A 1-night operant learning task without food-restriction differentiates among mouse strains in an automated home-cage environment

Authors
Esther Remmelink\textsuperscript{a,b,c}; Maarten Loos\textsuperscript{d}; Bastijn Koopmans\textsuperscript{a}; Emmeke Aarts\textsuperscript{c}; Sophie van der Sluis\textsuperscript{d}; Neuro-BSIK Mouse Phenomics Consortium\textsuperscript{e}; August B. Smit\textsuperscript{b}\textsuperscript{*}; Matthijs Verhage\textsuperscript{c}\textsuperscript{*}

\textsuperscript{a}Sylics (Synaptologics B.V.), Amsterdam, The Netherlands
\textsuperscript{b}Department of Molecular and Cellular Neurobiology, Center for Neurogenomics and Cognitive Research, Neuroscience Campus Amsterdam, VU University, Amsterdam, The Netherlands
\textsuperscript{c}Department of Functional Genomics, Center for Neurogenomics and Cognitive Research, Neuroscience Campus Amsterdam, VU University, Amsterdam, The Netherlands
\textsuperscript{d}Section Complex Trait Genetics, Department of Clinical Genetics, VU Medical Center, Amsterdam, The Netherlands
\textsuperscript{e}Collaborators: Brussaard AB, Borst JG, Elgersma Y, Galjart N, van der Horst GT, Levelt CN, Pennartz CM, Smit AB, Spruijt BM, Verhage M, de Zeeuw CI
\textsuperscript{*}These authors contributed equally to this work

Keywords
Operant conditioning; Instrumental learning; PhenoTyper; Automated home-cage; Mice

Published in:
Abstract

Individuals are able to change their behavior based on its consequences, a process involving instrumental learning. Studying instrumental learning in mice can provide new insights in this elementary aspect of cognition. Conventional appetitive operant learning tasks that facilitate the study of this form of learning in mice, as well as more complex operant paradigms, require labor-intensive handling and food deprivation to motivate the animals. Here, we describe a 1-night operant learning protocol that exploits the advantages of automated home-cage testing and circumvents the interfering effects of food restriction. The task builds on behavior that is part of the spontaneous exploratory repertoire during the days before the task. We compared the behavior of C57BL/6J, BALB/cJ and DBA/2J mice and found various differences in behavior during this task, but no differences in learning curves. BALB/cJ mice showed the largest instrumental learning response, providing a superior dynamic range and statistical power to study instrumental learning by using this protocol. Insights gained with this home-cage-based learning protocol without food restriction will be valuable for the development of other, more complex, cognitive tasks in automated home-cages.
Introduction

Humans, and other animals, have the cognitive skills to adjust their behavioral repertoire in the face of novel situations. The ability to change behavior based on its consequences, also known as instrumental learning, or operant conditioning, can be considered fundamental to many forms of cognitive functioning. Studying this form of learning in mice, in for instance panels of inbred lines or mutant mouse lines, is important for our understanding of the genetic mechanisms underlying this elementary aspect of cognitive functioning.

Appetitive operant conditioning is a form of instrumental learning in which the reinforcing stimulus is palatable. This is traditionally studied in mice by using an operant conditioning chamber in which the animals have to learn to respond with a lever press or nose poke to a stimulus in order to receive a food or liquid reward, delivered at a specific location. Tasks for more complex forms of cognition that are performed in operant chambers, like a reversal learning task that measures flexibility, or the 5 choice serial reaction time task that measures attention and impulsivity, use the same principle albeit with more complex schedules of reinforcement. Although operant testing provides in-depth insight into cognition, it unfortunately requires labor-intensive animal-handling, which may confound task outcome and can cause handling stress. Another drawback is that food-restriction protocols, regularly used in operant testing to motivate mice to perform, can affect behavioral responses differentially in different mouse strains. In particular, food restriction can be perceived as a stressor, as it has been shown to increase stress hormone levels, and can in this way influence task outcome. Additionally, food restriction can change circadian and task related activity patterns in rodents.

To increase throughput and reproducibility of behavioral screening, new fully automated testing strategies are desirable, e.g., testing mice in their home-cage with subsequent automated data analysis. Automatic tracking of spontaneous behavior of mice in their home-cage for extended periods without human interference can provide comprehensive and detailed analysis of naturalistic behavior. Additionally, testing mice in an automated home-cage produced consistent strain differences across laboratories.

In this study, the main experimental question was to design an operant conditioning procedure, without food-restriction. In our effort towards enhancing the efficiency of behavioral testing, we have included this task in a 1-week automated home-cage protocol that combines observations of spontaneous behavior during a habituation period, with 2 other tasks; an avoidance learning task and an anxiety task. Together, this protocol allows for the study of multiple behavioral domains; i.e. locomotor activity, learning and anxiety, in an environment that does not require experimenter intervention.
In this report, we describe the design and analysis of the 1-night operant conditioning procedure in this 7-day protocol that exploits the advantages of automated home-cage testing with diminished interfering effects of handling or food-restriction stress on task outcome and activity. These advantages also allow for increased reproducibility and scalability.

The task protocol builds on experimental procedures previously developed by de Heer et al.\textsuperscript{82} and uses food reward without prior food-restriction. After 3 days of habituation and analysis of spontaneous behavior\textsuperscript{66}, on the fourth day mice can receive a food reward by performing an instrumental response, i.e., climbing on their shelter. The task was distributed over multiple sessions to prevent satiety. The task started after a habituation period of 3 days because our previous home-cage experiments\textsuperscript{66}, as well as those of others\textsuperscript{41,61}, showed that it can take up to three days, depending on the mouse strain used, for activity parameters to stabilize.

We compared the behavior of C57BL/6J, BALB/cJ, and DBA/2J mice on this task and were able to detect instrumental learning within C57BL/6J and BALB/cJ mice by analyzing their locomotor patterns. Differences in the magnitude of the instrumental learning response were found between BALB/cJ mice and C57BL/6J mice.

**Materials and methods**

**Mice**

61 C57BL/6J, 27 BALB/cJ, and 32 DBA/2J male mice were obtained from Charles River Laboratories (L’Arbresle, France; European supplier of Jackson Laboratories) and maintained in the facilities of the NeuroBSIK consortium (VU University Amsterdam, The Netherlands or at Harlan Laboratories, Horst, The Netherlands). At an age of 8 - 12 weeks, mice were single housed on sawdust in standard Makrolon type II cages enriched with cardboard nesting material for at least one week prior to experiments, with water and food (2018 Teklad, Harlan Laboratories, Horst, The Netherlands) ad libitum (7:00/19:00 lights on/off; providing an abrupt phase transition). Experiments were carried out in accordance with the European Communities Council Directive of 24 November 1986 (86/609/EEC), and with approval of the Animal Experiments Committee of the VU University.

**Automated home-cage and testing protocol**

Testing was performed in an automated home-cage environment (PhenoTyper\textsuperscript{\textregistered} model 3000, Noldus Information Technology, Wageningen, The Netherlands) in which behavior was tracked by video and where hardware actions were triggered by the location of the mouse, as described in detail in Maroteaux et al.\textsuperscript{81}. Cages (L = 30 × W = 30 × H = 35 cm) were made of transparent Perspex walls with an opaque
A 1-night operant learning task without food-restriction

Perspex floor covered with bedding based on cellulose, and were equipped with a water bottle and a feeding station. A triangular shaped shelter with two entrances was fixed in one corner (H = 10 x D = 9 cm, non-transparent material). In the opposite corner an aluminum tube of a reward dispenser protruded into the cage. Food and water were provided ad libitum.

The top unit of each cage contained an array of infrared LEDs and an infrared-sensitive video camera used for video-tracking. The X-Y coordinates of the center of gravity (COG) of mice, sampled at a resolution of 15 coordinates per second, were acquired using EthoVision software (EthoVision HTP 2.1.2.0, based on EthoVision XT 4.1, Noldus Information Technology, Wageningen, The Netherlands) and processed to generate behavioral parameters using AHCODA analysis software (version 1.3.2, Synaptologics BV, Amsterdam, The Netherlands) and R, version 2.15.0. Two zones were digitally defined: an OnShelter zone on top of the shelter and a RewardZone in the corner where the pellet dispenser was positioned.

Mice were introduced to the PhenoTyper during the light phase (14:00h – 16:00 h) and housed in this cage without any further human interference. Video tracking and the testing protocol started at the onset of the first dark phase (19:00 h). The protocol (Fig. 6) started off with 3 days of habituation where spontaneous behavior was tracked. At the start of the 4th dark phase, the operant conditioning task became active. In this task, mice had to learn that a visit to the top of the shelter (i.e. an OnShelter visit) led to a reward (14 mg Dustless Precision Pellets F05684, 14 mg, Bio-Serve, Frenchtown, NJ, USA) being dropped by the reward dispenser in the corner opposite to the shelter (i.e. RewardZone). We refer to this task as an operant conditioning task and not a Pavlovian cue approach learning or autoshaping task, because mice had to make a specific operant response (OnShelter visit) to receive a reward. Furthermore, neither the cue light nor the shelter was at the location of the reward and therefore approach behavior, such as first described by Brown and Jenkins, could not be studied.

The operant protocol was not continuously active during the night, but split up in 10 sessions of 15 min, with session intervals of 1 h to prevent satiety for the rewards. Each session started off with a ‘free’ reward delivery. Mice could earn additional rewards by making OnShelter visits. These visits were rewarded according to a fixed ratio (FR1) schedule, but only when an OnShelter visit was followed by a visit of the reward zone (i.e. a RewardZone visit). There were no temporal restrictions for this RewardZone visit to occur. The start and duration of each session was indicated by a yellow cue light in the top unit that was on for the full duration of each session, which was intended to signal the possibility to earn food rewards. Mice were not food-restricted before the start of the task and regular chow was freely accessible during the task.
Statistical analysis

Behavior of mice during the task sessions on day 4 (Q-task) was compared to behavior during the same time bins on day 3 (Q-D3), and to 15-minute bins during the inter-session period (Q1-after, Q2-after, Q3-after, Q4-after) (Fig. 7). Differences in distance moved during all these bins were analyzed using a Friedman ANOVA, followed by a Wilcoxon Signed Rank test post-hoc analysis with Bonferroni correction that compared distance moved during all bins to distance moved during the task bin. OnShelter count data was analyzed using Poisson Generalized Linear Models to compare between strains, or Poisson Generalized Estimating Equations (GEE) to compare between bins within a strain.

Mice were required to visit OnShelter to earn additional rewards after the first free reward delivery. Learning of an association between OnShelter and reward delivery was therefore dependent on OnShelter exploration. To prevent hampering the detection of learning within groups of mice, mice were excluded from further analyses when they showed low levels of OnShelter exploration; i.e., fewer than 10 OnShelter visits during the entire task and absence of OnShelter visits in more than 5 sessions of the task.

All relative frequency measures of goal-directed behavior and learning were analyzed over bins or sessions using Generalized Estimating Equations (GEE) with a Binomial distribution. Relative time spent scores were log transformed and analyzed using a factorial repeated measures ANOVA or paired t-test. Comparisons of bin Q-D3 or Q1-after to Q-task, were performed with Bonferroni correction.

Post hoc power analyses were performed to determine the minimum number of mice needed to detect learning with a power of 80%. The analysis used resampling without replacement from a group of mice, with a sampling rate of 1000 and sample sizes ranging from 5 to 50. On each sample drawn, a binomial GEE analysis was performed to determine if a significant learning effect could be detected. The smallest sample size that gave a significant result 800 out of the 1000 samples drawn, was established to be the minimum number of mice needed to detect learning within that group.

Analysis were performed using R, version 2.15.0 or IBM SPSS Statistics 20. Error bars in graphs show the standard error of the mean.
A 1-night operant learning task without food-restriction

Figure 6 | Schematic overview of the Phenotyper and the task protocol. During the first 3 days, the mouse habituated to the cage. On day 4, the mouse was subjected to 10 operant conditioning sessions where it had to visit OnShelter to earn a food pellet. (1) RewardZone, (2) Drinking spout, (3) OnShelter zone, (4) Feeding station, (5) Pellet dispenser, (6) Yellow LED (cue light) in the top unit of the cage.

Figure 7 | Schematic representation of the operant protocol. The operant conditioning task consisted of 10 sessions of 15 min (quarters; Q) spread out over the dark phase of the 4th day. Each session was followed by an inter-session interval of 1 h, which was divided into 4 15-minute bins (Q1-after, Q2-after, Q3-after, Q4-after). Session 1 coincided with the start of the dark phase. Behavior of the mice during the 10 task bins on day 4 (Q-task) was taken together, averaged, and compared to behavior during the same time bins as the task on day 3 (Q-D3), and to the 15-minute bins during the inter-session interval.
Chapter 2

Results

In the 1-night home-cage operant learning task without food-restriction, mice had to make an instrumental response (i.e. move OnShelter) to receive a reward in a RewardZone. Their behavior was tracked during the 10 task sessions in the 4th dark phase in the automated home-cage and compared to behavior during the same time bins on day 3, and to the 15-min bins during the 1 h inter-session intervals. Below, the task and the analyses are exemplified for C57BL/6J mice, and their task performance is subsequently compared to that of BALB/cJ and DBA/2J mice.

Increases in OnShelter visits and distance moved due to the task

To assess the behavioral response to the task we recorded OnShelter visits and distance moved per 15 min of C57BL/6J mice for day 4 and, as a control, day 3 (Fig. 8). OnShelter visits were not increased during session 1 of the task compared to day 3 (Fig. 8a), as expected because the animals were naïve to the test and reward, but were clearly increased in all subsequent sessions. Averaging the data over all sessions (Fig. 8b) showed that OnShelter visits were almost doubled during Q-task compared to the same time bins on Day 3 (Q-D3) as well as compared to Q2-after, Q3-after and Q4-after (Fig. 8b; Main effect bin: $\chi^2(5) = 330.76, p < 0.001$; Post hoc for each bin compared to task, respectively $p = 0.000, 0.811, 0.000, 0.000, 0.000$). General activity, in terms of the total distance moved, followed a similar pattern to OnShelter visits; the distance travelled was increased during Q-task (Fig. 8c and d) but returned to Q-D3 levels during the inter-session interval (Main effect bin: $\chi^2(5) = 161.94, p < 0.001$; Post hoc for each bin compared to task, respectively $p = 0.000, 0.002, 0.000, 0.000, 0.000$). Both OnShelter visits and distance moved during Q-task were not significantly higher than during Q1-after, indicating that the behavioral response to the task lasted not more than 15 min beyond the session duration.

To investigate the effect of general activity on OnShelter visits, we analyzed whether distance moved significantly predicted the number of OnShelter visits during Q-task. Regression analysis indicated this was the case (F(1,60) = 327.04, p < 0.001, $R^2 = 0.85$). Taken together, the operant learning task evoked a clear behavioral response, however, it could not be excluded that the increase in OnShelter visits was the consequence of a general increase in activity, rather than a specific association of OnShelter visits with the reward.
A 1-night operant learning task without food-restriction

Mice respond to reward delivery with more direct movements
To discriminate between a general increase in activity and a specific association of OnShelter visits with the reward, we investigated the movement between OnShelter and RewardZone (OR movements) during Q-D3, Q-task and Q1-after. Four C57BL/6J mice, that showed extremely low levels of OnShelter visits during the task, were excluded from further analyses based on the exclusion criteria as described in the materials and methods section. If mice were motivated to pick up
the reward that followed an OnShelter visit during the task, they are expected to move more directly from OnShelter to the RewardZone during the task than during Q-D3 or Q1-after. Fig. 9a shows the tracks of OR movements during Q-D3, Q-task and Q1-after of a representative mouse. In this typical example, the mouse clearly made a higher number of direct movements during Q-task, as can be seen from tracks running diagonally through the cage from OnShelter to the RewardZone. Nevertheless, indirect OR movements also occurred during the task. In order to quantify whether OR movements were more direct during Q-task, movements were classified as direct OR movements (<60cm) or indirect OR movements (>60cm). This cutoff point was chosen because a movement track running along the walls covers about 60cm. The relative frequency of direct OR movements during Q-task was significantly higher compared to Q-D3 and Q1-after (Fig. 9b; GEE: Task versus Q-D3: \(\chi^2(1) = 124.24, p < 0.001\); Task versus Q1-after: \(\chi^2(1) = 124.53, p < 0.001\)). This implies goal-directed behavior of mice as a result of their motivation to pick up the reward after an OnShelter visit.

**Mice make instrumental responses**

In order to determine whether mice made instrumental responses during Q-task to induce a new reward delivery after obtaining the previous one, we examined whether mice made more direct movements from the RewardZone back to OnShelter (i.e., RO movements). During the task, mice made more direct moves between RewardZone and OnShelter compared to Q-D3 and Q1-after (Fig. 9d). Direct RO movements were significantly higher during Q-task than during Q-D3 and Q1-after (Fig. 9c; GEE: Task versus Q-D3: \(\chi^2(1) = 12.21, p < 0.001\); Task versus Q1-after: \(\chi^2(1) = 17.86, p < 0.001\)). This indicated that mice made instrumental responses to evoke reward delivery.

When mice learn, they are expected to increase the number of direct RO movements relative to the total number of RO movements over sessions. Indeed, mice increased their direct movements over sessions (Fig. 10; GEE: Session^2: \(\chi^2(1) = 19.69, p < 0.001\)). This increase appeared to be quadratic rather than linear. A reduced motivation towards the end of the dark phase (session 9 and 10) might have played a role in the detected quadratic learning curve. In addition, not all mice made RO moves during all sessions, probably due to variation in activity patterns among individual mice. The number of mice making RO movements followed the same activity pattern as observed in Fig. 8, and as previously reported, with most mice performing these movements during session 2 and 9, and least mice during session 7. This explains the fluctuation in error bars over sessions.
Figure 9 | OnShelter to RewardZone (OR) and RewardZone to OnShelter (RO) movements.  
(a) Visualization of all OR movements of 1 representative mouse during Q-D3, Q-task and Q1-after.  
(b) Relative frequency of direct OR movements during Q-D3, Q-task and Q1-after (response to reward).  
(c) Relative frequency of direct RO movements during Q-D3, Q-task and Q1-after (instrumental learning response).  
(d) Visualization of all RO movements of 1 mouse during Q-D3, Q-task and Q1-after.  
***Bonferroni corrected p < 0.001 compared to Q-task. n = 57. Error bars represent SEM.
Chapter 2

Strain specific performance
To investigate the suitability of the task for testing different strains of mice, we also tested BALB/cJ and DBA/2J mice on this protocol. BALB/cJ mice made a similar number of unconditioned (Q-D3) OnShelter visits as C57BL/6J mice (Fig. 11; $\chi^2(1) < 1$, ns), while DBA/2J mice made fewer ($\chi^2(1) = 322.81, p < 0.001$). During the task, BALB/cJ mice made more, and DBA/2J mice made fewer OnShelter visits than C57BL/6J mice (respectively: $\chi^2(1) = 242.55, p < 0.001$; $\chi^2(1) = 758.03, p < 0.001$).

Of the 32 DBA/2J mice tested, only 14 passed the inclusion criteria for further analysis as described in the experimental procedures, due to the low levels of OnShelter visits during the task as can be perceived in Fig. 11. OnShelter visits were not significantly increased during Q-task compared to Q-D3 in DBA/2J mice ($\chi^2(1) = 3.30, p = 0.069$). This is not surprising, as a prerequisite for associating OnShelter with reward delivery is the initial exploration of the top of the shelter. We concluded that this task is unsuitable for DBA/2J mice as these mice did not show sufficient levels of this type of exploration.

To investigate whether the presence of reward during the task had an effect on time spent feeding from the habituated ad libitum source of food, we used time spent in the zone around the Feeder (FeederZone) as a proxy for feeding behavior. Both C57BL/6J and BALB/cJ mice spent significantly less time in the FeederZone, as a percentage of time outside the shelter, during the task compared to Q-D3 (Fig. 11; Main effect of bin $F(1,85) = 98.86, p < 0.001$; No effect of genotype between C57BL/6J and BALB/cJ $F(1,85) = 1.68, p = 0.198$), underscoring the notion that mice of both these strains are motivated to eat the reward pellets in the presence of regular chow.

This reduction in time spent in the FeederZone during the task was absent in DBA/2J mice ($t(13) = -1.49, p = 0.159$). This was most likely due to the low levels of OnShelter visit made by these mice, and therefore low number of pellets earned, demonstrating that this task is unsuitable to test operant learning in DBA/2J mice.

Only 1 BALB/cJ mouse did not pass the inclusion criteria. BALB/cJ mice, like
C57BL/6J mice, made instrumental responses during the task as can be concluded from a selective increase of direct RewardZone to OnShelter movements during Q-task (Fig. 12a; GEE: Task versus Q-D3: $\chi^2(1) = 40.92, p < 0.001$; Task versus Q1-after: $\chi^2(1) = 20.37, p < 0.001$). BALB/cJ mice did not differ from C57BL/6J mice in the relative frequency of short moves during Q-task (GEE: $\chi^2(1) = 1.46, p = 0.23$).

However, the difference in relative frequency of direct RO moves between Q-D3 and Q-task, an index of instrumental learning, was significantly larger in BALB/cJ mice than in C57BL/6J mice, with an increase from 0.32 to 0.37 in C57BL/6J mice (Fig. 9c) and from 0.10 to 0.32 in BALB/cJ mice (Fig. 12a; GEE for strain*bin: $\chi^2(1) = 6.65, p = 0.01$). This stronger instrumental learning response in BALB/cJ mice compared to C57BL/6J is mostly due to the low relative frequency of direct RO moves during Q-D3. Thus, in this task, BALB/cJ show a larger dynamic range to study the instrumental learning response.

**Figure 11 | Number of OnShelter visits and relative time spent in the FeederZone.** The bar graph represents the number of OnShelter visits per strain during Q-D3 and Q-task, averaged over session. BALB/cJ (n = 27), C57BL/6J (n = 61), DBA/2J (n = 32). The line graph shows relative time spent in the FeederZone as a percentage of total time outside the shelter during Q-D3, Q-task and Q-after, averaged over session. BALB/cJ (n = 26), C57BL/6J (n = 57), DBA/2J (n = 14). Error bars represent SEM.

**Figure 12 | Direct RO moves of BALB/cJ mice.** (a) Relative frequency of direct (< 60cm) movements during Q-D3, Q-task and Q1-after (instrumental learning response). (b) Relative frequency of direct RO moves per session (learning curve). ***Bonferroni corrected $p < 0.001$ compared to Q-task. n = 26. Error bars represent SEM.
The learning curve over sessions of BALB/cJ mice (Fig. 12b) was best modeled by a linear curve with a significant increase of direct movements over sessions (GEE: Session: $\chi^2(1) = 10.32, p = 0.001$; Session$^2$: $\chi^2(1) = 1.95, p = 0.162$). A striking reduction in relative frequency of direct RO moves was perceived during session 8 and 9. As with C57BL/6J mice, the number of mice showing RO movements in each session, followed the general activity pattern (data shown in [66]), with fewer mice showing task activity with increasing session number. This explains the observed increase in error bars in the later sessions, however, not the high performance in session 10.

Due to the different shape of the learning curve of C57BL/6J and BALB/cJ mice in our study, it is difficult to directly compare the rate of learning between these two strains for the full 10 sessions. However, there was no difference in rate of learning during the initial instrumental learning sessions when comparing the linear part of the curve up to session 5 in C57BL/6J with the performance up to session 5 in BALB/cJ mice (GEE: effect of session: $\chi^2(1) = 3.87, p = 0.049$; effect of genotype: $\chi^2(1) = 2.41, p = 0.120$).

To determine the suitability of using C57BL/6J mice and BALB/cJ mice to study instrumental learning using the present task, power analyses were performed. These power analyses on the instrumental learning parameter (change in relative frequency of RO moves between Q-D3 and Q-task) used resampling without replacement from the two strains of mice in which the learning effect was present, as described in section 2.3. Performing this analysis by resampling from the 57 C57BL/6J mice in this study showed that approximately 30 mice were required to detect a significant learning effect with a power of 80% (Fig. 13). As a consequence of the larger difference in direct moves between Q-D3 and Q-task in BALB/cJ mice (Fig. 12a compared to Fig. 9c), the power analysis indicated that a sample size of only 10 BALB/cJ mice is already sufficient to detect a significant learning effect with a power of 95% (Supplementary Fig. S1).

Taken together, C57BL/6J and BALB/cJ mice do not show differences in learning curves, however, the BALB/cJ strain displayed a larger instrumental learning response, providing a superior dynamic range resulting in more statistical power to detect the effect of interventions on instrumental learning compared to C57BL/6J.

Figure S1 | Post hoc estimated power. Power is estimated by resampling from the C57BL/6J population ($n = 57$) and BALB/cJ population ($n = 26$) with a sampling rate of 1000 times at each particular sample size. The estimated power is the portion of the 1000 samples that gave a significant GEE result. The dashed line signifies a power of 80%.
Discussion

This report describes differences in behavior among 3 mouse strains in a novel operant learning task in an automated home-cage environment in the absence of distressing food-restriction. During the 1-night protocol, behavior that is part of the spontaneous exploratory repertoire during the previous days (i.e. making an OnShelter visit), was rewarded by the distribution of a reward into the reward zone of the home-cage. This protocol induced a clear increase in OnShelter visits in both C57BL/6J and BALB/cJ mice, but these visits remained infrequent in DBA/2J mice. Goal directed behavior towards the reward, as well as instrumental learning, were studied in C57BL/6J and BALB/cJ mice by assessing directional movements. We did not detect differences in initial learning curves between these strains, however, BALB/cJ mice showed a stronger instrumental learning response, providing a superior dynamic range and statistical power to detect the effect of interventions on instrumental learning compared to C57BL/6J.

Separating general activity from learning in the absence of food-restriction

Performance during operant learning was previously found to be affected by several factors, such as basal exploratory activity, stress and motivational factors. In this study, a home-cage based operant task without food-restriction was described. The task was designed as part of a 7-day home-cage based test battery and builds on a behavior that occurs naturally in most mice (climbing on the shelter), which was unrewarded during the home-cage habituation phase before the task and then rewarded during the task. Therefore, we here discuss the possible influence of general exploratory activity and the absence of food-restriction, on the performance of mice in our task.

The influence of general activity on task performance

During the unrewarded phase (day 1-3), large strain differences were observed in how often mice climb the shelter (Fig. 6 and 66). These differences are due to diverging ability or motivation to climb and explore the shelter. During the task, this form of exploratory behavior became rewarded and increased in C57BL/6J and BALB/cJ, but not DBA/2J mice. Increased general activity during the task, probably related to increased arousal due to the distribution of reward, was also observed. By employing a measure of instrumental learning that corrects for this increase in activity (i.e. relative frequency of direct RO movements), differences in general activity were effectively separated from differences in cognitive performance. Since OnShelter visits are a prerequisite for discovering the possibility to earn rewards, this task is not suited for mouse strains that show low initial levels of exploration of the top of the shelter, such as the DBA/2J strain.

A second factor that influences the study of cognitive performance in this task
is the general activity pattern of mice during the dark phase. Task-related activity followed previously observed general activity patterns and not all mice were active or showed task-related movements during each session, which led to large variation in performance between sessions. Probably, low motivation of ad libitum fed mice to change their activity patterns in order to receive food rewards, has played a role in this finding. Furthermore, the cue light may have been ineffective in enhancing activity during the task, due to its location (in the top unit rather than near the reward). As a consequence, group averages are required to observe learning effects. In some mouse strains these effects are robust and small and manageable groups suffice to detect significant task-related behavioral changes (i.e. <10 in BALB/cJ), whereas larger groups may be required for other strains.

**Testing in the absence of food-restriction**

Most reward based learning studies in mice apply food-restriction protocols to motivate the animals to perform, however, food-restriction can induce stress in rodents. Moreover, food-restriction can change task activity patterns dependent on the level of restriction used. Even though it might be apparent that food-restriction increases stress in rodents and can reverse or abolish strain difference in behavioral responses dependent on the level of restriction used, food-restriction can also aid in the detection of learning. Orsini et al. found that food restriction enhanced strain differences in non-reward based learning tasks. Forestell et al. provided evidence that food-restricted and non-restricted mice differ in the expression of learned behavior in an odor discrimination task for reward, but showed that this is not due to differences in learning. However, the expression of learning is certainly important as a readout of cognitive capabilities. Makowiecki et al. could only observe an effect of genotype in severely food-restricted mice (80% instead of 90% body weight level) and showed that mice on the lesser food restriction protocol took longer to learn the Y-maze visual discrimination task. Also, Kant et al. reported that without food-restriction, learning a maze for a food reward appears slow, even when the reward is highly desirable. Similarly, others have reported on a complete absence of operant behavior in ad libitum fed groups or performance that does not exceed chance levels.

The task protocol we describe here was designed to circumvent the confounding effects of food-restriction on task performance and activity patterns. Our operant performance measure relies on a comparison between locomotor activity during a habituated state (Q-D3) and during the task (Q-task). The specific changes in activity patterns we observed during the task can be attributed to the presence of a reward and are not confounded by changes in general activity due to food-restriction.

To prevent satiety and, thereby, a reduction in motivation, we chose to make the task discontinuous by creating multiple sessions with inter-session intervals of one hour. This prevents mice from collecting a vast amount of rewards in a short period of time. However, the lack of a strong motivating force might have been a factor
involved in the more subtle learning effect and variation in performance observed in C57BL/6J mice, as well as the absence of task related behavior in DBA/2J mice.

The data obtained in this home-cage test are in line with previous experiments in conventional non-home-cage based tests, in concluding that testing mice in a reward-based learning task without food-restriction affects the possibility to detect learning and may produce variable data\textsuperscript{106}. The resulting increase in variation in performance reduces the power to detect learning. Nonetheless, when BALB/cJ mice are used, learning can be clearly observed in a relatively small group of mice without the application of a stressful food-restriction protocol.

**Differential performance in BALB/cJ mice**

The initial learning curve over the first 5 sessions was comparable in BALB/cJ and C57BL/6J mice. In line with this, operant nose poke behavior in an operant chamber was reported to be similar in BALB/cJ mice to that of C57BL/6J mice\textsuperscript{106}. In touchscreen-based systems for learning, BALB/cJ mice need similar number of trials to reach a learning criterion as C57BL/6J mice in the simple operant conditioning phase of the task\textsuperscript{107,108}. In the present task, task-related behavior of BALB/cJ and C57BL/6J started to diverge in the last sessions (Fig. 5 and 7b). During these sessions, motivation might have differentially affected behavior in BALB/cJ mice compared to C57BL/6J mice. BALB/cJ mice have previously been found to cumulatively consume more sucrose than C57BL/6J mice over a 2 h period\textsuperscript{109}. In a behavioral chaining task, BALB/cJ mice responded more for sucrose pellets and task performance was sustained at a higher level over sessions than in C57BL/6J mice\textsuperscript{110}. These finding suggest that C57BL/6J mice satiate more quickly to sucrose.

The instrumental learning response, in terms of the change in relative frequency of direct RO moves during Q-D3 compared to Q-Task, was substantially larger in BALB/cJ mice compared to C57BL/6J, mostly due to the low relative frequency of direct RO moves during Q-D3. Thus, in this task, BALB/cJ have a larger dynamic range to study an instrumental learning response, with larger statistical power to detect learning as shown by the power analyses. In the avoidance learning task, that followed the operant conditioning task in our home-cage test battery protocol, C57BL/6J and DBA/2J mice showed a stronger cognitive response than BALB/cJ mice\textsuperscript{81}. The avoidance learning task uses light as a mild aversive stimulus to change the natural preference of mice for using one of the two shelter entrances. Strain differences in performance between these two learning task might have been caused by different sensitivity to, or motivational properties of, either sucrose rewards or light. Our home-cage protocol now enables us to study the effect of different types of reinforcers on the behavior of mice within 7 days and without changing the task environment.
Chapter 2

Conclusion
As demonstrated by this study, in an automated home-cage, learning can be studied without human intervention and distressing food-restriction within 1 night. However, the absence of food-restriction to enhance motivation has possibly reduced the power of our task to detect learning in all mutant mouse strains tested. Our fully automated 7-day protocol has decreased the amount of time and labor needed to study different behavioral dimensions. Moreover, it contributes to the standardization of behavioral screening and therefore positively contributes to efforts to enhance reproducibility of behavioral data. Insights gained with the elementary operant learning task within this protocol will be valuable for the development of other, more complex, cognitive tasks or task batteries in an automated home-cage.

Acknowledgements

We thank Rolinka van der Loo for operating the PhenoTyper systems and Ruud Wijnands for assistance, Noldus Information Technology for supplying software free of charge and hardware at cost price, Ben Loke, Cecilia Herrera, Raymond de Heer and Willem van der Veer for development of hardware, software and test scripts and Dr. Oliver Stiedl for valuable discussions. This work was supported by Agentschap NL (NeuroBSIK Mouse Phenomics Consortium, BSIK03053) and the Netherlands Organization for Scientific Research (VIDI-452-12-014 to S.v.d.S).