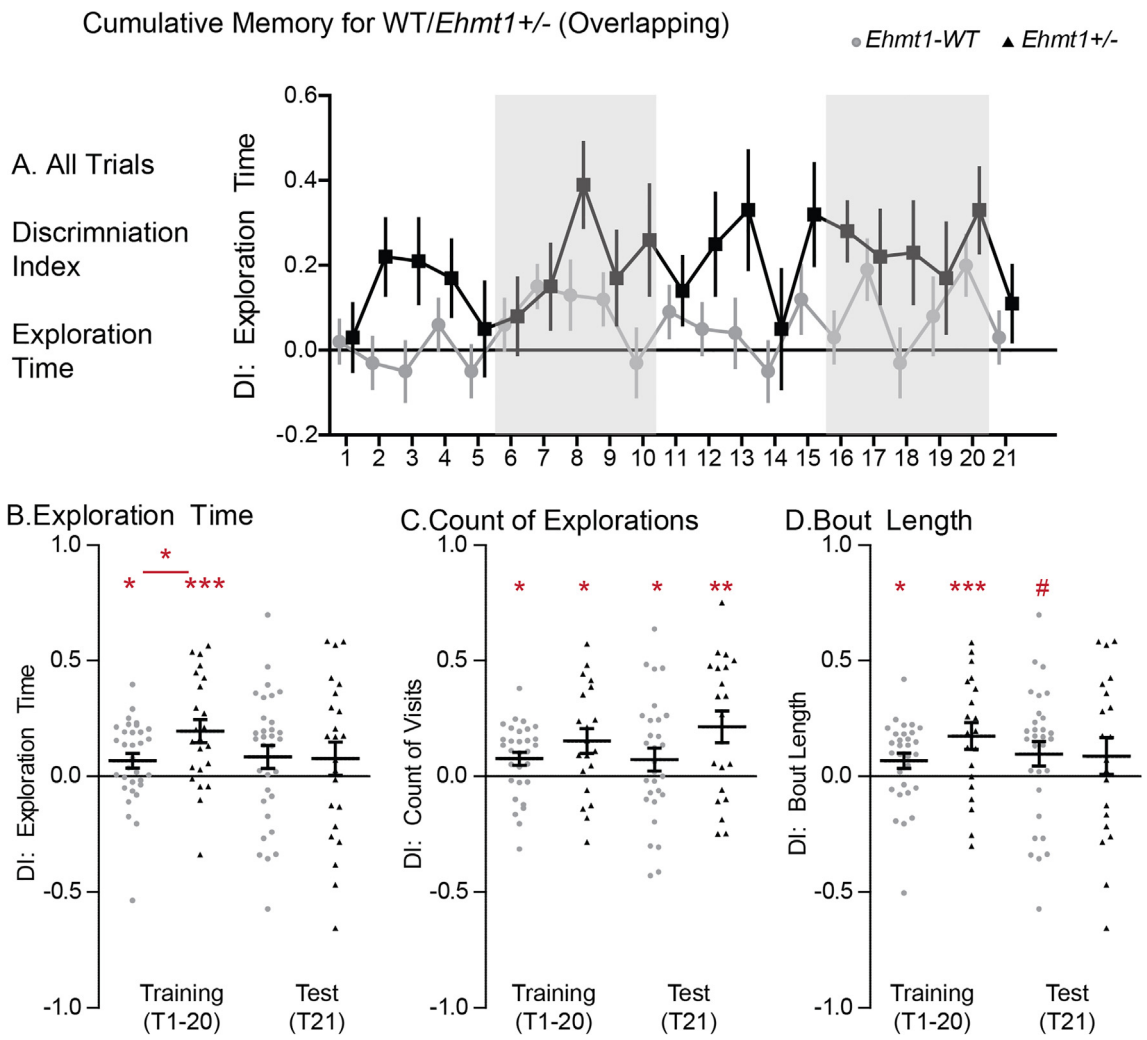


**Fig. 4.** Cumulative Memory (Overlapping condition): Discrimination Index for WT/ *Ehmt1*$^{+/-}$. Shown are the discrimination indecies for training (trials 1–20) and test (trial 21) calculated on A. exploration time for each trial seperatly (trial $F_{12.97,687.6} = 2.44$ p = 0.003, trialXgenotype $F_{12.97,687.6} = 0.1$ p = 0.46, genotype $F_{1,53} = 4.97$ p = 0.03), and averaged across training B. exploration time, C. count of exploration bouts, and D. average bout length. *T*-test *Ehmt1*$^{+/-}$ vs WT overlapping pt $t_{53} = 2.3$ p = 0.027; Test to chance Training: *Exploration Time* WT wilcoxon rank test p = 0.02; *Ehmt1*$^{+/-}$ *t*-test $t_{22} = 4.0$ p = 0.0007; *Count* WT *t*-test $t_{26} = 2.7$ p = 0.01; *Ehmt1*$^{+/-}$ *t*-test $t_{19} = 2.8$ p = 0.01; *Bout Length* WT wilcoxon rank test p = 0.015; *Ehmt1*$^{+/-}$ *t*-test $t_{19} = 3.0$ p = 0.008; Test to chance Test: *Count* WT *t*-test $t_{25} = 2.3$ p = 0.03; *Ehmt1*$^{+/-}$ *t*-test $t_{19} = 3.1$ p = 0.006; *Bout Length* overlapping WT *t*-test $t_{26} = 1.9$ p = 0.066; Data shown as mean and SEM, grey shading in I for individual training days.

shows that all models have relatively high specificity relative to their sensitivity (sensitivity ≈ 0.35–0.49, specificity ≈ 0.87–0.89). For the models trained on trials in the random condition, neither model (Random Forest, AUROC ≈ 0.56, p > 0.05 XGBoost, AUROC ≈ 0.59, p > 0.05) could predict the genotype based on single trials.

To illustrate which features tended to drive the classifiers performance, the top ten features for each model (including all conditions or just subconditions stable, random and overlapping) are shown. Their relative importances (Fig. 6B) were tested under the permutation distribution and also compared between genotype (Fig. 6C). A high relative feature importance means that a model uses that specific feature and all its potential interactions with other features. Interestingly, while no single feature dominated across all models the top features used in all, stable and random conditions were based on general exploration features such as total number of exploration bouts in the first 2 min of the trial. In contrast, such general exploration features became less weighted in the overlapping condition, in which memory-based features gained importance (e.g. DI at 3 min/4 min/5 min).

The increased weights of memory-based features would further support that while we see genotype-specific differences in general exploration features, differences in cumulative memory expression are more prominent and thus here drive automatic classification of WT and *Ehmt1*$^{+/-}$.

## 4. Discussion

In this study, we used the Object Space Task to characterize the behavioral phenotype of the *Ehmt1*$^{+/-}$ mouse model in detail. The task tests general as well as memory-specific exploration features, and contained both an simple and cumulative memory condition. By combining conventional behavioral analysis with automatic behavioral analysis via deep-learning networks, computational modelling of learning behavior across days, and a trial-by-trial behavioral classifier, we could elucidate a variety of behaviors influenced by genotype and their relative importance.

*Ehmt1*$^{+/-}$ animals showed (1) decreased total exploration time due to decreased number of object visits and increased corner sitting. Over training sessions, both parameters approached WT levels. (2) In contrast to WT, *Ehmt1*$^{+/-}$ were not above chance in the simple memory condition, but (3) did show increased cumulative memory expression. (4) Modelling revealed that *Ehmt1*$^{+/-}$ and WT showed similar learning rates in all conditions, but *Ehmt1*$^{+/-}$ showed increased memory

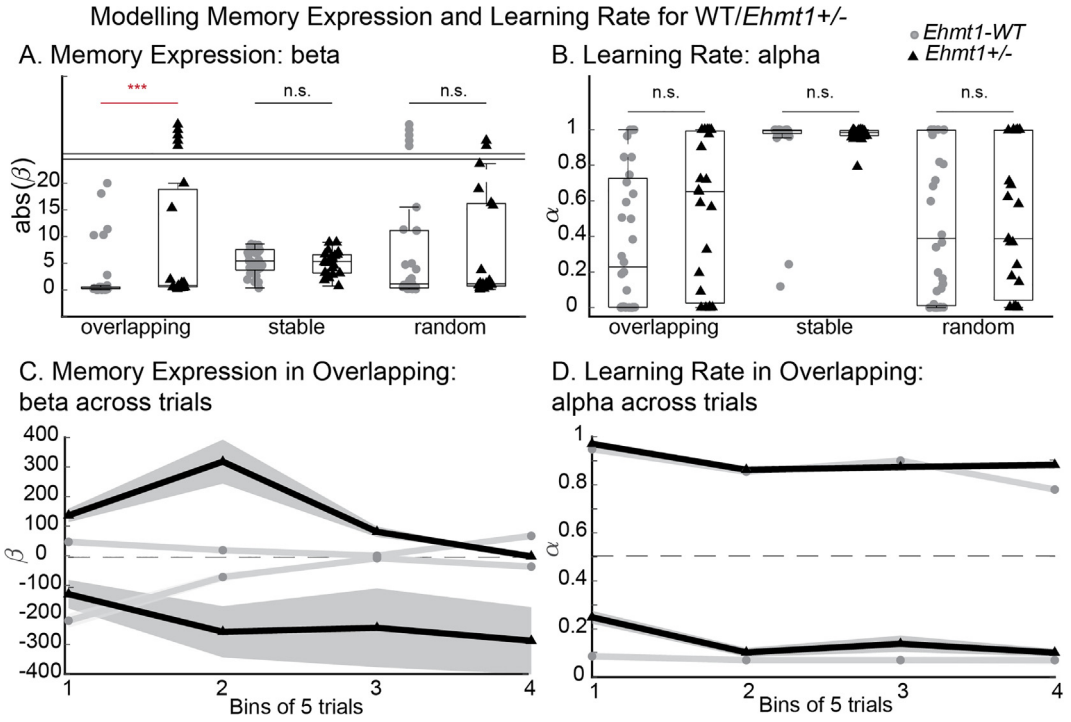


**Fig. 5.** Modelling Memory Expression and Learning Rate: A. Absolute memory expression (β) only showed a significant difference between the two genotypes in the overlapping condition (Kruskal-Wallis, Chi2 = 4.09, p = 0.043, values above 25 stacked above the line). B. No differences between WT and *Ehmt1*$^{+/-}$ was seen in the learning rate (α). Focussing on the overlapping condition memory, memory expression (C) and learning rate (D) plotted by bins of five trials with groups split by genotype and by neophilic/phobic preference (pos/neg β) or by α = 0.5. WT grey, *Ehmt1*$^{+/-}$ black. The grey shading indicates the standard error.

expression (β) in comparison to WT. This difference was specific to the cumulative memory condition (overlapping). (5) Likewise, computational video analysis with two different classifiers could differentiate between *Ehmt1*$^{+/-}$ and WT on a trial-by-trial basis. Relative weights revealed that memory-features in the cumulative condition showed a larger difference and drove the classifier more than general exploration features.

### 4.1. General exploration differences

Initially *Ehmt1*$^{+/-}$ mice visited the objects less frequently and remained more in the corners than WT, perhaps indicating increased anxiety. Balemans et al also showed that *Ehmt1*$^{+/-}$ mice display impaired social behavior as well as reduced exploration and increased anxiety when exposed to novel environments (Balemans et al., 2010). However, the Object Space Task allowed us to add to this finding, in that this increased anxiety seems to habituate over time: in the third week of training (fourth week of box exposure) there was no difference between WT and *Ehmt1*$^{+/-}$. Thus, *Ehmt1*$^{+/-}$ mice do not show persistently increased anxiety, instead they only need more time to habituate to novel situations.

The difference in exploration behavior is sufficient for a classifier to determine genotype on a trial-by-trial basis. Especially the amount of exploration by minute 2 seems to drive the classifier when considering the stable or random condition. In contrast, when determining genotype in the overlapping condition, memory-related behaviors (discrimination index at different time points) outweigh general exploration features. This indicates that while the genotypes differ in general exploration behaviors, their behavior differed even more in cumulative memory condition.

Automatic classification of genotype based on video analysis is becoming more popular and shows great potential for monitoring treatment outcomes in pre-clinical studies. Both 3D and 2D video techniques can be used (Mathis et al., 2018; Wiltschko et al., 2015). Our findings highlight the importance of considering which behaviors are recorded with the video-data: Behaviors with higher cognitive demands (such as our overlapping condition) may be more sensitive to genotype differences as we can show here.

### 4.2. Memory-specific exploration differences

In contrast to WT, *Ehmt1*$^{+/-}$ mice did not perform significantly above chance at test in our simple memory task (stable). This effect is similar to previous reports (Balemans et al., 2013), in which *Ehmt1*$^{+/-}$ mice demonstrated significantly reduced discrimination index compared to wildtype controls in a one-event object location memory test. This memory deficit may be due to hippocampal dysfunction since our simple memory – comparable to classic object location – most likely relies on this circuit (Squire et al., 2015). Other differences in hippocampal function have been found between *Ehmt1*$^{+/-}$ mice and littermate controls, such as increased excitability in CA1 neurons (Balemans et al., 2013; Frega, Linda, Keller, Gümüş-Akay, Mossink, van Rhijn, Negwer, Gunnewiek, Foreman, Kompier, Schoenmaker, van den Akker, Oudakker, Zhou, Kleefstra, Schubert, van Bokhoven, & Kasri, 2019). Of note, in addition to just remembering the last event the stable condition can also be solved with a cumulative memory strategy since the objects stay in the same placement for all trials. However, the dissociation in memory performance seen here with intact cumulative memory expression in overlapping, indicates that our animals did not employ this strategy.

Interestingly, *Ehmt1*$^{+/-}$ expressed a stronger cumulative memory than wildtype controls in the overlapping condition. Initial evidence for this effect came from the conventional behavior analysis and we confirmed by modelling learning rate and memory expression. The learning model also revealed that especially absolute β (memory expression) differed between genotypes and thus the strength of memory expression independent of neophobic and neophilic tendencies. This also explains why conventional analyses only weakly showed this difference in memory expression: *Ehmt1*$^{+/-}$ mice show more extreme discrimination values (closer to −1 and 1) than the WT. With the mix of neophobic

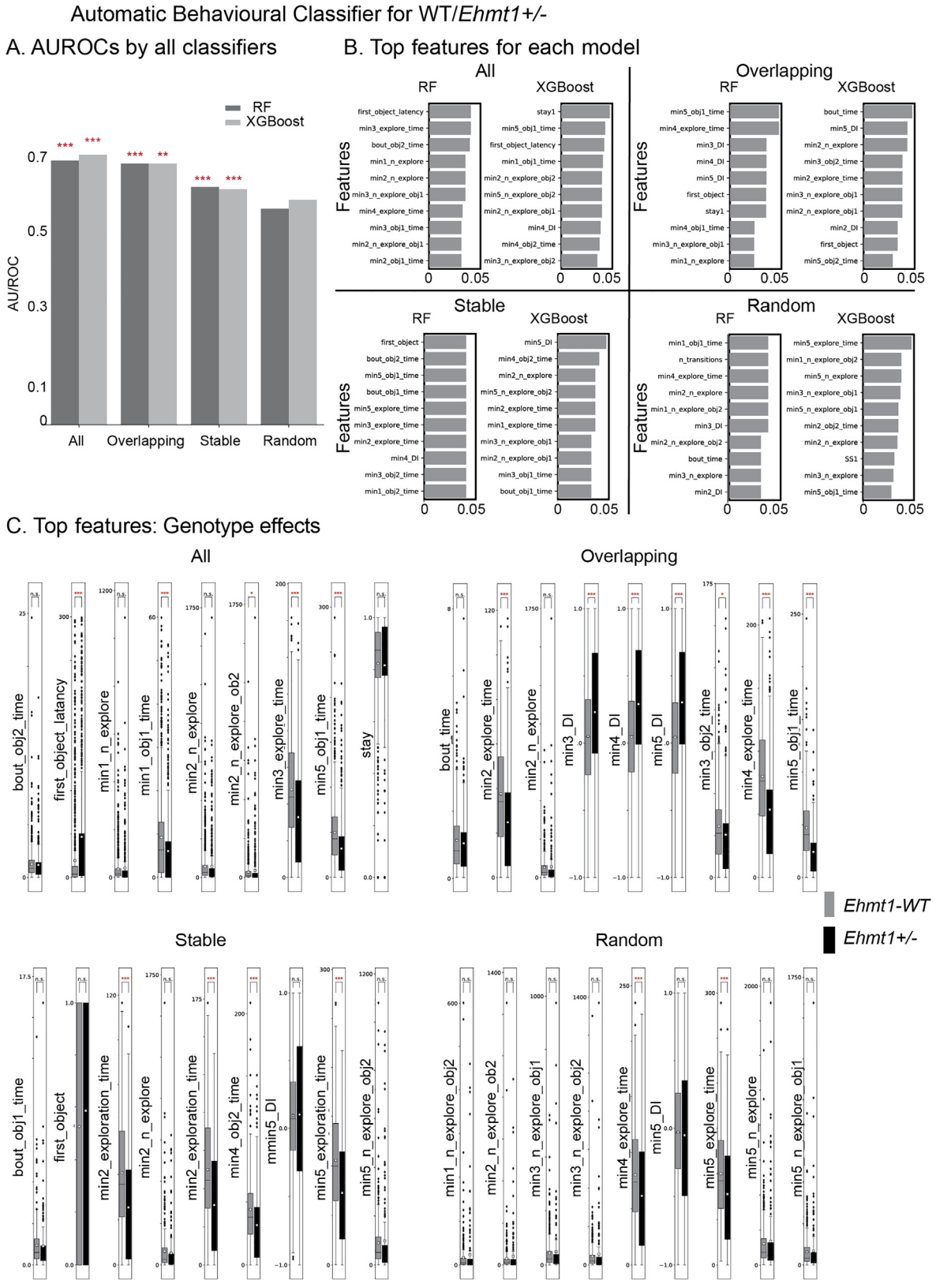


**Fig. 6.** Classifier for WT/ *Ehmt1*$^{+/-}$: A. AUROCs by all classifiers for all subsets of data (all conditions, stable, overlapping, random). Asterisks represent the p-value of its respective AUROC under the permutation distribution: B. Top 10 features used by each classifier and their relative feature importance per model for all subsets of data (all conditions, stable, overlapping, random). For each sub classifier we show the 10 features which dominated performance (thus had the highest weights/ feature importance and drive the discrimination) C. Box plots of the top nine features per model for all subsets of data (all conditions, stable, overlapping, random). While these nine features drove the classifiers, there is no one feature that dominates all or even just one classifier. Some of the features did individually show a significant difference between genotype, however not all of them even though they did contribute to the classification of genotype. This supports the notion that one single variable would not be sufficient for discriminating between genotype, instead a complex interaction is used by the classifier. The white square represents the mean. The difference of the mean between genotype was tested for each feature under the permutation distribution. * = ($p < 0.05$), ** = ($p < 0.01$), *** = ($p < 0.001$).

and neophilic tendencies that is present in both genotypes, the average discrimination index is similar between genotypes, disguising the difference in memory expression. In short, a distribution from −1 to 1 in *Ehmt1*$^{+/-}$ (and −0.5 to 0.5 in WT) will result in overall the same average but difference in standard deviation.

In contrast to memory expression, learning rate did not differ between the genotypes. Thus *Ehmt1*$^{+/-}$ did not simply show better memory or learned faster, instead the same memory strength was just expresed behaviourally more in these animals. Elucidating this difference is another argument why it is important to go beyond conventional analysis of behaviour when performing phenotyping.

Notably, other experiments also showed comparable performance between *Ehmt1*$^{+/-}$ and WT in some behavioral tasks that typically involve the cortex as well as the hippocampus, including spatial learning in the Barnes Maze (Balemans et al., 2013). Additionally, pattern separation learning in a touch-screen task is superior in *Ehmt1*$^{+/-}$ mice (Benevento et al., 2017). It is highly likely that successful extraction of statistical regularities in the overlapping condition of the Object Space task requires both the hippocampus and the neocortex (Moscovitch et al., 2016; Preston & Eichenbaum, 2013; Wang & Morris, 2010).

In sum, the dissociation between a deficit in simple and intact cumulative memory, may hint at a selective hippocampal and not cortical dysfunction in the *Ehmt1*$^{+/-}$ mouse model. Future experiments combining the Object Space Task with neural recordings may help elucidate the underlying mechanism.

### 4.3. Implications for phenotype

Autism is a complex condition characterized by impaired social behavior, perseverant behaviors and communication deficits (Association, 2013). Memory processes in autism are affected as well. Whereas episodic memory deficits have been found consistently (Boucher & Anns, 2018; Boucher, Mayes, & Bigham, 2012), semantic memory abilities and gist extraction may be equal or even superior in individuals with autism in comparison to control subjects (Beversdorf et al., 2000; Gaigg, Bowler, & Gardiner, 2014; Kurz et al., 2019; Parra et al., 2016). This dissociation of deficit across memory type, is reminiscent of our findings in the *Ehmt1*$^{+/-}$ mouse model, where we also found a deficit in simple memory, which may be capturing episodic-like memory deficits, but intact cumulative memory, which may be more semantic-like. Autistic children are known for their increased need to pattern-separate: toys will often be arranged by color and size. Perhaps the increased cumulative memory expression seen in *Ehmt1*$^{+/-}$ mice, which has underlying statistical regularities, is a reflection of this characteristic in mice. Increased anxiety especially in novel situations, is often seen in autism as well. In Ehmt1 mice, this phenotype was expressed in the decreased number of visits to the objects and more corner sitting in the task. The effect alleviated over time, indicative of habituation.

Overall, the more detailed characterization of memory and behavior with the Object Space Task in this model provides initial evidence that the phenotype may be reflecting autism more than intellectual disability as no deficits in cumulative memory were seen. This is in contrast to the human, in which intellectual disability features are more prominent but autism features are also present (Vermeulen et al., 2017). Until now, due to the moderate to severe intellectual disability in humans with Kleefstra syndrome, no specific assessment of memory has been performed which could be compared to our findings in the mouse model. To further classify the autistic features in the mouse model, tasks testing complex social interactions should be employed next.

### 4.4. The object space task for phenotyping

In this study, we used the Object Space Task for detailed behavioral characterization of the *Ehmt1*$^{+/-}$ mouse model. The advantage of the task in comparison to other behavioral assays is that two types of memories are tested in a controlled and comparable setting (simple and cumulative), that due to differences in underlying circuity can show a dissociation of deficits in neurological disorders. In addition to memory specific effects, general exploration, and movement patterns in an open-field environment can also be characterized in this task. When evaluating a behavioral phenotype, it is critical to avoid confounding effects such as increased anxiety or decreased mental flexibility. Thus, one should be cautious with one-trial evaluations of behavior. These factors are controlled for in the Object Space Task and thus more nuanced phenotypes that normally would be occluded by confounding factors can be measured. Finally, we could also show that the task can easily be combined with automatic video analysis, modelling learning behavior as well as a trial-by-trial classifier that allow the in-depth characterization of phenotype beyond conventional behavioral measures.

### 4.5. Caveats

Of note, in this study we only included male mice. Until now no sex differences have been reported for the Kleefstra syndrome, however in general spatial tasks such as likely also the Object Space Task tend to show both sex differences as well as an influence of the menstrual cycle in both humans and rodents (Cahill, 2006; Genzel et al., 2012; Muller et al., 2018; Saucier, Shultz, Keller, Cook, & Binsted, 2008; Wang et al., 2018). Thus a separate, more extensive study accounting for the rodent four-day menstrual cycle is warranted in the future.

## 5. Conclusion

In sum, we could show that Ehmt1$^{+/-}$ mice show increased cumulative-memory compared to controls, but show deficits in the simple control condition and increased habituation time to environments. We did so by combining conventional behavioral analysis with a session-based learning computational model and a trial-based classifier in the Object Space Task.

## Acknowledgements

We would like to thank the bachelor and master students, who ran the behavioural experiments: Joanne Igoli, Gülberk Bayraktar, Nikkie Cornelisse, Luc Nijssen, Elif Gizem Kain, Betul Tamer, Stefanos Louizo, Joris van Hout, Joost Pleune, Anne Mulder, Alysha Maurmair, Niamh Hemingway. Ronny Eichler for technical support and Veronica Pellis for labelling of videos for DeepLabCut. This work was supported by the Branco Weiss Society in Science grant to Lisa Genzel.

## Author contributions

E.S. helped design and perform the experiments, supervised the students and did the conventional behavioral analysis; AA, AS, INL helped execute the experiments and supervise the students; S.S. made the classifier, M.K. made the learning model, M.N. performed the genotyping, N.N.K. provided the mouse mode,; L.G. designed the experiment, supervised the students, performed the conventional analysis, supervised the model and classifier, and wrote the first draft of the manuscript. All authors helped revise the manuscript.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nlm.2020.107265.

## References


Association, A. P. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC.

Balemans, M. C., Huibers, M. M., Eikelenboom, N. W., Kuipers, A. J., van Summeren, R. C., Pijpers, M. M., ... Van der Zee, C. E. (2010). Reduced exploration, increased anxiety, and altered social behavior: autistic-like features of euchromatin histone methyltransferase 1 heterozygous knockout mice. *Behavioural Brain Research, 208*(1), 47–55. https://doi.org/10.1016/j.bbr.2009.11.008.

Balemans, M. C., Kasri, N. N., Kopanitsa, M. V., Afinowi, N. O., Ramakers, G., Peters, T. A., ... Van der Zee, C. E. (2013). Hippocampal dysfunction in the Euchromatin histone methyltransferase 1 heterozygous knockout mouse model for Kleefstra syndrome. *Human Molecular Genetics, 22*(5), 852–866. https://doi.org/10.1093/hmg/dds490.

Balemans, M. C. M., Ansar, M., Oudakker, A. R., van Caam, A. P. M., Bakker, B., Vitters, E. L., ... van Bokhoven, H. (2014). Reduced Euchromatin histone methyltransferase 1 causes developmental delay, hypotonia, and cranial abnormalities associated with increased bone gene expression in Kleefstra syndrome mice. *Developmental Biology, 386*(2), 395–407. https://doi.org/10.1016/j.ydbio.2013.12.016.

Benevento, M., Iacono, G., Selten, M., Ba, W., Oudakker, A., Frega, M., ... Nadif Kasri, N. (2016). Histone methylation by the Kleefstra syndrome protein EHMT1 mediates homeostatic synaptic scaling. *Neuron, 91*(2), 341–355. https://doi.org/10.1016/j.neuron.2016.06.003.

Benevento, M., Oomen, C. A., Horner, A. E., Amiri, H., Jacobs, T., Pauwels, C., ... Kasri, N. N. (2017). Haploinsufficiency of EHMT1 improves pattern separation and increases hippocampal cell proliferation. *Scientific Reports, 7*, 40284. https://doi.org/10.1038/srep40284.

Beversdorf, D. Q., Smith, B. W., Crucian, G. P., Anderson, J. M., Keillor, J. M., Barrett, A. M., ... Heilman, K. M. (2000). Increased discrimination of “false memories” in autism spectrum disorder. *Proceedings of the National Academy of Sciences, 97*(15), 8734–8737.

Boucher, J., & Anns, S. (2018). Memory, learning and language in autism spectrum disorder. *Autism & Developmental Language Impairments, 3*. https://doi.org/10.1177/2396941517742078.

Boucher, J., Mayes, A., & Bigham, S. (2012). Memory in autistic spectrum disorder. *Psychological Bulletin, 138*(3), 458–496. https://doi.org/10.1037/a0026869.

Cahill, L. (2006). Why sex matters for neuroscience. *Nature Reviews Neuroscience, 7*(6), 477–484.

Carreira, J., & Zisserman, A. (2017, 21–26 July 2017). Quo vadis, action recognition? A new model and the kinetics dataset. In Paper presented at the 2017 IEEE conference on computer vision and pattern recognition (CVPR).

Frega, M., Linda, K., Keller, J. M., Gümüş-Akay, G., Mossink, B., van Rhijn, J.-R., ... Kasri, N. N. (2019). Neuronal network dysfunction in a human model for Kleefstra syndrome mediated by enhanced NMDAR signaling. *bioRxiv,* 585596. https://doi.org/10.1101/585596.

Gaigg, S. B., Bowler, D. M., & Gardiner, J. M. (2014). Episodic but not semantic order memory difficulties in autism spectrum disorder: Evidence from the Historical Figures Task. *Memory, 22*(6), 669–678. https://doi.org/10.1080/09658211.2013.811256.

Genzel, L., Kiefer, T., Renner, L., Wehrle, R., Kluge, M., Grözinger, M., ... Dresler, M. (2012). Sex and modulatory menstrual cycle effects on sleep related memory consolidation. *Psychoneuroendocrinology, 37*(7), 987–989. https://doi.org/10.1016/j.psyneuen.2011.11.006.

Genzel, L., Schut, E., Schroder, T., Eichler, R., Khamassi, M., Gomez, A., ... Battaglia, F. (2019). The object space task shows cumulative memory expression in both mice and rats. *PLoS Biology, 17*(6).

Genzel, L., & Wixted, J. T. (2017). Cellular and systems consolidation of declarative memory. In N. Axmacher, & B. Rasch (Eds.). *Cognitive neuroscience of memory consolidation* (pp. 3–16). Cham: Springer International Publishing.

Gupta-Agarwal, S., Franklin, A. V., Deramus, T., Wheelock, M., Davis, R. L., McMahon, L. L., & Lubin, F. D. (2012). G9a/GLP histone lysine dimethyltransferase complex activity in the hippocampus and the entorhinal cortex is required for gene activation and silencing during memory consolidation. *Journal of Neuroscience, 32*(16), 5440–5453.

Iacono, G., Dubos, A., Meziane, H., Benevento, M., Habibi, E., Mandoli, A., ... Stunnenberg, H. G. (2018). Increased H3K9 methylation and impaired expression of Protocadherins are associated with the cognitive dysfunctions of the Kleefstra syndrome. *Nucleic Acids Research, 46*(10), 4950–4965. https://doi.org/10.1093/nar/gky196.

Kleefstra, T. (2005). Disruption of the gene Euchromatin Histone Methyl Transferase1 (Eu-HMTase1) is associated with the 9q34 subtelomeric deletion syndrome. *Journal of Medical Genetics, 42*(4), 299–306. https://doi.org/10.1136/jmg.2004.028464.

Kleefstra, T., Brunner, H. G., Amiel, J., Oudakker, A. R., Nillesen, W. M., Magee, A., ... van Bokhoven, H. (2006). Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome. *American Journal of Human Genetics, 79*(2), 370–377. https://doi.org/10.1086/505693.

Kleefstra, T., Kramer, J. M., Neveling, K., Willemsen, M. H., Koemans, T. S., Vissers, L. E. L. M., ... van Bokhoven, H. (2012). Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *American Journal of Human Genetics, 91*(1), 73–82. https://doi.org/10.1016/j.ajhg.2012.05.003.

Kleefstra, T., Schenck, A., Kramer, J. M., & van Bokhoven, H. (2014). The genetics of cognitive epigenetics. *Neuropharmacology, 80*, 83–94.

Kramer, J. M., Kochinke, K., Oortveld, M. A., Marks, H., Kramer, D., de Jong, E. K., ... Schenck, A. (2011). Epigenetic regulation of learning and memory by Drosophila EHMT/G9a. *PLoS Biology, 9*(1), e1000569. https://doi.org/10.1371/journal.pbio.1000569.

Kurz, E. M., Conzelmann, A., Barth, G. M., Hepp, L., Schenk, D., Renner, T. J., ... Zinke, K. (2019). Signs of enhanced formation of gist memory in children with autism spectrum disorder – A study of memory functions of sleep. *Journal of Child Psychology and Psychiatry, 60*(8), 907–916.

Martens, M. B., Frega, M., Classen, J., Epping, L., Bijvank, E., Benevento, M., ... Nadif Kasri, N. (2016). Euchromatin histone methyltransferase 1 regulates cortical neuronal network development. *Scientific Reports, 6*, 35756. https://doi.org/10.1038/srep35756.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience, 21*(9), 1281–1289.

Moscovitch, M., Cabeza, R., Winocur, G., & Nadel, L. (2016). Episodic memory and beyond: The hippocampus and neocortex in transformation. *Annual Review of Psychology, 67*, 105–134.

Muller, N., Campbell, S., Nonaka, M., Rost, T. M., Pipa, G., Konrad, B. N., ... Genzel, L. (2018). 2D:4D and spatial abilities: From rats to humans. *Neurobiology of Learning and Memory, 151*, 85–87. https://doi.org/10.1016/j.nlm.2018.04.012.

Parra, M. A., Cubelli, R., Bellamy, K. J., Abrahams, S., Avila, C. L., Castro-Jaramillo, L. D., & Della Sala, S. (2016). Gist-based illusions within and across stimulus modalities in autism spectrum disorder. *Memory, 24*(3), 295–305. https://doi.org/10.1080/09658211.2015.1004349.

Plaisted, K., O'Riordan, M., & Baron-Cohen, S. (1998). Enhanced visual search for a conjunctive target in autism: A research note. *Journal of Child Psychology and Psychiatry, 39*(5), 777–783.

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology, 23*(17), R764–R773. https://doi.org/10.1016/j.cub.2013.05.041.

Saucier, D. M., Shultz, S. R., Keller, A. J., Cook, C. M., & Binsted, G. (2008). Sex differences in object location memory and spatial navigation in Long-Evans rats. *Animal Cognition, 11*(1), 129–137. https://doi.org/10.1007/s10071-007-0096-1.

Shinkai, Y., & Tachibana, M. (2011). H3K9 methyltransferase G9a and the related molecule GLP. *Genes & Development, 25*(8), 781–788.

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor Perspectives in Biology, 7*(8), a021766. https://doi.org/10.1101/cshperspect.a021766.

Stewart, D. R., & Kleefstra, T. (2007). The chromosome 9q subtelomere deletion syndrome. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics, 145C*(4), 383–392. https://doi.org/10.1002/ajmg.c.30148.

Vermeulen, K., Egger, J. I. M., Janzing, J. G. E., van Dongen, L., van Bokhoven, H., Kleefstra, T., & Staal, W. G. (2017). The context of symptom measures: Interpretation and clinical diagnosis of autism spectrum disorders in intellectual disabilities. *Journal of the American Academy of Child & Adolescent Psychiatry, 56*(7), 618–619. https://doi.org/10.1016/j.jaac.2017.05.009.

Wang, S. H., & Morris, R. G. (2010). Hippocampal-neocortical interactions in memory formation, consolidation, and reconsolidation. *Annual Review of Psychology, 61*(49–79), C41–C44. https://doi.org/10.1146/annurev.psych.093008.100523.

Wang, W., Le, A. A., Hou, B., Lauterborn, J. C., Cox, C. D., Levin, E. R., ... Gall, C. M. (2018). Memory-related synaptic plasticity is sexually dimorphic in rodent hippocampus. *Journal of Neuroscience, 38*(37), 7935–7951. https://doi.org/10.1523/jneurosci.0801-18.2018.

Willemsen, M. H., Willemsen, M. H., Vulto-van Silfhout, A. T., Nillesen, W. M., Wissink-Lindhout, W. M., van Bokhoven, H., ... Kleefstra, T. (2011). Update on Kleefstra syndrome. *Molecular Syndromology, 2*(3–5), 202–212. https://doi.org/10.1159/000335648.

Wiltschko, A. B., Johnson, M. J., Iurilli, G., Peterson, R. E., Katon, J. M., Pashkovski, S. L., ... Datta, S. R. (2015). Mapping sub-second structure in mouse behavior. *Neuron, 88*(6), 1121–1135. https://doi.org/10.1016/j.neuron.2015.11.031.