**Accelerated development of habits in a mouse model of glutamatergic dysfunction relevant to schizophrenia**

*dmb 15052017v4.1*

Abbreviated title: *Gria1-/-* mice demonstrate facilitated stimulus-response learning.

Anushka B.P. Fernando1, David J Sanderson2, Stephen Cuell1, Anthony Dickinson3 Rolf Sprengel4, 5Paul J Harrison, Mark E Walton1, David M Bannerman1.

1 Department of Experimental Psychology, University of Oxford, Oxford. England.

2 Department of Psychology, University of Durham. Durham. England

3 Department of Psychology, University of Cambridge, Cambridge. England

4Max-Planck Institute of Medical Research, Department of Molecular Neurobiology, D-69120 Heidelberg, Jahnstrasse 29, Germany.

5Department of Psychiatry, University of Oxford, and Oxford Health NHS Foundation Trust, Warneford Hospital, Oxford OX3 7JX

Corresponding authors:

Anushka Fernando, now Department of Neuroscience, Physiology and Pharmacology, University College London, Gower Street, London, WC1E 6BT, +447765251080, [abpfern@gmail.com](mailto:abpfern@gmail.com).

David Bannerman, Department of Experimental Psychology, University of Oxford, Oxford. OX1 3UD, +441865271426, [david.bannerman@psy.ox.ac.uk](mailto:david.bannerman@psy.ox.ac.uk).

*Key words:* GluA1, AMPA receptors, synaptic plasticity, goal-directed, spatial learning

**ABSTRACT (250)**

Genome-wide association studies and post-mortem investigations have implicated the GluA1 subunit of the AMPAR in schizophrenia. GluA1, coded by the *Gria1* gene, plays an important role in AMPAR trafficking and synaptic plasticity. *Gria1-/-* mice therefore represent an important experimental tool for investigating the role of GluA1 dysfunction in behaviours relevant to schizophrenia. Here we investigated the effects of GluA1 deletion on the balance between goal-directed (model based) and habitual (model-free) modes of behaviour, which is altered in schizophrenics. In Experiment 1 we demonstrate that while the behaviour of wild-type mice was goal-directed during performance of a modified spatial reference memory radial maze task (assessed by a devaluation procedure with satiety), *Gria1-/-* mice exhibited habitual behaviour. In Experiment 2 we used a two-lever, operant task and showed that both wild-type and *Gria1-/-* mice demonstrated sensory-specific devaluation, necessary for goal-directed behaviour. In Experiment 3 we found that although *Gria1-/-* mice were initially goal-directed in their behaviour on a single lever operant task, they later became habitual in responding after continued training. In contrast, WT mice remained goal-directed and sensitive to devaluation throughout. Our data therefore suggest that the increased propensity for habitual behaviour in *Gria1-/-* mice likely reflects differences in associative learning and the rate at which asymptotic levels of performance are attained. This may have important implications for understanding the tenacity of delusions in psychotic disorders like schizophrenia.

250.

**INTRODUCTION**

Delusions are false beliefs that are thought to reflect the formation of aberrant associations (REFS: Kapur, Gray). A key feature of psychosis is the tenacity of delusions, with aberrant beliefs often proving highly resistant to re-evaluation, despite often overwhelming contrary evidence. As such, these beliefs could be considered to exist in isolation from any overall model of the world and thus the tenacity of delusions could be thought to resemble habits, driven by stimulus-response mechanisms (as opposed to goal-directed behaviours mediated by action-outcome learning; REFS). Consistent with this possibility, schizophrenic patients have been shown to exhibit deficits in utilising action-outcome associations and display an increased propensity for habitual responding in laboratory situations (REFS-Voss et al., 2010; Morris et al., 2015). However, very little is known about the processes and mechanisms that might lead to this increased propensity for model-free, habit like behaviour in schizophrenic subjects.

In recent years, GWAS have identified over 100 genes linked to schizophrenia (Ripke et al., 2013; PGC, 2014). Many of these genes are associated with synapses and synaptic plasticity, including *Gria1* which codes for the GluA1 subunit of the AMPA receptor (also known as GluR-A or GluR1; see also Eastwood & Harrison REFS). GluA1 may be particularly important for the trafficking of additional AMPARs into the post-synaptic membrane to strengthen synapses and *Gria1-/-* mice exhibit deficits in synaptic plasticity (Malinow & Malenka; Kessels & Malinow; Zamanillo et al., 1999; Hoffman et al., 2002; Romberg et al., 2009; Eriksen et al., 2010). *Gria1-/-* mice therefore represent a useful experimental tool for studying the role of GluA1 and synaptic plasticity in behaviours relevant to schizophrenia.

*Gria1-/-* mice are impaired in a form of short-term memory that might underlie short-term habituation (Sanderson et al., 2007; 2011; Sanderson & Bannerman, 2011). In contrast, these mice are perfectly able to form long-term, associative memories. For example, they learn normally in the Morris watermaze (Zamanillo et al., 1999; Reisel et al., 2002), and during acquisition of a spatial reference memory version of the radial maze task in which they must learn to discriminate between always rewarded and never rewarded arms (Schmitt et al., 2003; 2005). If anything, *Gria1-/-* mice may even acquire long-term spatial memories faster than wild-type controls (Schmitt et al., 2003; Sanderson et al., 2009). This is all the more surprising given that synaptic plasticity is impaired in these mice (see Bannerman et al., 2014 for discussion).

Previous studies have also shown that *Gria1-/-* mice exhibit an increased propensity to behave in a model-free, habit like state (Johnson et al., 2005; 2007 see also Mead & Stephens, 2003). For example, in one study mice were trained on a double runway task to run for two distinct rewards that differed in terms of their sensory properties (e.g. flavour, texture etc.). Following training, one of the food rewards was devalued by pre-exposure to satiety in the home cage just prior to test trials on the runway task, performed in extinction. Wild-type mice ran more slowly down the runway associated with the devalued reward. In contrast, *Gria1-/-* mice continued to run at the same fast speed in both the devalued and non-devalued arms (Johnson et al., 2005). Therefore, their behaviour was insensitive to the devaluation procedure and thus more habitual than that of the controls. However, neither the generality nor the aetiology of this increased propensity to display model-free, habitual responding in *Gria1-/-* mice has been investigated.

In the present study we first investigated whether *Gria1-/-* mice would display increased habitual behaviour, and thus resistance to devaluation, on a modified version of the spatial reference memory radial maze task that we have used previously (see Figure 1; Schmitt et al., 2003). Rather than learning to discriminate between always rewarded and never rewarded arms, mice were trained to associate different arms of the radial maze with different food rewards (e.g. Noyes grain pellets versus sucrose solution). Devaluation was then assessed after mice had been allowed to consume one of the rewards to satiety in their home cages.

We next assessed the aetiology of habit formation in *Gria1-/-* mice. In a second experiment, we investigated whether *Gria1-/-* mice could ever access a sensory specific representation of the reward and behave in a goal-directed manner. To this end we used an operant choice procedure with two levers that were reinforced with independent interval schedules in order to maintain the variability in the action-outcome relationship, sustaining goal-directed behaviour even after extensive training and experience of the contingencies (Dickinson, 1985; Rescorla and Colwill, 1985; Kosaki and Dickinson, 2010).

Finally, in Experiment 3 we tracked the development of habitual responding in a longitudinal study with repeated devaluation tests at different stages during acquisition. It is generally considered that early on when learning a new task, behaviour is goal directed. However, once the task is learnt, behaviour transitions to habits. Therefore, it is possible that the increased propensity for habitual responding in *Gria1-/-* mice, compared to wild-types, reflects the rate of associative learning in these animals. The series of experiments described here set out to test this hypothesis.

850

**METHODS**

*Subjects*

Age-matched GluA1 knockout (*Gria1-/-*) and wild-type (WT) mice were used in all experiments (see Table 1 for number of mice in each experiment). The mice were bred from heterozygous *Gria1-/+*parents at the Department of Experimental Psychology, University of Oxford. Genetic construction, breeding and subsequent phenotyping were conducted as described previously by Zamanillo et al. (1999). Mice used in Experiment 1 had previously taken part in an appetitively motivated, auditory discrimination experiment conducted in standard Med Associates mouse operant boxes (Sanderson et al., *2017*). Separate groups of experimentally naïve mice were used for Experiments 2 and 3.

Mice were housed in same-sex groups (n=1-5 mice per cage), in white plastic cages with secured metal lids, and maintained on a 12-hr light/dark cycle, with lights on at 07.00AM. Before training commenced, mice were reduced to ~85% of their free-feeding weights (weight range: 20-33g), and then maintained on this schedule throughout behavioural testing. The experimenter was blind to the genotype of the mice and all testing occurred during the light stage of the diurnal cycle. Experiments complied with the U.K. Animals (Scientific Procedures) Act, 1986.

*Experiment 1:* *Investigating habit formation in Gria1-/-* *mice in a spatial reference memory radial maze task*

*Training*: Prior to commencing the radial maze task, mice received eight days of pre-training on a Y-maze in their colony holding room in order to familiarise them to the general experimental procedure (e.g. running on an open maze for food rewards; see Supplementary methods: Pre-training). Prior to the start of the radial-arm maze task, mice were also exposed to grain-based pellets and sucrose solution on separate occasions in their home cages.

Following eight days of familiarisation, mice commenced the radial-arm maze task. Testing was conducted in a novel room with different extra-maze spatial cues available. Mice received twenty sessions of radial maze training in total. Each session consisted of a forced visit to each of the six arms of the maze (i.e. 6 visits (or trials) in total). The mouse was placed on the maze with just one goal arm accessible, and allowed to run down the arm and consume the reward. One session was conducted daily throughout training, with five sessions every week. Of the six arms of the radial maze, three goal arms contained a single 20-mg grain-based pellet (Research Diets, New Brunswick, New Jersey). The other three goal arms were baited with 1 ml of sucrose solution (20% wt/vol). The allocation of a particular reward to specific goals arms, defined by their allocentric spatial location, was pseudo-randomised, with no more than two adjacent arms containing the same food reinforcer. Combinations of arms containing the same reinforcer were counterbalanced across both sex and genotype. Importantly, the allocation of food to arms was maintained across all training trials such that the spatial cues associated with a given arm reliably predicted the particular food type (pellet/sucrose). The order in which mice could visit an arm was also pseudo-randomised, with no more than two adjacent arms being visited in succession. The order was the same across mice within-sessions, but was changed between sessions with the constraint that an order of arms could not be repeated throughout training.

Each session consisted of a mouse being placed onto the central platform for 10s. The first guillotine door (at the entrance to a goal arm) was then opened, and the mouse was allowed to enter that arm and to consume the food. After leaving the arm, there was a period of 10s between one door closing and the next door opening. This pattern was continued until all 6 arms had been visited. After a final period of 10s, mice were then returned to their home-cage. The maze was then cleaned with paper towels and water. Mice were allowed a maximum of 300s to enter an arm. If mice failed to enter the arm within 300s, an omission of responding was recorded. The session then continued with the next arm in the sequence being opened. The maze was periodically rotated by 60° between sessions (clockwise or anticlockwise pseudorandomly) so that intra-maze cues were irrelevant for solving the discrimination across days. Following each of the final six sessions, mice were familiarised to the apparatus that would be used during feeding to satiety for the subsequent outcome-devaluation test (see Supplementary methods: Habituation to devaluation apparatus).

*Devaluation and Extinction test*: After twenty sessions of radial maze training, mice received an outcome devaluation test. Mice were pre-fed with either 8g of grain-based pellets or provided with bottles of 20% sucrose solution for 120 min immediately prior to testing on the radial maze. Devaluation took place in squads. Each mouse within a squad was individually housed within a cage identical to the holding cages. The cage contained a black plastic lid, into which pellets could be placed (grain-based-devaluation). A plastic drinks bottle was also available into which either 20% sucrose solution (sucrose-devaluation) or water (grain-based-devaluation) could be provided. Weights of both the food and the drink bottle were recorded before and after devaluation (see Supplementary methods: Food devaluation). The weights of the mice were also recorded. The food to which an animal was devalued (grain vs. sucrose) was counterbalanced across sex and genotype.

Immediately after this devaluation procedure, mice received a series of five sessions in extinction, conducted back-to-back (i.e. within the same day), on the radial maze. The method of running these test trials was the same as for the training stage, with two exceptions. First, no food or sucrose solution was present in the food-wells on the radial maze. Second, animals were only allowed 180s, rather than 300s, to enter an arm and reach the food-well. The order of testing was manipulated so that grain-based-devaluation occurred on test days 1 and 4, with sucrose-devaluation on test days 2 and 3. Half of the mice (counterbalanced across sex and genotype) were first exposed to a devalued arm, whilst the remaining half was initially exposed to a non-devalued arm. The sequence of arms was created so as to counterbalance the order of exposure to devalued and non-devalued arms across mice. The sequence of arms for the extinction trials was the same for all mice within a given session. However, the sequence on any given test session had not been used previously, either throughout the initial training or during the devaluation test.

*Experiment 2: Investigating goal-directed behaviour in Gria1-/-* *mice using a two lever operant choice procedure*

The experimental procedure is outlined in Table 2. Two cohorts of mice were run in the same manner except that the first cohort received an extra 3 days of training due to a programming error.

*Pre-exposure*: Food-deprived mice were exposed to both reinforcers on separate days in their home cages prior to any training, either by placing a dish with pellets in the home cage or switching the water bottle for sucrose solution for six hours. Mice then received magazine training during which 15 grain-based pellets and 15 drops (20ul) of 20% sucrose solution were experienced sequentially on a random interval, 60s schedule during a 30min session. The order of reinforcers was selected from a pseudo-randomised list.

Mice then underwent daily pre-training sessions with a single lever during which each press was reinforced with either a single grain pellet or the delivery of sucrose solution. Two sessions were conducted per day; one during which a particular lever was reinforced with pellets and one during which the other lever was reinforced with the sucrose solution. The order of sessions (i.e. grain vs. sucrose) was varied across days. The lever and reinforcer designations (e.g. left lever produces grain, right lever produces sucrose) were counterbalanced for genotype, sex and squad, and remained constant throughout the experiment for a given mouse. Initially, the mice were required to make 15 lever presses during a 30-min session for each reinforcer. Once this criterion was met, the schedule was increased to 30 lever presses in a 30-min session. All mice reached these criteria within five days.

*Training.* Following this pre-training phase, mice then received training with both levers present. Each lever press was reinforced with its designated reinforcer on an independent random interval (RI) 30-s schedule during a 30-min session. This meant that every second there was a 1/30 chance of reinforcement. If this probability was met, the next lever press was rewarded with a reinforcer. Following a single session of this concurrent RI 30-s training, the probability of receiving a reinforcer was reduced to .017 so that pressing the two levers was now reinforced by concurrent RI 60-s schedules. Moreover, throughout this concurrent training, a reinforcer was made available whenever 180 s had elapsed without the subject making an appropriate lever press in order to impose a limit on the maximum period of non-reinforcement. Training with this concurrent RI 60-s schedule continued for five more sessions.

*Devaluation and Extinction test.* Mice received a pre-feeding, devaluation session (e.g. with either grain or sucrose), followed by a 5-min extinction test in the operant chambers. The pre-feeding session (e.g. grain or sucrose) was conducted in empty home cages in the testing room. Each mouse was presented with either a dish filled with 5g of grain based pellets or a drinking bottle containing 20% sucrose solution. Animals were allowed to freely consume this reinforcer for 1hr 15 min. Dishes and bottles were weighed prior to, and following, the devaluation session.

The 5-min extinction test followed immediately in the operant chambers with the house-light and fan on throughout. Both levers were extended for the duration of the session. Responses on the lever now had no consequence. The next day mice were pre-fed with the alternate reinforcer (e.g. sucrose vs. grain), and then received a second extinction test in the operant chambers with both levers extended. The order of reinforcer devaluation across the two days was counterbalanced for genotype, sex and squad.

*Experiment 3: Investigating the rate of habit formation in Gria1-/-* *mice*

The experimental procedure is outlined in Table 3. Food-deprived mice were first exposed to grain based pellets by placing a dish with pellets in the home cage for six hours. Mice were also habituated to a drinking bottle containing 20% sucrose solution for 1.5hrs.

Magazine training was conducted for grain based pellets only (given the results from Experiment 2; see *Results*). In parallel, exposure to sucrose solution was conducted in an empty home cage for the same period of time. Following this pre-exposure phase, mice were placed in operant chambers and exposed to the lever manipulandum. The start of the session began with the illumination of the house-light and the extension of either the left or right lever, with the identity of this training lever counterbalanced across subjects, for sex and genotype.

*Pre-training.* Mice then underwent daily pre-training sessions with a single lever during which each press was reinforced with a pellet. Lever designation (i.e. right or left) was counterbalanced for sex and genotype, but remained constant throughout the experiment for a given mouse. Initially, the mice were required to make 15 lever presses in a 30-min session. Once this criterion was met, it was increased to 30 lever presses in a 30-min session. All mice reached these criteria within five days. Mice were also exposed to sucrose solution in empty home cages for the same period of time across the five days. The order of session/reinforcer exposure (i.e. instrumental training for grain based pellets or exposure to sucrose solution) was counterbalanced each day between mice and across the five days of training.

*Training.* The schedule was then changed to the RI 30-s schedule (plus 180-s limit) used in Experiment 2 for a single 30-min session. This was then followed by two 30-min sessions of training in which the schedule was now switched to a leaner RI 60-s schedule (with the 180-s limit) as also used in Experiment 2. As well as these instrumental training sessions with the pellet reinforcer, each mouse again also received parallel sessions on the same day in which it was placed in an empty home cage and allowed to consume freely the 20% sucrose solution from a standard drinking bottle for 30 min. The order of instrumental training and sucrose exposure sessions was randomised across days of training. These sucrose sessions served to familiarise the mice to the sucrose solution prior to the use of this solution as a control for general satiety in the reinforcer devaluation tests.

*Devaluation and Extinction tests:* Following this training, the mice received a pre-feeding devaluation session with either the pellets or sucrose, followed by a 5-min extinction session in the operant chambers during which responses on the training lever did not deliver any reinforcer. The pre-feeding session (e.g. grain vs. sucrose) was conducted in empty home cages in the testing room. Each mouse was presented with either a dish filled with 5g of grain based pellets or a drinking bottle containing 20% sucrose solution. Animals were allowed to consume freely this reinforcer for 1hr 15 min. Dishes and bottles were weighed prior to, and following, the devaluation session.

The 5-min extinction test followed immediately in the operant chambers with the house-light and fan on throughout. The trained lever was extended for the duration of the session. Responses on the lever now had no consequence. The next day mice were pre-fed with the alternate reinforcer (e.g. sucrose vs. grain), and then received a second extinction test in the operant chambers. The order of reinforcer exposure in the extinction tests was counterbalanced across genotype and sex.

After this first set of devaluation/extinction tests, the mice received two further cycles of instrumental training followed by pairs of reinforcer devaluation/extinction tests (see Table 3). The procedure was identical to that used during the first cycle except for the fact that the training consisted of just the two RI 60-s sessions prior to Extinction tests 2 and 3. Thus, in total, following RI30 training (Day 0), mice received RI60 lever training on Days 1-2, 5-6 and 9-10 with devaluation/extinction tests on Days 3-4, 7-8 and 11-12.

*Data analysis*

*Experiment 1:* The critical measure was the speed at which mice ran down the arms. During training the mean running speed was calculated across trials in which mice entered the arms. Therefore, if mice failed to enter an arm on a given trial, the data were considered missing and the mean was calculated from the remaining successful arm entries. In the test stage the running speed was calculated by averaging across the total number of completed runs across the five trials. A Type-I error rate of 0.05 was adopted for all reported statistical analyses for all three experiments in this study.

*Experiment 2*: A mixed ANOVA was conducted on the rates of responding during the two extinction tests with Devaluation as a within subjects factor (devaluation vs non-devaluation), and Genotype (KO vs WT), Sex (Male vs Female) and Cohort as between subjects factors. This was followed by the same analysis with reinforcer (Grain vs Sucrose) as an additional within subjects factor. Furthermore, separate analyses were also conducted for responding only on the grain-associated lever (grain devalued vs grain not devalued). A rejection criterion of p<0.05 was used and the Huynh-Feldt adjustment was applied if sphericity was violated. Pairwise comparisons were Sidak corrected.

*Experiment 3*: A mixed ANOVA was employed with devaluation/extinction test number as a within subjects factor (Test 1 vs Test 2 vs Test 3), and Genotype (*Gria1-/-* vs WT) and Sex (Male vs Female) as between subjects factors. The ANOVA was performed on the ratio of the rate of responding during each devaluation/extinction test. This ratio was calculated as the rate of lever pressing during the 5-min extinction test divided by the rate of responding during the first 5 min of the previous baseline, training session (i.e. Day 2, 4 or 6 of training). Again, a rejection criterion of p<0.05 was used and the Huynh-Feldt adjustment was applied if sphericity was violated. Pairwise comparisons were Sidak corrected.

**RESULTS**

**Experiment 1: *Gria1-/-* mice are insensitive to devaluation after training on a spatial reference memory radial maze task**

We first assessed whether *Gria1-/-* mice would exhibit habitual behaviour after training on a spatial reference memory radial maze task (Schmitt et al., 2003), using a modified paradigm in which different goal arms were consistently paired with different reinforcers (see Figure 1). *Gria1-/-* mice were insensitive to devaluation after training on the spatial reference memory radial maze task, suggesting that their behaviour was habitual, in contrast to the goal-directed behaviour exhibited by the wild-type controls

*Acquisition.* The speed at which mice ran down the goal arms paired with sucrose and grain-based pellets is shown in Figure 2a. Running speeds significantly increased over training blocks (F(9,180) = 20.01, p < .001). Overall, *Gria1-/-* mice ran significantly slower than WT mice (F(1,20) = 5.03, p < 0.04). Both groups ran significantly faster in the arms paired with sucrose than in the arms paired with grain-based pellets (main effect of reinforcer - F(1,20) = 52.31, p < 0.001). Furthermore, the effect of reinforcer also interacted with training block (F(9,180) = 2.55, p < 0.01), reflecting the gradual acquisition of the ability to discriminate between the goal arms associated with the two different rewards by mice of both genotypes.

*Gria1-/-* mice completed fewer trials overall during training than WT mice (WT: median = 120, interquartile range = 119.75-120; *Gria1-/-* mice: median = 114, interquartile range = 110.25-120). However, this effect did not reach statistical significance (U(12,12) = 44.5, p = 0.11). Including the number of trials completed as a covariate in the analysis of the running speeds did not change the pattern of results (F(1,19) = 15.45, p < 0.002), although number of trials completed did significantly correlate with running speeds.

*Devaluation/Extinction Test.*

Analysis was conducted by first converting the raw running speed data to a ratio of the sum of devalued and non-devalued speeds (i.e. = devalued/[devalued + non-devalued]) for each mouse. A ratio below 0.5 indicates that the speed of running in the devalued arms was slower than for the non-devalued arms (i.e. the mice exhibited successful devaluation). A ratio of 0.5 indicates equal running speeds in devalued and non-devalued goal arms.

There was a significant difference in the running speed ratio between WT and *Gria1-/-* mice (F(1,20) = 5.60, p = 0.028, see Figure 2b). There was no effect of sex, nor any interaction of factors (largest F value = 1.64, p > 0.2). The ratio was significantly below chance (0.5) for WT mice (t(11) = 2.63, p = 0.024), indicating that they ran more slowly down the devalued arms. In contrast, the running speed ratio for the *Gria1-/-* mice did not differ significantly from chance (t < 1, p > 0.3; see Figure 2b), demonstrating that they continued to run down devalued and non-devalued arms at the same speed. Importantly, further analysis showed that WT and *Gria1-/-* mice ran at a similar speed down the non-devalued arms (WT mean = 10.00 cm / s ±1.16 S.E.M.; *Gria1-/-* mice mean = 9.20 cm / s ±1.81 S.E.M.; F < 1). There was no effect of sex and no interactions of factors (F values < 1).

Overall, *Gria1-/-* mice completed less arm runs than WT mice during the devaluation/extinction tests (Maximum of 30 runs (5 sessions, with six arm runs per session); WT: median = 30, interquartile range = 29-30; *Gria1-/-* mice: median = 16.5, interquartile range = 10.75-24.5; U(12,12) = 19, p = 0.001). However, there was no effect of devalued vs. non-devalued arms on the number of trials completed for either group (WT: non-devalued median = 15, interquartile range = 14-15; devalued median = 15, interquartile range = 15-15, W(12) = 2.5, p = 0.16; *Gria1-/-* mice: non-devalued median = 7.5, interquartile range = 6-12.25; devalued median = 9, interquartile range = 5-11.75, W(12) = 20.5, p =0.81.

**Experiment 2: *Gria1-/-* mice are sensitive to devaluation of a reinforcing outcome during an operant choice test**

Radial maze performance in *Gria1-/-* mice was insensitive to devaluation, demonstrating that their behaviour was habitual and reliant on stimulus-response associations, consistent with previous studies in this mouse line (Johnson et al., 2005; 2007 and see also Mead & Stephens 2003). It was suggested previously that this could reflect an important fundamental role for GluA1 in the formation or accessing of a representation of the sensory-specific incentive motivational properties of an appetitive reward (Johnson et al., 2005; 2007). Alternatively, it could be argued that *Gria1-/-* mice acquired the radial maze task faster (e.g. Schmitt et al., 2003), and had become habitual at the time of devaluation testing, whereas the WT mice were still goal-directed in their behaviour. We therefore investigated whether *Gria1-/-* mice are ever able to access a sensory specific representation of the reward and behave in a goal-directed manner, by using a two lever, operant choice procedure (Dickinson, 1985; Rescorla and Colwill, 1985; Kosaki and Dickinson, 2010).

*Acquisition*: Analysis of rates of lever pressing during the last day of RI60 training revealed no differences in responding between the two cohorts of mice that were run separately (Lever x Cohort = F(1,23)=4.1 p>.05 N.S.) The two cohorts were therefore analysed together for the subsequent devaluation tests. Both *Gria1-/-* mice and WT controls responded equally for both pellets and sucrose (Reinforcer F<1 p>0.3 N.S., Genotype F<1 p>0.4 , Reinforcer x Genotype F(1,27)=1.7 p>0.2 N.S.). Male and female mice did not significantly differ in their rate of responding for sucrose or grain (Reinforcer x Sex F<1p>0.7, Reinforcer x Genotype x Sex F<1 p>0.9) although, male mice did respond more than female mice during acquisition (Sex F1,27)=5.5 p<.05).

*Devaluation/Extinction Test*: Both *Gria1-/-* mice and WT controls reduced their rate of responding on the lever associated with the devalued outcome compared to responding on the lever associated with the non-devalued outcome (main effect of Devaluation - F(1,23)=9.4 p<.01; see Figure 3a). There was a significant main effect of Genotype (F(1,23)=7.2 p<.05), reflecting higher overall rates of lever pressing in the *Gria1-/-* mice but importantly, however, there was no genotype by devaluation interaction (F(1,23)=0.2, p>0.9 N.S.). There was a main effect of Sex (F(1,23)=10.6 p<.005), reflecting a higher rate of responding in male mice with respect to female mice (males: mean = 13.5 lever presses per min, s.e.m=1.2; females: mean=8.7, s.e.m=0.8 ), but no other significant main effects or interactions (Day F(1,23)=3.1 p>0.8 N.S., Day x Devaluation F(1,23)=3.3 P>0.8 N.S., Day x Devaluation x Genotype F(1,23)=1.4 p>0.2 N.S., Day x Devaluation x Sex F(1,23)=2.8 p>0.1 N.S. All other interactions F’s<1, p’s>0.4 N.S.)

Closer inspection of the data revealed that the significant main effect of devaluation was being driven almost completely by devaluation with the grain reinforcer. Indeed, further analysis of these data with a second ANOVA now including Reinforcer (Grain vs Sucrose) as an additional factor revealed a significant main effect of Reinforcer (F(1,23)=45.5 p<.001), and a significant Reinforcer x Devaluation interaction (F(1,23)=33.4 p<.001), confirming that the mice were responding differently with the two reinforcers. In particular, it was notable that the sucrose reinforcer did not sustain sufficiently high levels of lever responding in either group in order to meaningfully assess devaluation (see Means and S.E.Ms Table 4).

Given this observation, a further separate analysis was conducted which was limited solely to responding on the lever associated with grain rewards, and comparing performance following both grain and sucrose devaluation sessions. Figure 3B depicts the rates of lever pressing on the lever associated with grain reward following devaluation of grain (*Devalued*) compared with rates of lever pressing on the same lever when grain was not devalued (i.e. following prior exposure to sucrose; *Non-devalued*). As in the previous full analysis presented above, both *Gria1-/-* and WT mice significantly reduced their responding on the grain-associated lever following devaluation of grain, compared to responding on the same lever when sucrose rather than grain was provided prior to the test (main effect of Devaluation F(1,23)=11.6 p<.005). Importantly, again there was again no Genotype x Devaluation interaction (F<1 p>0.6 N.S.), suggesting a similar devaluation effect in both groups of mice. Again there was a significant main effect of Genotype (F(1,23)=9.5 p<.01) and of Sex (F(1,23)=6.2 p<.05), reflecting higher response rates overall in knockout and male mice, but no other significant main effects or interactions (Devaluation x Genotype x Sex F(1,23)=1.7 P>0.2, all other main effects and interactions F<1, p>0.3). Importantly, this result demonstrates that *Gria1-/-* mice *can* demonstrate sensitivity to devaluation of a reinforcing outcome. Thus, these mice do have access to representations of the sensory-specific aspects of the rewards, and so *can* be goal-directed in certain situations.

**Experiment 3: *Gria1-/-*** **mice show accelerated development of habitual lever pressing.**

Given these results of Experiment 2, an alternative possibility is that the increased propensity for model-free, habitual behaviour in *Gria1-/-* mice reflects the rate, or the nature, of associative learning such that the knockout animals become habitual faster or more readily than the WT controls. We therefore tracked the development of habitual responding in WT and *Gria1-/-* mice, in a longitudinal study with repeated devaluation tests at different stages during acquisition. Given the results of Experiment 2, in which only lever pressing for grain exhibited a devaluation effect, in this next experiment all animals were trained to lever press for grain rewards and performance was assessed during extinction tests after pre-exposure to either grain or sucrose.

*Acquisition*: Analysis of rates of responding during the first five mins of the three baseline training days prior to each extinction test (Days 2, 6 and 10) revealed no differences in the rates of responding between these three training days, nor any interaction of Training Day with Sex or Genotype (all F’s<1, p’s>.4). Main effects of Genotype (F(1,44)=4.7 p<.04) and Sex (F(1,44)=5.7 p<.05) were observed with statistical analysis reflecting higher levels of responding in knockout and female mice respectively (see supplementary material for acquisition data:Table 2).

*Devaluation and Extinction tests:*

Figure 4A depicts lever responding during the 3 extinction tests that followed devaluation of the grain based reinforcer. When pre-fed with grain pellets just prior to the extinction test sessions, WT mice showed reduced levels of lever pressing, compared to the previous baseline training day, on all three test sessions. In contrast, the *Gria1-/-* mice displayed reduced lever pressing during the first extinction session, but then displayed levels of responding which were comparable to the previous baseline training days during extinction tests 2 and 3.

Statistical analysis revealed a significant difference in the rate of responding between genotypes across the 3 extinction tests following devaluation of the grain reinforcer (Test x Genotype F(2,88)=6.7 p<.005). Post hoc comparisons indicated a significant difference in the rate of responding between *Gria1-/-* mice and WT mice during the last two extinction tests only (*Gria1-/-* mice vs WT mice: Test 1 p>0.5 (NS); Test 2 p<0.01; Test 3 p<.05). Thus *Gria1-/-* mice responded at a higher rate with respect to WT mice following devaluation of the grain reinforcer, but only during the last two extinction tests (conducted after 4 and 6 days of training respectively), and not during the first extinction test conducted after two days of training. There was also a main effect of genotype F(1,44)=4.8 p<.05), but no effect of Sex (F(2,44)=2.2 p>0.1), nor any interaction of Sex with any other factors (F<1 p>0.8). Notably, WT mice did not differ in their rate of responding across the three extinction tests (pairwise comparisons all p’s>0.9), whereas *Gria1-/-* mice significantly increased their rate of responding between the first and last extinction tests (Test 1 vs Test 3 p<.001). These results support the notion that overtraining in *Gria1-/-* mice resulted in a greater contribution of S-R supported lever pressing that was insensitive to devaluation of the grain based reinforcer during the last two extinction tests. ( group 1 t tests).

*Gria1-/-* mice and WT mice were initially sensitive to the devaluation of the contingent grain reinforcer. This is reflected in their ratio of responding during the first extinction test which is less than 1 (group1 t test) for both *Gria1-/-* and WT mice during the first extinction test i.e. responding less than baseline responding (WT mice ratio:, *Gria1-/-* mice ratio: ). Need to discuss.

Figure 4B shows lever responding during the extinction tests following prior exposure to the sucrose reinforcer which acts as a control condition. Both groups actually tended to respond more after pre-feeding with sucrose than on the previous baseline training day (ratio of baseline > 1). Statistical analysis of response rates revealed a progressive increase in responding across the 3 extinction tests (Test F(2,88)=5.2 p<.01), but no differences in rates of responding between WT and *Gria1-/-* mice (Genotype F(1,44)=1.1 p>0.3; Test x Genotype F<1 p>0.7, Test x Genotype x Sex F<1 p>0.9 all N.S.). A main effect of Sex was revealed F(1,44)=4.7 p<.05), reflecting a higher rate of responding in male mice (Male:mean:1.6, s.e.m=0.09; female: mean 1.3. s.e.m=0.09). However, this did not interact with any other factor (Test x Sex F(2,88)=1.7 P>0.1; Sex x Genotype F(1,44)=1.1 P>0.3).

**DISCUSSION**

In this study we provide evidence that *Gria1-/-* mice transition more rapidly from goal-directed to habitual behaviour than WT mice. First, *Gria1-/-* mice were insensitive to prior devaluation of a food outcome during performance on a modified spatial reference memory radial maze paradigm (Experiment 1). These findings are consistent with previous studies assessing action-outcome learning in *Gria1-/-* mice, demonstrating habitual behaviour in the knockouts compared to goal-directed behaviour in wild-types (Johnson et al., 2005; 2007). However, we also show for the first time that under certain conditions *Gria1-/-* mice *can* be goal-directed when presented with two levers, each with independent interval schedules of reinforcement (Experiment 2). Consistent with this ability to engage in goal-directed behaviour, *Gria1-/-* mice also exhibited devaluation of lever responding, and hence evidence of model-based behaviour, after 2 days of training on a single lever, operant task (Experiment 3). However, this experiment also revealed an accelerated development of habits in the *Gria1-/-* mice over the course of further single lever training with respect to WT mice (after 4 and 6 days of training). Taken together, these results demonstrate that although *Gria1-/-* mice are capable of goal-directed behaviour, and thus sensitive to devaluation of the outcome of their actions, their behaviour transitions to habits more readily than for WT controls. These results provide a potentially important insight into the aetiology of the increased habitual, model-free behaviour in *Gria1-/-* mice, which may have relevance to the neurobiological mechanisms underlying the tenacity of delusions in psychosis.

*Gria1-/-* *mice and impairments in goal-directed spatial memory performance*

In Experiment 1 *Gria1-/-* mice and wildtype (WT) controls were trained over 20 days to enter sequentially all six arms of a radial maze that contained either grain based pellets or sucrose solution, as defined by the extramaze spatial cues. Mice were then pre-fed with one of the reinforcers prior to a probe test on the maze, performed in extinction, during which the speeds at which they ran down the arms associated with the devalued and non-devalued reinforcers, were assessed. If *Gria1-/-* mice are sensitive to changes in the value of an outcome and thus goal-directed in their behaviour (Dickinson 1985; deWit and Dickinson 2009), then their speed to run down an arm associated with the devalued outcome should be decreased. This was not the case. Instead, *Gria1-/-* mice were insensitive to the devaluation procedure and their mean speed of running down arms associated with the devalued outcome was unaffected. Thus, the behaviour of the *Gria1-/-* mice was habitual at the time of the extinction test sessions (Dickinson 1985; Balleine and Dickinson 1998).

This failure of *Gria1-/-* mice to demonstrate a difference in running speed for devalued and non-devalued arms was not due to an inability to discriminate between the rewards, or to discriminate between the arms of the maze, or a problem with spatial learning per se. It is clear from the data obtained during the acquisition phase of the radial maze task (see Figure 2a) that both groups ran significantly faster down the sucrose arms compared to the arms containing grain pellets during training. Thus, both groups could discriminate between the two different rewards and between the goal arms. Furthermore, we have shown previously that *Gria1-/-* mice are perfectly capable of learning to discriminate between always-rewarded and never-rewarded goals arms during the standard spatial reference memory task, using the very same radial maze in our laboratory (Schmitt et al., 2003; 2005).

*GluA1 is not essential for forming or accessing a representation of the sensory-specific incentive value of a reward.*

In Experiment 2 we tested whether, under certain circumstances, *Gria1-/-* mice can ever exhibit goal-directed behaviour. To this end we employed an operant choice procedure that has been shown to produce goal-directed behaviour despite extensive training and experience of the contingencies (Dickinson 1985; Colwill and Rescorla 1985; Kosaki and Dickinson 2010). Mice were presented with two levers that were reinforced with independent interval schedules that thus maintain variability in the action-outcome relationship. Both WT and *Gria1-/-* mice exhibited devaluation of an operant response associated with the reward that had been consumed to satiety just prior to the test session.

Thus, *Gria1-/-* mice *can* be goal-directed and so do have knowledge of the action-outcome associations which are sensitive to changes in the value of the outcome. Thus, *Gria1-/-* mice are not solely dependent on a habit-mediated system as findings from previous studies in these mice might have led us to believe (Johnson et al., 2005; 2007). Indeed, it was suggested previously that GluA1 might be required for encoding the relationship between the sensory-specific aspects of reward and their incentive value by contributing to relevant neural circuits including the basolateral amygdala (Johnson et al., 2005; 2007). The data from both Experiment 2 and the first extinction test in Experiment 3 demonstrate that the GluA1 subunit is not essential for goal directed behaviour. Instead, our data suggest that GluA1 deletion alters the balance between goal-directed and habitual forms of behavioural control.

*Changes in the associative learning underlie habitual behaviour in Gria1-/-* mice

Instrumental action theory (Dickinson 1985) and computational theories of behaviour (Dolan and Dayan 2013) posit two systems that mediate behaviour: a goal-directed/ model-based system; and a habitual/ model-free system. These systems are thought to summate in their mediation of an instrumental behaviour but differ in their relative contributions to the action in question. The contribution of each system is considered to be dependent, in part, on the variability of the action-outcome relationship (although other factors may also influence this distribution; see Dolan and Dayan 2013). When the animal can compute a correlation ratio of the action-outcome relationship, it can exhibit goal-directed behaviour. When there is no variability in the action-outcome contingency (or there is extreme variability); animals are unable to compute a correlation ratio and therefore behaviour is driven by contextual stimuli (i.e. behaviour is habitual: note this may also be the case if the variability in the action-outcome contingency exceeds a particular level (Mark’s reference in comment box).

In *Gria1-/-* mice, goal-directed learning may be surpassed by an increased contribution of S-R learning (i.e. habitual control). Several explanations could account for a putative increase in the contribution of S-R behaviour with over-training in *Gria1-/-* mice with respect to WT mice (e.g. in Experiment 3). This increased contribution could simply be due to *Gria1-/-* mice learning at a faster rate than WT mice. *Gria1-/-* mice may therefore reach a state of reduced or zero variability (i.e. asymptotic performance) more quickly, and thus they become habitual more readily. Indeed, there are prior examples of enhanced learning in *Gria1-/-* mice (e.g. Schmitt et al., 2003; Sanderson et al., 2009). It has been suggested that GluA1 deletion can facilitate associative learning by increasing the perceived temporal contiguity between events as a result of reducing short-term habituation processes (Sanderson and Bannerman, 2012). It is possible that an enhancement in associative learning accelerates the transition from goal-directed to habitual behaviour. Alternatively, rather than a quantitative change in the rate of learning, GluA1 deletion could qualitatively alter the balance between action-outcome and stimulus-response learning in favour of the latter. We have shown previously that GluA1 deletion can increase the perceived salience of environmental stimuli above the levels seen for a novel stimulus (e.g. Sanderson et al., 2011) which could selectively facilitate stimulus-response learning. This imbalance between stimulus-response and action outcome learning may only become apparent later in training (e.g. by extinction tests 2 and 3 in Experiment 3).

It might at first appear that the increased propensity for habitual behaviour in *Gria1-/-* mice seen in Experiment 3 could, in part, be a result of the higher rate of behavioural output seen in these animals. Increasing the level of training of an action-outcome association has been shown to result in behaviour which is insensitive to devaluation and thus habitual in nature (Adams 1982). *Gria1-/-* mice responded at a higher rate overall during the acquisition of the single lever task (may need other examples where they do respond at higher rates to keep this included). A greater experience with the instrumental contingencies resulting simply from the increased behavioural output of the *Gria1-/-* mice may have resulted in these mice experiencing the lack of variability in the action-outcome contingencies earlier in training than for the WT mice.

Importantly, the results of Experiment 1 could be used argue against such an account. Although it is difficult to quantify reward rate per unit time on the spatial radial maze task, it is worth noting that the experimental design that we adopted was intended to match exposure to the contingencies in the maze for WT and *Gria1-/-* mice by giving all of the mice 6 forced visits per session (one to each of the 6 goal arms). In fact, if anything, the *Gria1-/-* mice completed marginally fewer runs and ran more slowly than the controls during radial maze training. Thus, their behavioural output (at least as far as the maze task was concerned) was arguably reduced in this particular situation, yet their propensity to develop habitual behaviour was still enhanced.

Further investigation of the neurobiology underlying the accelerated development of habitual behaviour in *Gria1-/-* mice is now required. Lesions and neuropharmacological inactivation studies have implicated the infralimbic prefrontal cortex (Kilcross and Coutureau 2003; Smith and Graybiel 2013; Haddon and Killcross et al 2011; Schmitzer-Torbet et al 2015) and dorsolateral striatum (Yin et al 2006; Smith and Graybiel 2013; Shan et al 2015; Smith and Graybiel 2016; O’Hare et al 2016) in mediating habitual behaviour. It is not known how these neural circuits are altered as a consequence of putative changes in associative learning in *Gria1-/-* mice. Moreover, the role of the dopamine system in these shifts between goal-directed and habitual responding also merits investigation. Previous studies have provided evidence for retarded clearance of extracellular dopamine in the striatum of anaesthetised *Gria1-/-* mice compared to wild-type controls, and our own recent studies have found increased phasic striatal dopamine responses in *Gria1-/-* mice using fast-scan cyclic voltammetry (Boerner et al., in preparation). Future studies could track these striatal dopamine signals during the development of habits in these mice.

*Place cells, grid cells and representing “goal space”*

Finally, the data from our radial maze study (Experiment 1) might have important implications for understanding the functional significance of place cells and grid cells in the hippocampal formation. As mentioned earlier, numerous previous studies have demonstrated normal associative long term spatial memory in *Gria1-/-* mice (Zamanillo et al., 1999; Reisel et al., 2002; Schmitt et al., 2003; 2005). In fact, under some circumstances *Gria1-/-* mice actually form long-term spatial memories faster than WTs (Schmitt et al., 2003; see also Sanderson et al., 2009). This previous demonstration of normal, or even enhanced, associative long-term spatial memory is somewhat surprising given the fact that previous studies have reported both abnormal place cell and grid cell function in *Gria1-/-* mice (Resnik et al. 2012; Allen et al. 2012).

Although there is clearly a strong relationship between the firing of hippocampal pyramidal cells and the spatial location of the animal, it is still not clear precisely what information is conveyed when these place cells fire, nor how that information is used to solve different spatial memory tasks. Previous demonstrations of impaired place and grid cells in *Gria1-/-* mice, despite normal or even enhanced spatial memory performance, appear to break the link between place cell/place field fidelity and spatial memory function (although we should not discount the possibility that place cell and grid cell function might be normal if recorded during performance of these spatial reference memory tasks rather than, for example, in a linear track). Nevertheless, the present data raise the intriguing possibility that, rather than providing a representation of space per se, place cells and grid cells provide a representation of “goal space” (REFS: Behrens, Lisman, Reddish, Dupret, others??? Tank). This might be of particular importance for enabling the hippocampal formation to select between conflicting or competing goals, or response options (Gray and McNaughton 1982; Gray & McNaughton, 2000????; Bannerman et al., 2014; Lisman????), although the hippocampus itself may not be required for goal-directed behaviour per se (Corbit and Balleine paper).

*Conclusions*

GWAS and post-mortem brain studies have suggested an important link between the GluA1 AMPAR subunit, an essential player in certain forms of synaptic plasticity, and schizophrenia. We have shown previously that *Gria1-/-* mice exhibit deficits in short-term habituation. This can, under certain circumstances, lead to inappropriately high levels of attention being paid to environmental stimuli, and thus act as a potential driver of aberrant salience (Sanderson et al., 2011; Barkus et al., 2014), which has been strongly linked to psychosis in disorders including schizophrenia (Kapur, 2003; others??). In addition, we have reported that these deficits in short-term habituation processes can alter associative learning in *Gria1-/-* mice, both qualitatively and quantitatively (REFS). Here we show that GluA1 deletion accelerates the rate at which behaviour transitions from being goal-directed and model based, to being habitual and model free. Taken together, these findings in *Gria1-/-* mice demonstrate that GluA1 dysfunction could contribute not only to aberrant salience and inappropriate associations being encoded, which could form the basis of generating false beliefs, but also to behaviour that is resistant to rapid re-evaluation. Understanding the relationship between these two dysfunctions might provide a window into the tenacity of delusions in schizophrenia.

(2171)

**FIGURE LEGENDS**

*Figure 1:* Schematic representing the radial maze arrangement for assessing (a) standard, spatial reference memory task during which mice learn to discriminate between always rewarded and never-rewarded arms (see Schmitt et al., 2003; 2005), and (b) the modified version of the task used in Experiment 1 during which mice are trained to discriminate between goal arms associated with different reinforcers (grain Noyes pellets versus sucrose solution). The same radial maze was used for both experiments.

*Figure 2:* TO BE COMPLETED!!!

The speed at which mice ran down the goal arms paired with sucrose and grain-based pellets is shown in Figure 2A.

Figure 2A: Running speeds of and WT mice and Gria-1- mice significantly increased over training blocks for both reinforcers.

Figure 2B: Running speed ratio differed between Gria-1-mice and WT mice. WT mice ran significantly slower for the devalued reinforcer whereas Gria-1-mice did not differ from chance running at the same speed for both devalued and non-devalued reinforcers. (Data are means ± S.E.Ms).

TO BE INCLUDED-Simple main effects analysis of this interaction revealed that the effect of food type was significant on blocks 2, 5, 6, 7, 9, 10 (F(1,20) values > 8.7, p values < 0.01). On the remaining blocks the effect failed to reach significance (F(1,20) values < 3.3, p values > 0.08. The effect of block also interacted with sex (F(9,180) = 0.025), reflecting that there was a trend for male mice to run faster than females in the first half of training, but slower than females in the second half. No other main effects or interactions were significant (p values > 0.1).

Figure 2 comment- stop x-axis in panel “a” at 10 (not 12)

Figure 3A

Both Gria-1- mice and WT mice demonstrated a reduced rate of responding specifically on the lever associated with the devalued reinforcer during five min choice tests in extinction This was observed despite extended training with both levers. Rate of responding was calculated as lever presses per min (Data are means ± S.E.Ms).

Figure 3B

Both Gria-1- mice and WT mice demonstrated a reduced rate of responding on the grain-associated lever following devaluation of the grain based pellet despite extended training. This is compared to responding on the same lever during an extinction test where the grain had not been devalued. Rate of responding was calculated as lever presses per min (Data are means ± S.E.Ms).

Figure 4A

Gria-1- mice reduced their rate of lever pressing during single-lever extinction tests 2 and 3 following devaluation of the contingent grain-based pellets. Rate of lever pressing calculated as the rate of responding during the five min extinction test divided by rate of responding during the first five min of the previous training session. (Data are means ± S.E.Ms, dashed lines = rate of responding during training).

Figure 4B

Gria-1- mice and WT showed no suppression of lever–press responding following devaluation of the control sucrose reinforcer. Rate of lever pressing calculated as the rate of responding during the five min extinction test divided by rate of responding during the first five min of the previous training session. (Data are means ± S.E.Ms, dashed lines= rate of responding during training)

Allen, K., Rawlins, J. N. P., Bannerman, D. M., & Csicsvari, J. (2012). Hippocampal place cells can encode multiple trial-dependent features through rate remapping. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, *32*(42), 14752–14766. http://doi.org/10.1523/JNEUROSCI.6175-11.2012

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, *37*(4-5), 407–419.

Bannerman, D. M., Sprengel, R., Sanderson, D. J., McHugh, S. B., Rawlins, J. N. P., Monyer, H., & Seeburg, P. H. (2014). Hippocampal synaptic plasticity, spatial memory and anxiety. *Nature Reviews Neuroscience*, *15*(3), 181–192. http://doi.org/10.1038/nrn3677

Barkus, C., Sanderson, D. J., Rawlins, J. N. P., Walton, M. E., Harrison, P. J., & Bannerman, D. M. (2014). What causes aberrant salience in schizophrenia? A role for impaired short-term habituation and the GRIA1 (GluA1) AMPA receptor subunit. *Molecular Psychiatry*, *19*(10), 1060–1070. http://doi.org/10.1038/mp.2014.91

Colwill, R. M., & Rescorla, R. A. (1985). Instrumental responding remains sensitive to reinforcer devaluation after extensive training. *Journal of Experimental Psychology. Animal Behavior Processes*, *11*(4), 520–536. http://doi.org/10.1037//0097-7403.11.4.520

de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal-human translational models. *Psychological Research*, *73*(4), 463–476. http://doi.org/10.1007/s00426-009-0230-6

Dickinson, A. (1985). Actions and Habits: The Development of Behavioural Autonomy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *308*(1135), 67–78. http://doi.org/10.1098/rstb.1985.0010

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, *80*(2), 312–325. http://doi.org/10.1016/j.neuron.2013.09.007

Gray, J. A., & McNaughton, N. (2003). The Neuropsychology of Anxiety. Oxford University Press. http://doi.org/10.1093/acprof:oso/9780198522713.001.0001

Haddon, J. E., & Killcross, S. (2011). Inactivation of the infralimbic prefrontal cortex in rats reduces the influence of inappropriate habitual responding in a response-conflict task. *Neuroscience*, *199*, 205–212. http://doi.org/10.1016/j.neuroscience.2011.09.065

Hoffman, D. A., Sprengel, R., & Sakmann, B. (2002). Molecular dissection of hippocampal theta-burst pairing potentiation. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(11), 7740–7745. http://doi.org/10.1073/pnas.092157999

Johnson, A. W., Bannerman, D. M., Rawlins, N. P., Sprengel, R., & Good, M. A. (2005). Impaired outcome-specific devaluation of instrumental responding in mice with a targeted deletion of the AMPA receptor glutamate receptor 1 subunit. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, *25*(9), 2359–2365. http://doi.org/10.1523/JNEUROSCI.4146-04.2005

Johnson, A. W., Bannerman, D., Rawlins, N., Sprengel, R., & Good, M. A. (2007). Targeted deletion of the GluR-1 AMPA receptor in mice dissociates general and outcome-specific influences of appetitive rewards on learning. *Behavioral Neuroscience*, *121*(6), 1192–1202. http://doi.org/10.1037/0735-7044.121.6.1192

Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *The American Journal of Psychiatry*, *160*(1), 13–23. http://doi.org/10.1176/appi.ajp.160.1.13

Killcross, S., & Coutureau, E. (2003). Coordination of actions and habits in the medial prefrontal cortex of rats. *Cerebral Cortex (New York, N.Y. : 1991)*, *13*(4), 400–408.

Kosaki, Y., & Dickinson, A. (2010a). Choice and contingency in the development of behavioral autonomy during instrumental conditioning. *Journal of Experimental Psychology. Animal Behavior Processes*, *36*(3), 334–342. http://doi.org/10.1037/a0016887

Kosaki, Y., & Dickinson, A. (2010b). Choice and contingency in the development of behavioral autonomy during instrumental conditioning. *Journal of Experimental Psychology. Animal Behavior Processes*, *36*(3), 334–342. http://doi.org/10.1037/a0016887

Mead, A. N., & Stephens, D. N. (2003). Involvement of AMPA receptor GluR2 subunits in stimulus-reward learning: evidence from glutamate receptor gria2 knock-out mice. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, *23*(29), 9500–9507.

Morris, R. W., Quail, S., Griffiths, K. R., Green, M. J., & Balleine, B. W. (2015). Corticostriatal control of goal-directed action is impaired in schizophrenia. *Biological Psychiatry*, *77*(2), 187–195. http://doi.org/10.1016/j.biopsych.2014.06.005

O'Hare, J. K., Ade, K. K., Sukharnikova, T., Van Hooser, S. D., Palmeri, M. L., Yin, H. H., & Calakos, N. (2016). Pathway-Specific Striatal Substrates for Habitual Behavior. *Neuron*, *89*(3), 472–479. http://doi.org/10.1016/j.neuron.2015.12.032

Reisel, D., Bannerman, D. M., Schmitt, W. B., Deacon, R. M. J., Flint, J., Borchardt, T., et al. (2002). Spatial memory dissociations in mice lacking GluR1. *Nature Neuroscience*, *5*(9), 868–873. http://doi.org/10.1038/nn910

Resnik, E., McFarland, J. M., Sprengel, R., Sakmann, B., & Mehta, M. R. (2012). The effects of GluA1 deletion on the hippocampal population code for position. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, *32*(26), 8952–8968. http://doi.org/10.1523/JNEUROSCI.6460-11.2012

Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J. L., Kähler, A. K., Akterin, S., et al. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nature Genetics*, *45*(10), 1150–1159. http://doi.org/10.1038/ng.2742

Romberg, C., Raffel, J., Martin, L., Sprengel, R., Seeburg, P. H., Rawlins, J. N. P., et al. (2009). Induction and expression of GluA1 (GluR-A)-independent LTP in the hippocampus. *The European Journal of Neuroscience*, *29*(6), 1141–1152. http://doi.org/10.1111/j.1460-9568.2009.06677.x

Sanderson, D. J., & Bannerman, D. M. (2011). Competitive short-term and long-term memory processes in spatial habituation. *Journal of Experimental Psychology. Animal Behavior Processes*, *37*(2), 189–199. http://doi.org/10.1037/a0021461

Sanderson, D. J., Good, M. A., Skelton, K., Sprengel, R., Seeburg, P. H., Rawlins, J. N. P., & Bannerman, D. M. (2009). Enhanced long-term and impaired short-term spatial memory in GluA1 AMPA receptor subunit knockout mice: evidence for a dual-process memory model. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *16*(6), 379–386. http://doi.org/10.1101/lm.1339109

Sanderson, D. J., Gray, A., Simon, A., Taylor, A. M., Deacon, R. M. J., Seeburg, P. H., Sprengel, R., Good, M. A., Rawlins, J. N. P., & Bannerman, D. M. (2007a). Deletion of glutamate receptor-A (GluR-A) AMPA receptor subunits impairs one-trial spatial memory. *Behavioral Neuroscience*, *121*(3), 559–569. http://doi.org/10.1037/0735-7044.121.3.559

Sanderson, D. J., Gray, A., Simon, A., Taylor, A. M., Deacon, R. M. J., Seeburg, P. H., Sprengel, R., Good, M. A., Rawlins, J. N. P., & Bannerman, D. M. (2007b). Deletion of glutamate receptor-A (GluR-A) AMPA receptor subunits impairs one-trial spatial memory. *Behavioral Neuroscience*, *121*(3), 559–569. http://doi.org/10.1037/0735-7044.121.3.559

Sanderson, D. J., Sprengel, R., Seeburg, P. H., & Bannerman, D. M. (2011). Deletion of the GluA1 AMPA receptor subunit alters the expression of short-term memory. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *18*(3), 128–131. http://doi.org/10.1101/lm.2014911

Schmitt, W. B., Deacon, R. M. J., Seeburg, P. H., Rawlins, J. N. P., & Bannerman, D. M. (2003). A within-subjects, within-task demonstration of intact spatial reference memory and impaired spatial working memory in glutamate receptor-A-deficient mice. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, *23*(9), 3953–3959.

Schmitt, W. B., Sprengel, R., Mack, V., Draft, R. W., Seeburg, P. H., Deacon, R. M. J., et al. (2005). Restoration of spatial working memory by genetic rescue of GluR-A-deficient mice. *Nature Neuroscience*, *8*(3), 270–272. http://doi.org/10.1038/nn1412

Schmitzer-Torbert, N., Apostolidis, S., Amoa, R., O’Rear, C., Kaster, M., Stowers, J., & Ritz, R. (2015). Post-training cocaine administration facilitates habit learning and requires the infralimbic cortex and dorsolateral striatum. *Neurobiology of Learning and Memory*, *118*, 105–112. http://doi.org/10.1016/j.nlm.2014.11.007

Shan, Q., Christie, M. J., & Balleine, B. W. (2015). Plasticity in striatopallidal projection neurons mediates the acquisition of habitual actions. *The European Journal of Neuroscience*, *42*(4), 2097–2104. http://doi.org/10.1111/ejn.12971

Smith, K. S., & Graybiel, A. M. (2013). A dual operator view of habitual behavior reflecting cortical and striatal dynamics. *Neuron*, *79*(2), 361–374. http://doi.org/10.1016/j.neuron.2013.05.038

Smith, K. S., & Graybiel, A. M. (2016). Habit formation coincides with shifts in reinforcement representations in the sensorimotor striatum. *Journal of Neurophysiology*, *115*(3), 1487–1498. http://doi.org/10.1152/jn.00925.2015

Voss, M., Moore, J., Hauser, M., Gallinat, J., Heinz, A., & Haggard, P. (2010). Altered awareness of action in schizophrenia: a specific deficit in predicting action consequences. *Brain*, *133*(10), 3104–3112. http://doi.org/10.1093/brain/awq152

Yin, H. H., Knowlton, B. J., & Balleine, B. W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behavioural Brain Research*, *166*(2), 189–196. <http://doi.org/10.1016/j.bbr.2005.07.012>

Zamanillo, D., Sprengel, R., Hvalby, O., Jensen, V., Burnashev, N., Rozov, A., et al. (1999). Importance of AMPA receptors for hippocampal synaptic plasticity but not for spatial learning. *Science (New York, N.Y.)*, *284*(5421), 1805–1811.