Neurocognitive Correlates of Learning in a Visual Object Recognition Task

7 AUTHORS, INCLUDING:

Ion Juvina
Wright State University

49 PUBLICATIONS   **336** CITATIONS

Priya Ganapathy
Utopia Compression

11 PUBLICATIONS   **6** CITATIONS

Gautam Kunapuli
Utopia Compression

23 PUBLICATIONS   **114** CITATIONS

Tejaswi Tamminedi
Utopia Compression

10 PUBLICATIONS   **12** CITATIONS

Neurocognitive Correlates of Learning in a Visual Object Recognition Task

Ion Juvina[1], Priya Ganapathy[2], Matt Sherwood[3], Mohd Saif Usmani[4], Gautam Kunapuli[2], Tejaswi Tamminedi[2], Nasser Kashou[4]

[1]Department of Psychology, Wright State University, Dayton, OH, USA
[2]UtopiaCompression Corporation, Los Angeles, CA, USA
[3]Wright State Research Institute, Wright State University, Dayton, OH, USA
[4]Department of Biomedical, Industrial, and Human Factors Engineering, Wright State University, Dayton, Oh, USA
`ion.juvina@wright.edu`

Abstract. Preliminary results of a longitudinal study aimed at understanding the neurocognitive correlates of learning in a visual object recognition task are reported. The experimental task used real-world novel stimuli, whereas the control task used real-world familiar stimuli. Participants practiced the tasks over 10 weeks and reached a high level of accuracy. Brain imaging data was acquired in weeks 2, 6, and 10 and eye-tracking data was acquired in the other seven weeks. Quantitative and qualitative changes in brain activity were observed over the course of learning and skill acquisition. Generally, in the experimental task, brain activity increased at week 6 and decreased at week 10, whereas in the control task, brain activity decreased at week 6 and further decreased at week 10 compared to week 2. New clusters of brain activity emerged at week 6 in the experimental task. Eye-fixation and pupil-dilation data showed that fast learners tend to inspect the stimuli more thoroughly even after a response was given. These results are used to inform the development of computational cognitive models of visual object recognition tasks.

Keywords: Longitudinal study, learning, brain imaging, eye tracking.

1   Introduction and Background

Practice related changes in brain activity have been reported in a variety of laboratory tasks and paradigms. Here we report preliminary results form a study aimed at investigating patterns of brain activity associated with practice and learning in a visual object recognition task. What distinguishes this study from similar ones is the use of real-world, complex stimuli (i.e., military aircraft) and a variety of measurements, such as accuracy, response time, eye fixations, pupil size, and fMRI BOLD signal. The purpose of this investigation is to inform the development of computational cognitive models (Anderson, 2007) that can be used to suggest instructional interventions that maximize learning and engagement and can be embedded in intelligent, adaptive tutoring systems.

Anderson, Betts, Ferris, and Fincham (2010) have put forward a compelling argument for the value of combining behavioral and brain imaging data with computational cognitive modeling for the purpose of informing the development of intelligent tutoring systems.

The most robust result in the literature is that repeated practice is associated with reductions in activity in task-specific and task-general brain regions (Chein & Schneider, 2005). However, this result applies only when there is a consistent mapping between stimuli and responses and no change in strategy is expected to occur with practice and learning. Increases in activity with practice are reported in paradigms that aim to develop skills (e.g. mirror reading, Poldrack & Gabrieli, 2001) or capacities (e.g., working memory, Olesen, Westerberg, & Klingberg, 2004). In these paradigms, strategy shifts are expected and even seen as a desirable effect of training. We expect our task to be more prone to strategizing than simpler laboratory tasks and conceive of strategy learning as an intrinsic component of skill acquisition. We expect participants to develop strategies to inspect stimuli and their features, search for relevant information, encode, keep, and retrieve information in/from long-term memory. These activities will likely be associated with changes in neural activity that can be detected with fMRI. Given the purpose of our study, we are not just interested in quantitative changes such as increases or decreases in activity with learning. Qualitative changes are more interesting because we can learn something about the structure of thought, for example, whether a participant changes strategies or allocates different resources to the task at hand. In addition, since our main purpose is to inform the development of cognitive models, the temporal dimension of task performance is very important. We assume that the brain reacts differently at different stages of skill acquisition. Thus, the question is not whether activation increases or decreases but rather when it increases and when it decreases.

A number of theoretically informed regions of interest (ROI) were defined based on the literature on neural correlates of practice and learning (e.g., Anderson, Bothell, Fincham, Anderson, Poole, & Qin, 2011; Borst & Anderson, 2014; Supekar, Swigart, Tenison, Jolles, Rosemberg-Lee, Fuchs, & Menon, 2013). These were brain regions associated with visual recognition (fusiform gyrus, middle occipital gyrus), manipulations of spatial representations (posterior parietal), storage and retrieval of declarative memories (hippocampus, prefrontal cortex), cognitive control (anterior cingulate), motor control (areas around the central sulcus), automaticity (basal ganglia), and workload (insula). We also used a control ROI that was known a priori to be insensitive to the experimental manipulation – the auditory cortex.

1.1 Tasks

In the experimental task (Figure 1A), participants saw an aircraft image and had to select its name out of four options. There were 75 different aircraft images in total. The control task (Figure 1B) was similar except it used familiar stimuli. There were 52 different control images in total.
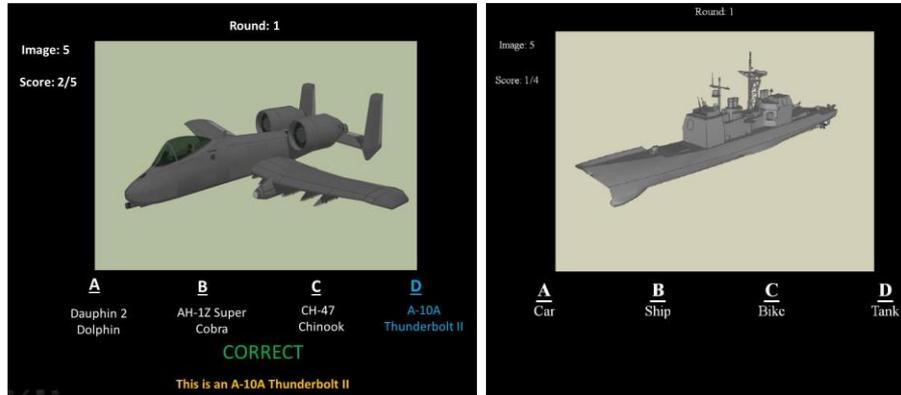
Fig. 1. A (left side): the experimental task and B (right side): the control task

2 Method

2.1 Participants

Fifteen participants were recruited for this study from Wright State University's undergraduate and graduate student population. Throughout the 10-week duration of the study, over 50% of the participants dropped out of the study or were excluded for various reasons. This attrition rate is not uncommon for longitudinal studies.

2.2 Design

Table 1. Layout of the experimental design

Week	Session type	Number and type of rounds	Number and type of trials
1	Behavioral	1 Training + 1 Testing	2 x (75 Air + 52 Ctrl)
2	Imaging	1 Testing	1 x (75 Air + 52 Ctrl)
3	Behavioral	3 Training + 1 Testing	4 x (75 Air + 52 Ctrl)
4	Behavioral	3 Training + 1 Testing	4 x (75 Air + 52 Ctrl)
5	Behavioral	3 Training + 1 Testing	4 x (75 Air + 52 Ctrl)
6	Imaging	1 Testing	1 x (75 Air + 52 Ctrl)
7	Behavioral	3 Training + 1 Testing	4 x (75 Air + 52 Ctrl)
8	Behavioral	3 Training + 1 Testing	4 x (75 Air + 52 Ctrl)
9	Behavioral	3 Training + 1 Testing	4 x (75 Air + 52 Ctrl)
10	Imaging	1 Testing	1 x (75 Air + 52 Ctrl)

A pilot study was run with five participants to determine the length of the study necessary to achieve asymptotic performance in terms of both accuracy and response time for an average participant. Based on what we learned from the pilot study, we decided to include 10 sessions in the main study, one session per week, and insert the brain imaging sessions at weeks 2, 6, and 10. Table 1 shows a schematic of the design.

2.3 Apparatus

Eye fixations and pupil size were recorded using a video-based eye tracker (EyeLink). Neuroimaging was performed using a 1.5 Tesla MR scanner (General Electric Excite HDX; General Electric, Milwaukee, Wisconsin) with an eight-channel head coil. Visual stimuli were projected onto a screen positioned at the foot end of the bore. A mirror affixed to the head coil enabled the participants to view the screen. A fiber-optic button response unit was used to record participants' behavioral performance.

3 Results and discussion

Here we present exploratory analyses and preliminary results, mainly descriptive statistics and visualizations. More detailed quantitative analyses, including inferential statistics, will be presented at the conference and in a subsequent journal paper.

3.1 Pilot study results

The purpose of the pilot study was to get a sense of how many sessions were needed in order to achieve asymptotic performance. In addition, we were interested to learn about the participants' strategies and how they organized their learning in the aircraft task. Figure 2 shows how (A) accuracy and (B) response time changes as a function of session. It is not clear whether the average performance has reached an asymptote, particularly with regard to response time. Based on this observation we decided to increase the number of sessions to 10 to ensure solid learning and skill acquisition.

After task completion, we debriefed the participants with regard to their learning strategies. A variety of strategies were reported such as: directly associating aircraft shapes and names, finding distinctive features of aircraft to aid with distinguishing among similar aircraft, selecting most memorable part of a name (e.g., the word "sea", animal names) and discarding apparently irrelevant parts (e.g., numbers), using features and labels to group aircraft in categories, look for logical associations (e.g., helicopters without wheels have "sea" in their name), give ad-hoc names to particular shapes, etc. We concluded that the task is conducive to strategizing and performance in this task may be a function of how participants organize their learning (strategic learning and executive control) in addition to visual processing and memory per se.
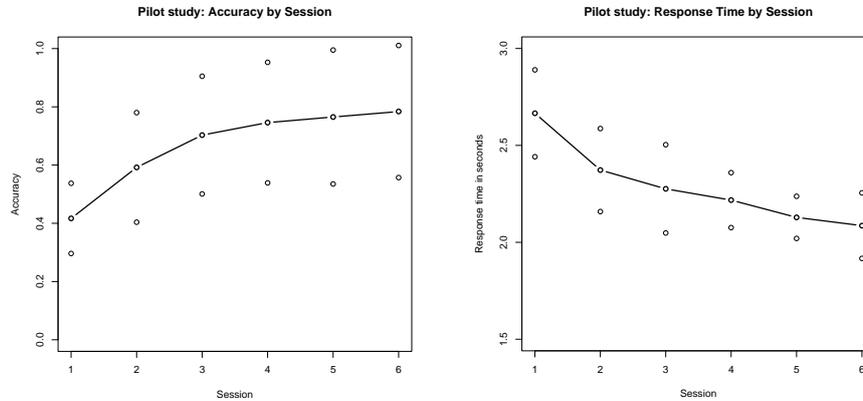
Fig. 2. (A) Accuracy (left plot) as a function of session. (B) Response time (right plot) as a function of session. Error bars are 95% confidence intervals.

3.2 Behavioral results

Extending the study to 10 sessions proved to be a good decision because performance continued to improve after session 6, although it did not reach the maximum possible level (100% accuracy). Since the accuracy variable was somewhat bimodal in distribution, we chose to show descriptive statistics for two ad-hoc groups that we call "fast learners" and "slow learners", respectively. We show the two groups separately to illustrate an interesting difference in learning strategy (see also the section on eye tracking results); we do not claim that the two groups are statistically distinct.

Figure 3 shows how accuracy increases over the 10 sessions, starting from slightly above chance level (25%) in session 1 and ending at almost ceiling level (100%) in session 9 for fast learners. Figure 4 shows reductions in response time associated with practice in the two groups. Performance in the fMRI scanner (sessions 2, 6, and 10) tends to be lower than expected based on the learning trajectory (i.e., accuracy is lower and response time is higher in the scanner than out of scanner). This effect has been attributed to specific attentional and motor deficits caused by the scanner environment and the scanning procedure (Van Maanen, Forstmann, Keuken, Wagenmakers, & Heathcote, 2015).

The tipping point of the separation between the two groups is around sessions 3 and 4, where accuracy of the fast learners increases significantly faster than that of the slow learners. The confidence intervals are also wider around this point, suggesting that the cause of separation between groups is also an important determinant of individual differences among learners. The response time pattern (Figure 4) suggests an interesting strategic difference between the two groups: fast learners tend to be more deliberative; they spend more time inspecting the stimuli in the first 3 sessions than slow learners (see the section of eye tracking results for a corroboration of this interpretation).
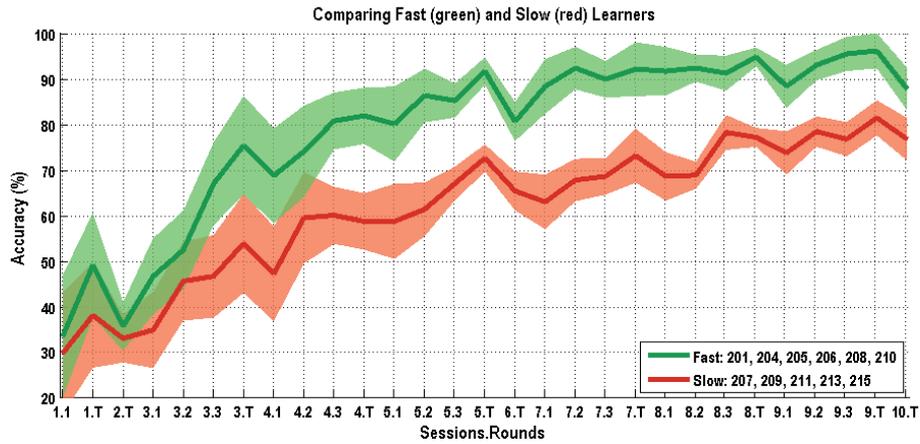
Fig. 3. Accuracy by session and round for fast and slow learners: For each session, the training rounds are indexed numerically (1 to 3) and the testing round is indexed with the letter "T". The shaded areas represent 95% confidence intervals.
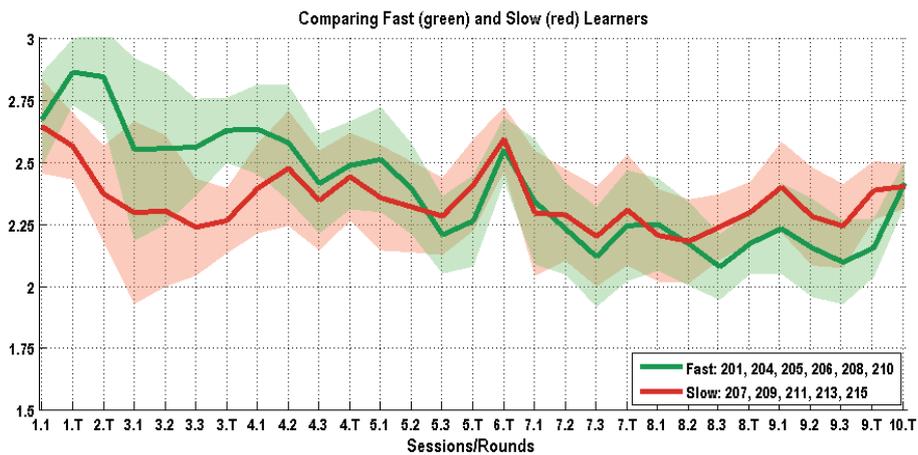

Fig 4. Response time (in seconds) by session and round for fast and slow learners. Eye tracking results

Eye-tracking data provided additional insights into the learning strategies of slow and fast learners. We analyzed the amount of time each participant looked at each of the following five components of the interface: object image, correct name option, incorrect name options, feedback, and progress (i.e., image index and score). We grouped the eye tracking data in two classes based on when it occurred in a given trial: (1) before a response was made and (2) after feedback was provided. Figure 5 shows the eye fixation patterns for (A) fast and (B) slow learners in trials in which they gave correct responses. The interface components of interest are shown on the X-axis. Each vertical bar represents a behavioral session (note: eye tracking data was not collected in the

fMRI sessions). We notice that fast learners look longer at the aircraft image than slow learners, particularly after a response was made. In contrast, slow learners look longer at feedback and progress regions. Another interesting observation is that fast learners tend to increase the time they spend looking at the object as they learn. This increase occurs around sessions 3 and 4, which coincides with the moment in which fast learners clearly differentiate themselves from slow learners. Arguably, the extra time in which fast learners inspect the object after a correct response was given is spent rehearsing and attempting to consolidate their memories by use of mnemonic strategies.
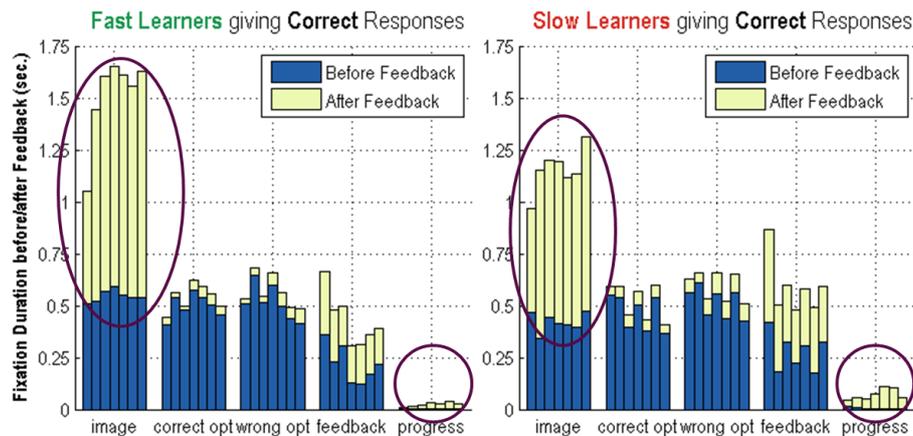

Fig. 5. Eye fixation patterns for fast and slow learners in trials in which a correct response was given.

We also analyzed changes in pupil diameter as a function of task and practice. Pupil dilation has been interpreted as an index of cognitive resources allocated to the task at hand (Granholm, & Steinhauer, 2004; Siegle, Steinhauer, Stenger, Konecky, & Carter, 2003). We assumed that pupil size of a given individual should be higher in the aircraft task than in the control task and took the difference between the two measures (here referred to as pupil dilation for brevity) to represent a normalized index of an individual's resource allocation strategy. Figure 6A shows that pupil dilation is higher in fast learners and peaks around sessions 3 and 4. Paradoxically, in some sessions, slow learners show "negative" pupil dilation, that is, higher pupil size for control vehicles than for aircraft. This is obviously an inappropriate resource allocation strategy that may be responsible for the poorer performance observed in slow learners. Figure 6B shows that pupil dilation in session 1 strongly predicts aircraft accuracy over the entire experiment ($r = 0.63$, $p = 0.04$). Recall from Figure 3 that aircraft accuracy in session 1 was similar in the two groups and close to chance level. This result suggests that the overall performance in this study was determined to a large extent by the participants' willingness to allocate cognitive resources to the aircraft task at the outset of the study.
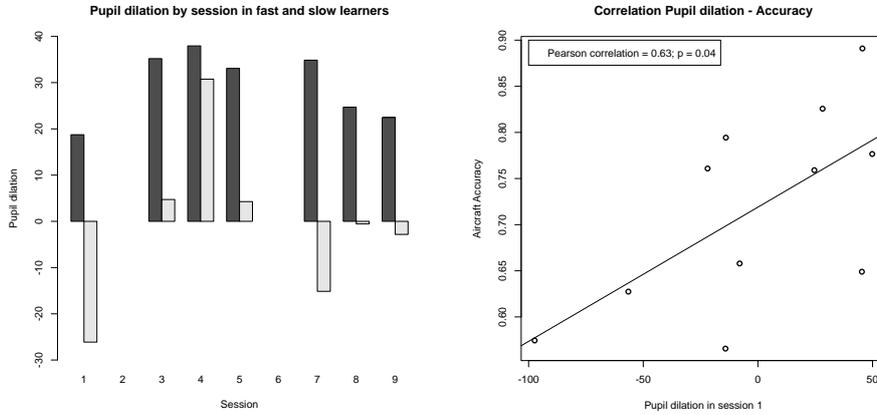
Fig. 6. A (left side): Pupil dilation by session (except fMRI sessions 2, 6, and 10) in fast learners (dark bars) as compared to slow learners (light bars). B (right side): Correlation between pupil dilation in session 1 and overall aircraft accuracy.

3.3 Region of interest analysis

Overall, across all brain regions of interest, in the control task, we see the usual pattern of activation reported in the literature: activation decreases with practice (Chein & Schneider, 2005). In the aircraft task, we see a different pattern: activation increases in the learning phase from week 2 to week 6 and decreases in the consolidation phase from week 6 to week 10 (Figure 7A). This distinction is even clearer in the activation extent data (Figure 7B). These data represent to what extent (percentage) a region is activated above threshold.
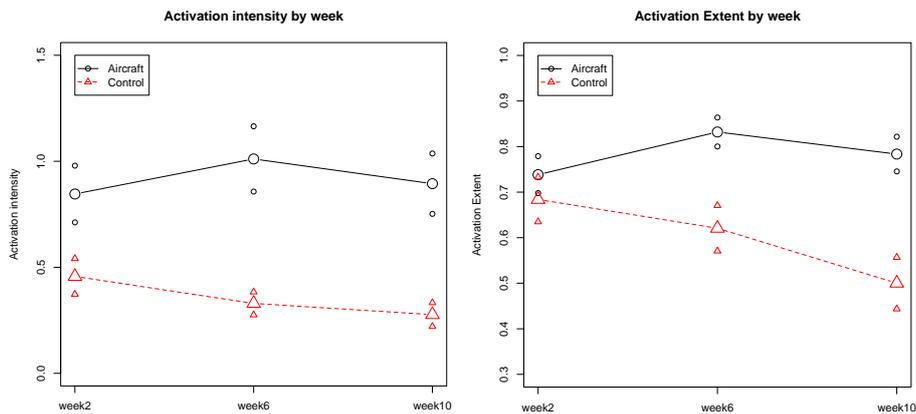

Fig. 7. A (left side): Activation intensity by fMRI session (week). B (right side): Activation extent by fMRI session (week). Error bars represent 95% confidence intervals.

We investigate further the distinction between the aircraft task and the control task in each region of interest (see Figure 8). We see that the distinction is more pronounced in particular regions of interest and almost inexistent in others. The largest differences are seen in fusiform gyrus, middle occipital gyrus, and superior parietal lobule. These are areas involved in developing visual object representations and operating on them. Part of this effect can be a "set size effect" considering that there are much more aircrafts than control vehicles. More interesting are the differences seen in the inferior, middle, and medial frontal gyri thought to reflect retrieval and cognitive control operations.
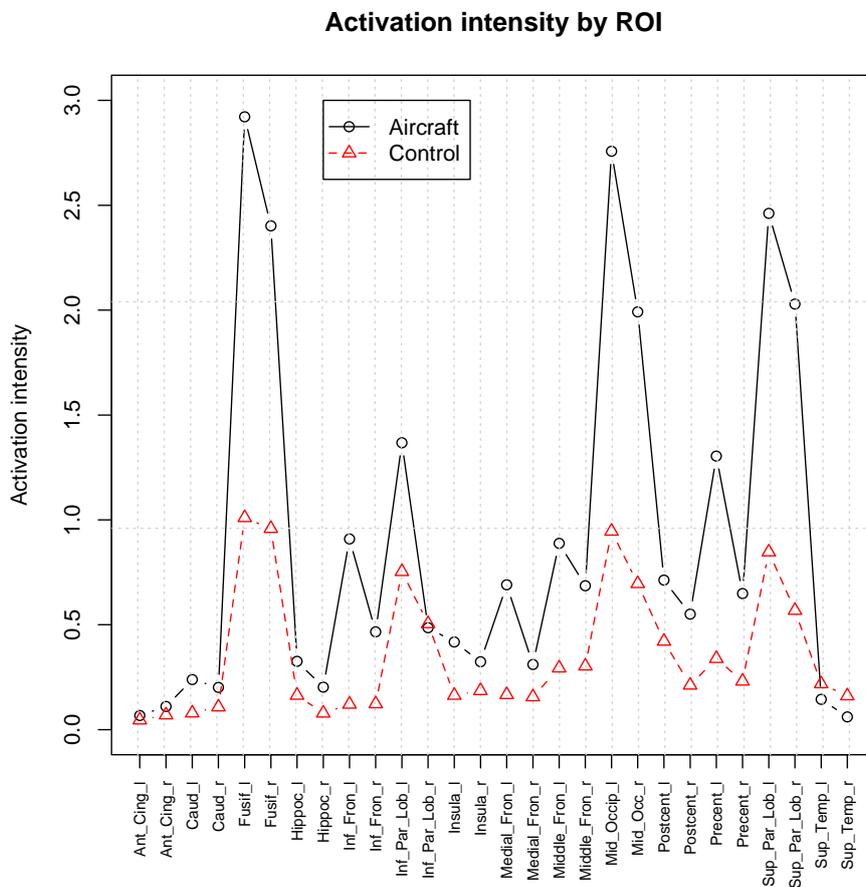
Activation intensity by ROI

Fig. 8. Activation intensity by brain region of interest in the aircraft and control tasks

Thus, there seem to be more than just quantitative differences between the two tasks. The participants seem to employ different strategies in the aircraft task as compared to the control task. The control task seems to be executed in a more automatic way, while the aircraft task is performed in a more controlled and deliberate manner.

Next we analyzed the dynamics of activation in the two tasks as the participants advanced in their learning. In the aircraft task, we notice increases in activation from week 2 to week 6 in caudate and hippocampus reflecting development of procedures (or rules) and declarative memories, respectively. The anterior cingulate cortex (ACC) also shows an increase in activation at week 6. This probably reflects the increased conflict and interference that occur with learning to distinguish between somewhat similar objects and features. The control task shows different dynamics. Activation is much lower in magnitude to start with and decreases with practice. There is not much change of activation in areas related to memory and strategy development or deployment (ACC, caudate, hippocampus, and frontal gyri), reflecting the high level of automaticity reached by this task.

3.4   Whole-brain exploratory analysis

A theoretically agnostic exploratory analysis was performed to uncover potential activation patters that were not found in previous research or may be specific to our task and experimental setup. We identified all clusters of activity (i.e., adjacent voxels activated above threshold) as a function of task (aircraft vs. control) and session (weeks 2, 6, and 10). For each cluster, we determined its size (i.e., number of voxels), identified all the local maxima, and listed the brain regions that these maxima belonged to. The overall pattern of changes in brain activity is consistent with the one observed in the ROI analysis (see section 3.4).

With regard to the content of the clusters, the general trend that emerges across individuals is as follows. The first cluster tends to capture visual and representational areas in the occipital, temporal and parietal lobes, and cerebellum. The second cluster and the subsequent ones capture frontal areas (inferior frontal gyrus, middle frontal gyrus, superior frontal gyrus, and cingulate gyrus), motor areas (precentral and postcentral gyri), and subcortical areas (e.g., thalamus, basal ganglia). In the aircraft task, we see a qualitative shift from week 2 dominated by occipital areas to week 6 and week 10 showing the emergence of the fusiform gyrus and other areas in the temporal lobe (inferior and middle temporal gyri) reflecting complex object representation and memory. Frontal and parietal areas involved in cognitive control, interference resolution, attention, and memory retrieval tend to be activated throughout the entire study in the aircraft task. In the control task, at week 2 the pattern looks much like the aircraft task, but by week 6 we see the emergence of medial frontal gyrus, possibly indicating a strategy shift. By week 10, we see the emergence of posterior parietal areas and the decline of frontal and temporal involvement, which suggests that by week 10 the strategy might be based on direct representations bypassing memory and control.

3.5   Correlations of brain activation and behavioral performance

We examined correlations between brain activity in our predefined regions of interest and behavioral performance (i.e., accuracy in identification of aircraft images). Most of these correlations were not statistically significant due to the very low number of par-

ticipants (7). However, an interesting pattern emerged from the analysis of the correlations from all regions of interest (Figure 9). The correlations tend to be negative at week 2, turn to positive at week 6, and decrease significantly approaching zero at week 10. This pattern is similar for both activation intensity (Figure 9A) and activation extent (Figure 9B).
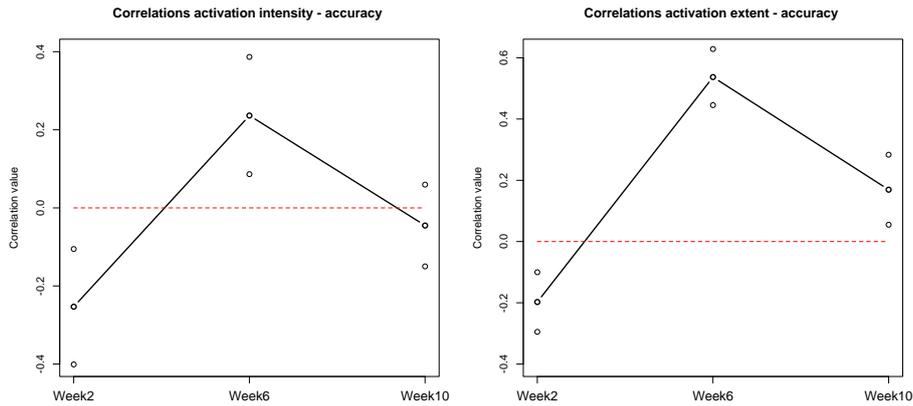

Fig. 9. Pattern of correlations between brain activation (A, left side) intensity and (B, right side) extent and task performance (aircraft accuracy) at start, midpoint, and end of study. Error bars represent 95% confidence intervals.

The negative correlations at week 2 can be interpreted as indicating higher and perhaps inefficient metabolic expenditure for low performers and relatively lower metabolic expenditure for high performers. This can also be associated with higher anxiety or stress for low performers. By week 6, all participants have significantly increased their accuracy. Their brain activity at this level could reflect how much of relevant brain resources participants recruit for the task. The participants who are able to recruit more of these resources perform better, which explains the positive correlation. By week 10, most participants reach a high level of performance (around 80%), the task has become less challenging for everybody, which explains the decrease in the magnitudes of the correlations. Thus, the participants who gradually recruit more brain resources as they learn the task tend to perform better. When their performance reaches a high level, they may decrease the level of metabolic expenditure. The participants who show relatively higher levels of brain activity and relatively lower performance in the early stages of learning tend to learn at a slower rate.

4   Conclusion

The preliminary analyses reported here suggest that learning progresses differently in the two tasks. The control task reaches high levels of automation, which is associated with decreases in most brain regions. The aircraft task recruits more brain resources as the participants learn to master it, particularly from brain regions involved in memory,

representation, and control. Learners would benefit from a tutoring process aimed at aiding memory encoding and retrieval, making correct associations, and resolving ambiguity and interference among representations and associations.

REFERENCES

1. Anderson, J.R. (2007). How Can the Human Mind Occur in the Physical Universe? New York: Oxford University Press.
2. Anderson, J.R., Betts, S., Ferris, J., & Fincham, J.M. (2010). Neural imaging to track mental states while using an intelligent tutoring system. *Proceedings of the National Academy of Sciences*, *107*(15), 7018-7023.
3. Anderson, J.R., Bothell, D., Fincham, J.M., Anderson, A.R., Poole, B., & Qin, Y. (2011). Brain Regions Engaged by Part- and Whole-task Performance in a Video Game: A Model-based Test of the Decomposition Hypothesis. *Journal of Cognitive Neuroscience*, *23*(12): 3983–3997.
4. Borst, J. P., & Anderson, J. R. (2014). Using the ACT-R Cognitive Architecture in combination with fMRI data. An Introduction to Model-Based Cognitive Neuroscience. Springer, New York.
5. Chein, J.M. & Schneider, W. (2005). Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. Cognitive Brain Research, *25*(3), 607-623.
6. Granholm, E., & Steinhauer, S. R. (2004). Pupillometric measures of cognitive and emotional processes. *International Journal of Psychophysiology*, 52(1), 1–6.
7. Olesen, P.J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, *7*(1) 75-79.
8. Poldrack, R.A. & Gabrieli, J.D. (2001). Characterizing the neural mechanisms of skill learning and repetition priming: evidence from mirror reading. *Brain*. *124*:67-82.
9. Siegle, G. J., Steinhauer, S. R., Stenger, V. A., Konecky, R., & Carter, C. S. (2003). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. NeuroImage, 20(1), 114–124.
10. Supekar, K., Swigart, A. G., Tenison, C., Jolles, D. D., Rosemberg-Lee, M., Fuchs, L., & Menon, V. (2013). Neural predictors of individual differences in response to math tutoring in primary-grade school children. *PNAS Early Edition*, 1-6.
11. Van Maanen, L., Forstmann, B.U., Keuken, M.C., Wagenmakers, E-J., & Heathcote, A. (2015). The impact of MRI scanner environment on perceptual decision-making. *Behavior Research Methods*, 1-17.