

Current state			Action	Additional condition	Reward	New state	
Position	Sound	Enviro.				Sound	Enviro.
N	-	-	-	5s passed since reward consumed	0	L or R or N	-
L (or R)	L (or R)	-	P	lever pressed before within the trial	0	-	Pellet
L (or R)	R (or L)	-	P		0	-	Light
N	N	-	-	animal for 1.5s in the nose poke hole	0	-	Pellet
N	N	-	L (or R or F)		0	-	Light
Any	-	Pellet	F		1	-	-
Any	-	Light	Any	5s passed since light turned on	0	-	-

**Supplementary Table 1. State transitions in the simulations of the actor-critic model.**