SUPPLEMENTARY METHODS

Mice

A total of 73 male R62 mice were used (Cohort 1, used at 11-16 weeks of age, \( n=22 \) R6/2, \( n=25 \) WT; Cohort 2, used at 11-14 weeks of age, \( n=15 \) R6/2, \( n=11 \) WT). They were housed in groups of mixed genotype, six to ten per cage, in a temperature controlled room (22°C) under diurnal conditions (12-h light/12-h dark). All testing occurred at a regular time during the light period. The animals were food deprived so that body weight did not fall below 85% of their free feeding weight throughout the experiment with water available \textit{ad libitum}.

Apparatus

The apparatus consisted of an infrared touchscreen (Craft Data Ltd., Bucks, UK) and a standard modular testing chamber housed within a sound-attenuating box. The box was fitted with a fan, an illuminated pellet receptacle fitted with head entry detectors, a 14mg pellet dispenser, a 3W houselight and a tone generator (Med Associates Inc., Vermont, USA).

The stimuli were presented on the touchscreen monitor. The touchscreen assembly used infrared sensors, so the mouse was not required to exert any pressure on the monitor screen in order for a nose-poke to be detected. A Perspex ‘mask’ containing two response windows approximately 1.6 cm from the floor of the chamber, in which the stimuli were displayed, was positioned in front of the touchscreen in order to stop the mouse accidentally triggering the touchscreen with its tail.

Behavioural Procedures

For both experimental groups, the overall design was similar. Mice were food deprived but maintained so that their body weight was no less than 85% of their free-feeding body
weight. All testing was carried out with the operator blind to the genotype of animals. Preliminary training and behavioural testing were carried out in 12 automated touchscreen testing chambers. **Pretraining.** Mice were trained, through several iterative stages, to touch stimuli on the screen to obtain reward. Initial sessions habituated mice to the apparatus; in subsequent sessions mice learned to associate the delivery of pellets with the sounds of the tone and the onset of the magazine light. Finally, mice learned to 'nose poke' trial-unique stimuli, when one stimulus was presented at a time, in order to obtain food reward.

**Two-choice discrimination task.** Mice were presented with pairs of stimuli, one the correct S+ and the other the incorrect S-. A nose poke to the S+ resulted in a tone, magazine light and a reward pellet. Incorrect responses were followed by a 5-s time-out in which the house light was extinguished, followed by a correction procedure in which the stimulus display was repeated until the mouse made a correct response. Performance was measured by calculating the percent correct choices per session of 30 (non-correction) trials. Two pairs of brightness-matched stimuli (see below) were selected for discrimination training (see below).

Mice in Cohort 1 were given 10 daily sessions of 30 two-choice discrimination trials as described above using pair A (below). On Session 11 the stimuli were reversed (S- became the new S+) and mice were given a further 10 sessions of 30 two-choice discrimination trials with this new stimulus-reward contingency. Following completion of these 10 sessions, mice were given 28 sessions of a new two-choice discrimination.

Mice in Cohort 2 were given 10 sessions of 30 two-choice discrimination trials using Pair A. At Session 11 the S+ and S- were reversed and mice were given a further 12 sessions of 30 discrimination trials. Mice were then given 14 sessions of 30 trials with a novel S+ (pair B below).
**Statistical analysis**

Behavioural data were subjected to analysis of variance (ANOVA, using the Statistics Package for Social Scientists (SPSS) software package, version 13.03.2) with one between subject factor (Genotype) and with repeated measures on one within subject factor (Day/Block of trials) as appropriate. In cases of a significant interaction (Genotype x Age/Day/Block of trials) Sidak’s test was used for multiple independent *post hoc* pairwise comparisons between transgenic and wild-type mice at each relevant age, day or block of trials. A critical value for significance of $P < 0.05$ was used throughout the study.

Pair A

Pair B