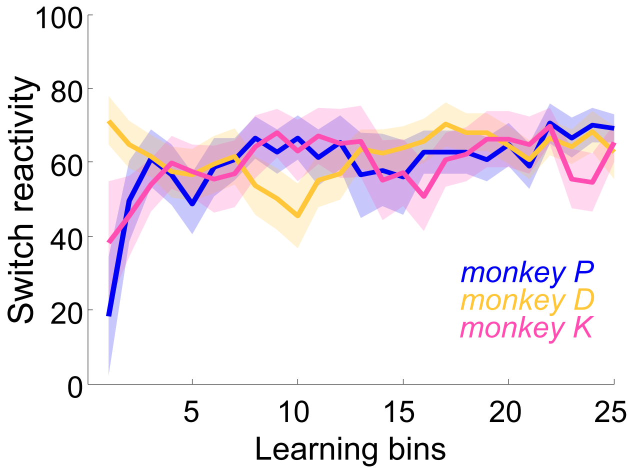

Figure S1. ***Switch reactivity across learning bins in the IT.*** The Switch reactivity score can be thought of as the size of the drop in performance between trials Switch-2 and Switch 0 (noting that Switch-1 is by definition 100%). Monkeys rapidly learn to adapt to Switches in this task, in which they are signaled by new stimuli. Shaded areas represent sem. Monkey P, blue; monkey D, yellow and monkey K, pink.

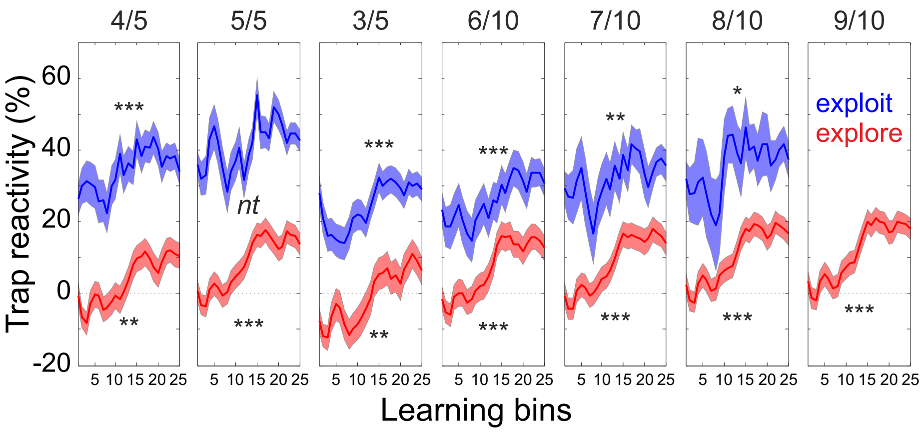

Figure S2. ***Effects of modulation of the exploration/exploitation criterion on Trap reactivity over learning bins in the IT.*** The first panel shows the criterion that is currently used in this paper (4/5 : exploitation when 4 correct trials out of 5). Other panels show the effects of modulation of the number of correct trials used for the criterion (either 5/5 or 3/5) and modulations for another sliding window (6/10 to 9/10). For visibility, data are concatenated accross monkeys. Stars indicate significant increase of Trap reactivity with learning bins (glm, Trials x Learning_bins, ***: p<0.001; **: p<0.01; *: p<0.05; nt: not tested).Statistics are missing forTrap reactivity in the 5/5 panel because all performance are at 100% in exploitation for the 5/5 criterion preventing from using the test. Data are not shown for exploitation with the 9/10 criterion because of an insufficient number of trials. Red: exploration (trials with performance <criterion), blue : exploitation (trials with performance >criterion).

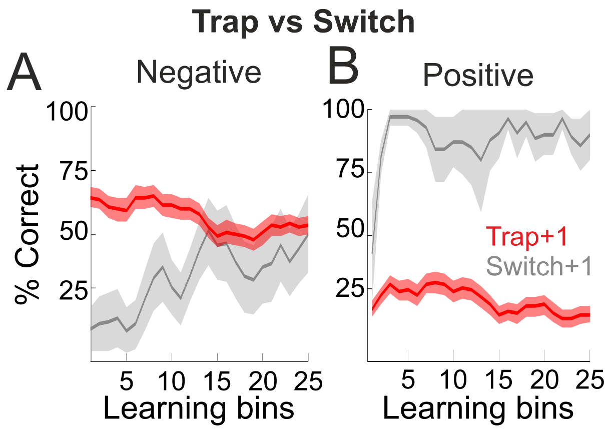

Figure S3. Comparison of performance at trial Trap+1 (red) and Switch+1 (gray) after a **A**: negative or **B**: positive feedback (concatenated data from the 3 monkeys).

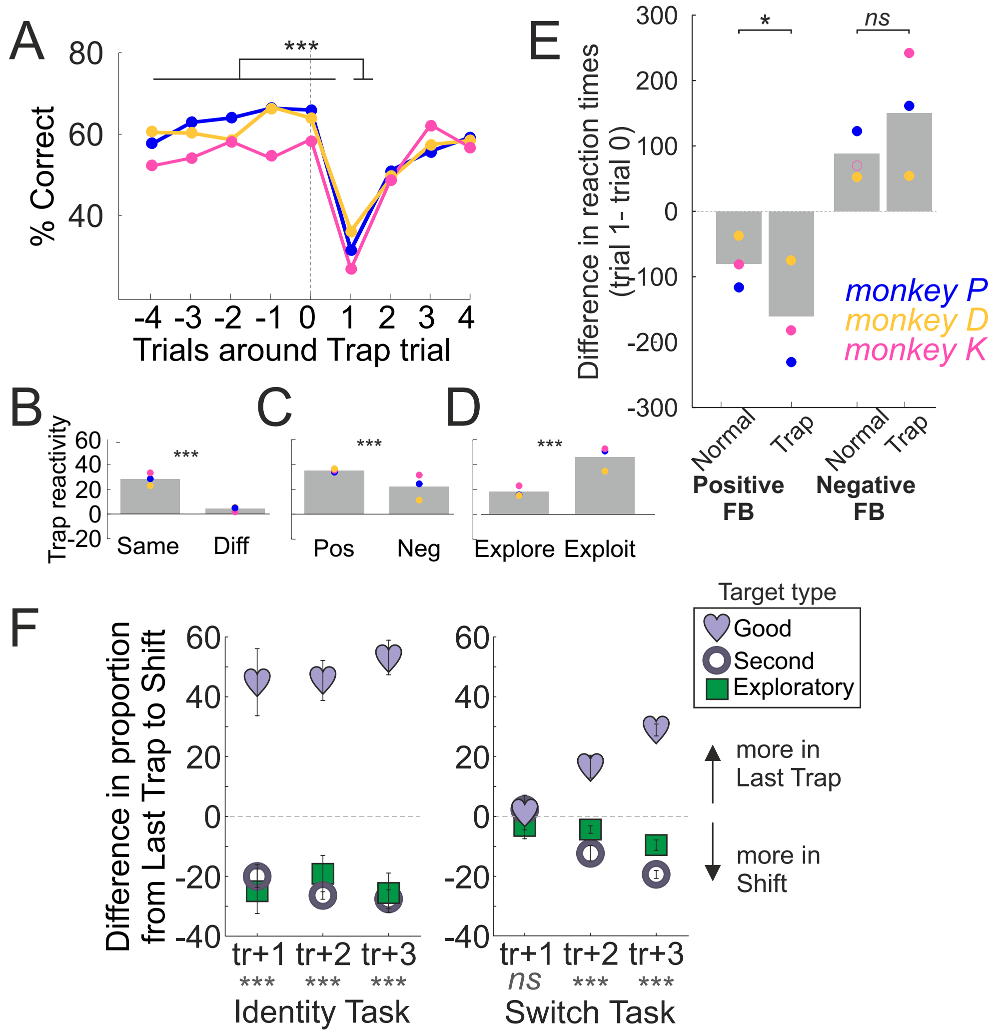

Figure S4. ***Trap reactivity is maintained in ST***: Percentage correct calculated for trials around Trap trial (trial 0). This Trap reactivity maintains similar properties to the IT as follows: **B.** Trap reactivity present for trials of the same stimulus to the Trap, **C.** Trap reactivity greater for Positive than Negative Traps, and **D.** Trap reactivity greater in exploitation. **E.** Effect of feedback on reaction times, with exactly the same calculations and conventions as **Figure 4D**. Post-correct speeding is present and greater after Trap trials. Post-error slowing is present in all but one case (open circle: non-significant difference between trials 0 and 1; closed circles – significant difference with $p<0.001$). Post-error slowing is not significantly increased for Trap trials in this task. ***F. Difference in proportion of chosen target following Last Trap versus Switch, for the 3 following trials, for the 2 tasks.*** The 3 'Target types' refer to the rules of the problem of the Last Trap. "Good" indicates the correct target for the given stimulus; "Second" indicates the other correct target of the mapping, that is incorrect for the given stimulus; and "Exploratory" indicates the third target, that is never correct in the mapping in which the Last Trap occurred. The pattern of target selection after a Last Trap or a Switch did differ between the two tasks. Monkeys immediately differentiated Traps from Switches in the IT, even one trial after unexpected feedback, by using the cue of new stimuli. In the ST, however, this differentiation was absent in trial+1, as should be expected given the Switch is not cued. But after trial+2, monkeys had sufficient information to distinguish Trap from Switch trials, and quickly made the distinction. As such, monkeys were indeed sensitive to the lack of switch cue, but adapted their behaviour so rapidly that there were no significant differences in errors to criterion. P-values correspond to the interaction Proportions x LastTrapOrSwitch. Bars represent sem across monkeys. ***: $p<0.001$; *: $p<0.05$; ns: non-significant.