university of groningen

University Medical Center Groningen

University of Groningen

Relating ACT-R buffer activation to EEG activity during an attentional blink task

van Vugt, M. K.

Published in:
11th International Conference on Cognitive Modeling - ICCM2012

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Early version, also known as pre-print

Publication date:
2012

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
van Vugt, M. K. (2012). Relating ACT-R buffer activation to EEG activity during an attentional blink task. In N. Rußwinkel, U. Drewitz, & H. van Rijn (Eds.), 11th International Conference on Cognitive Modeling - ICCM2012 Universitaetsverlag der TU Berlin.

Relating ACT-R buffer activation to EEG activity during an attentional blink task

Marieke K van Vugt (m.k.van.vugt@rug.nl)
ALICE, Nijenborgh 9
9747 AG Groningen
The Netherlands

Abstract

While a clear relation has been established between ACT-R and activity in fMRI, little is known about whether ACT-R has also correlates in EEG activity. Because of its superior temporal resolution compared to fMRI, EEG could potentially be used to adjudicate between model versions that differ in time courses of module activation, even while generating qualitatively similar patterns of behavioural data. On the other hand, ACT-R could form a much-needed source for hypotheses about interactions between brain areas (synchronization) in EEG data. I discuss a method to find such a mapping between ACT-R and EEG buffers, and apply it to data from an attentional blink experiment (Martens et al., 2006). I show preliminary EEG correlates of ACT-R modules and discuss broader implications of this approach for both cognitive neuroscience and cognitive modeling with ACT-R.

Keywords: EEG; ACT-R; attentional blink; oscillations

Introduction

There is a growing interest in using neural activity to help in constraining cognitive models and for cognitive models to help understand the brain. One of the models for which this brain-to-model mapping has worked very well is the ACT-R cognitive architecture (Anderson, Fincham, Qin, & Stocco, 2008). Multiple experiments have verified this mapping, and conversely, the mapping of ACT-R to fMRI (functional magnetic resonance imaging) has given rise to interesting neural predictions.

Despite the success of the mapping between ACT-R and fMRI, there has not been a comparable mapping between ACT-R and EEG (electroencephalography) data. EEG differs from fMRI in that it has a much higher temporal resolution (on the order of milliseconds) compared to the supra-second resolution of fMRI. This increase in temporal resolution of EEG compared to fMRI is countered by a decrease in spatial resolution. While fMRI is very well-suited to answer questions about what parts of the brain are associated with the different ACT-R modules, EEG could answer questions about differences in their time courses of activation. This is interesting because candidate cognitive models could differ in the time course of activation of various buffers, but this may not yield to observable differences in behaviour. An example of this concerns the question of how long the retrieval buffer takes to turn off after activation. Varying the retrieval buffer's decay time does not lead to qualitatively different predictions for behaviour in most experiments. Nevertheless, these models could potentially still be distinguished with a tool like EEG, which has a very high temporal resolution.

While EEG is most conventionally analyzed in terms of event-related potentials, i.e., the average electric field measured in an electrode in response to a certain event, another way is to examine electrical activity in different frequency bands that need to necessarily be time-locked to an event. It has been proposed that such oscillatory activity can be used to communicate and bind information across different parts of the brain (e.g., Singer, 1993). To have a more comprehensive grasp of EEG activity, we will consider both oscillatory and non-oscillatory EEG in our work.

Although a no mapping has been made between EEG activity and all ACT-R modules, some authors have proposed electrophysiological correlates for the production system that forms the core of ACT-R. For example, Zylberberg, Dehaene, Roelfsema, and Sigman (2011) propose that the ACT-R production system is similar to the Global Neural Workspace hypothesis in that the cognitive system selects productions serially from a set of sensory, memory, and motor options. Selection is mediated by mutual inhibition between neurons that increase in activation until a threshold is reached. Notably, "production selection resembles single decision making" (Zylberberg et al., 2011). A lot is known about the neural correlates of making a single decision between multiple alternatives which provides hypotheses for the neural correlates of production selection ("deciding" between productions). I have previously proposed that evidence accumulation is associated with power of oscillatory activity in the 4–9 Hz theta band in EEG (van Vugt, Simen, & Cohen, 2011) and crossing a threshold with the Lateralized Readiness Potential (an EEG potential consisting of the imbalance between the left and right-hemisphere central electrodes C3 and C4 that is thought to arise from motor cortex, see Figure 1; Simen, van Vugt, Balci, Freestone, & Polk, 2010; van Vugt, Simen, Nystrom, Holmes, & Cohen, submitted). Simen et al. (2010) also proposed that production selection would be associated with the Lateralized Readiness Potential.

In this study, we look for the electrophysiological correlates of a larger set of ACT-R modules in an attentional blink task, for which a well-established ACT-R model exists (Taatgen, Juvina, Schipper, Borst, & Martens, 2009). In an attentional blink task (Luck, Vogel, & Shapiro, 1996) participants see a very rapid stream of visual stimuli, and have to detect what letters were presented in this stream of digits. The main finding of interest in this task is that while participants can see two letters if they occur far apart or in direct succession, they often fail to see the second letter if it is separated from the first by one or two intervening digits. It is as if attention blinks after seeing the first letter. ACT-R accounts for the attentional blink phenomenon by assuming there is an over-exertion of control. If, when a target is recognized in the stream of stimuli, a control rule is triggered in the production
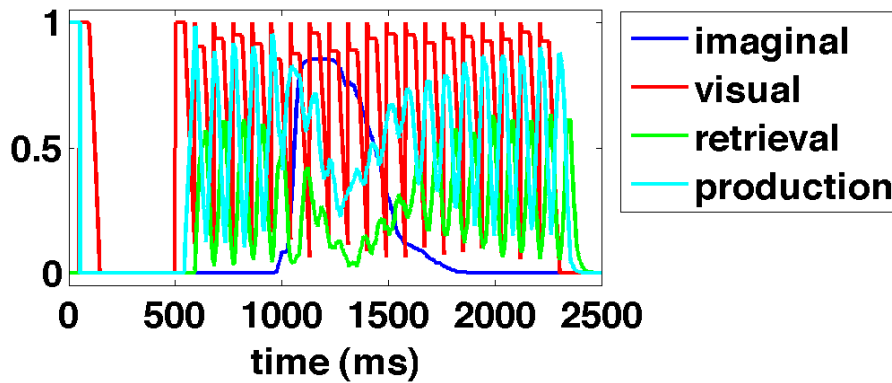
Figure 1: Example module activation probabilities for the ACT-R model of the attentional blink (lag 3, correct trials). Blue: imaginal module. Red: visual module. Green: retrieval module. Cyan: production module. These module activations were used to create the ACT-R regressors that were correlated with the EEG data.

module that suspends target detection, then this can create an attentional blink because the imaginal buffer is not open for receiving another target to consolidate during the "suspend target detection" time.

There have been two main findings in EEG studies of the attentional blink: an increase in the P3 event-related component (the P3 is a positive potential occurring approximately 500 ms after a stimulus onset at parietal electrode sites), and a decrease in gamma oscillation synchronization. The increase in the P3 has also been associated with an increase in 4–9 Hz theta oscillation reset, and has been thought to reflect over-investment of attentional resources in the first target stimulus (Slagter et al., 2007), This phenomenon may be similar to the over-exertion of control posited by the ACT-R model, and may be associated with the imaginal module. The decrease in gamma synchronization was predicted by the Global Neural Workspace model by Dehaene, Sergent, and Changeux (2003), which as discussed above, shares conceptual commonalities with ACT-R. Gamma oscillations are periodic activity observable in the EEG at a frequency of 28–90 Hz. Gamma oscillations have been associated with many things, including visual attention and consciousness (Varela, Lachaux, Rodriguez, & Martinerie, 2001). According to Dehaene's model, when gamma synchronization decreases, it makes the visual stimulus less accessible to consciousness (Gaillard et al., 2009), and hence the participant will frequently fail to report that s/he has seen the stimulus. In terms of ACT-R, this may reflect an inability of the visual stimuli to enter the imaginal buffer.

My goal is to examine whether we can find neural correlates of ACT-R during the attentional blink in EEG data. Guided by the above observations, I predict that activation of the imaginal module, which is crucial for the attentional blink effect, is correlated with 4–9 Hz theta oscillations and the P3 EEG component. I further predict that the gamma synchronization decrease that is also associated with the attentional blink reflects a disconnection between the visual module and

the retrieval module, such that items entering the visual module cannot be compared to memory (chunk activation from items in the visual buffer cannot spread to chunks in declarative memory during a retrieval request). Nevertheless, in testing these hypotheses, I will look at all frequency bands because there exist other plausible hypotheses and the field is relatively unexplored.

Methods

Task: I used existing data from an attentional blink task (Martens, Munneke, Smid, & Johnson, 2006) to study the electrophysiological correlates of ACT-R. In this task, participants see a very rapid stream of visual stimuli, presented for 90 ms each. Their task is to report whether there are letters present in the stream, and if so, which letters those are. The data reported here are from the 14 blinkers in the study by Martens et al. These EEG data were collected at the University of Groningen with a 64-channel EEG system (Twente Medical Systems, Enschede, The Netherlands) and a sample rate of 250 Hz.

Analysis: EEG data were analyzed with the EEG toolbox, a set of Matlab scripts developed in the laboratory of Michael Kahana (e.g, van Vugt, Schulze-Bonhage, Litt, Brandt, & Kahana, 2010) and custom-written scripts. I used this toolbox to extract data for every channel in our EEG setup. I concatenated the time series for each trial lengthwise into one long time series to be correlated with the ACT-R model time series. I then used Morlet wavelets (van Vugt, Sederberg, & Kahana, 2007) to create representations of the EEG data in six distinct frequency bands: 2–4 Hz delta, 4–9 Hz theta, 9–14 Hz alpha, 14–28 Hz beta, 28–48 Hz low gamma and 48–90 Hz high gamma (van Vugt et al., 2010). For this frequency-transformed data, I used the same concatenation procedure to create time series of the trial EEG for each standard frequency band.

To correlate ACT-R's predicted module dynamics to EEG data, I created regressors (van Vugt et al., 2011). Regressors

Table 1: Predictions for the neural correlates of ACT-R modules based on the cognitive neuroscience literature. Note that the speech module would create large artifacts in EEG activity, making it difficult to find correlates for this buffer. Reported fMRI correlates are based on Anderson et al (2008) and Borst et al (2011).

[1] Keeping track of subgoals and intentions. [2] ACC = Anterior Cingulate Cortex. [3] Storing and retrieving declarative information. [4] Auditory perception.

Module	fMRI region	EEG component
Motor module	Motor cortex	Central beta oscillations
Vision module	Fusiform gyrus	Posterior gamma oscillations
Imaginal module[1]	Parietal cortex (Intraparietal Sulcus)	Parietal theta oscillations
Goal module	ACC[2]	Frontal theta/gamma oscillations
Retrieval module[3]	Lateral inferior prefrontal	Hippocampal theta oscillations
Speech module	Artifacts	Artifacts
Aural module[4]	Secondary auditory cortex	Central gamma oscillations
Production selection	Head of caudate	Lateralized Readiness Potential

are fMRI terminology for a time series of interest that is used as the independent variable in a regression to find pieces of neural data that correspond to these dynamics. In this case, the data patterns of interest are ACT-R module activations (visual, production, retrieval, and imaginal). I ran the attentional blink ACT-R model (Taatgen et al., 2009) 250-350 times (corresponding to the number of trials in the dataset) and computed the average activation for different model conditions: lag 3 and 8, and correct and incorrect responses. These average activations therefore reflect the probability of a module being active. 'Lag' refers to the number of stimuli between the first and second target (letter) in the digit stream that the participant has to remember. An attentional blink is likely to occur for lag 3, but not lag 8 trials. Correct trials refer to trials in which both targets were reported correctly. Trials in which the first target was missed were removed from the analysis because in that case it is not clear what the reason is for missing the second target if that occurs.

For every trial that a participant did, I inserted the averaged module activation for the condition corresponding to that trial. This led to an activation time series during the whole tasks for every ACT-R module that, after subsampling to the EEG sample rate (250 Hz), had the same length as the EEG data. These were the time series that I could use to regress the EEG time series on, to obtain for every module an estimate of how well it correlated with the different frequency bands, and which channels were most strongly involved in this correlation. Instead of using a simple multiple regression, I used a canonical correlation analysis, which finds weights on the regressors (electrode time series) that maximize the correlation between the regressors and the ACT-R time series. Note that only the EEG data were wavelet-transformed–the ACT-R activation time courses were not.

Predictions: I correlated the regressors representing ACT-R's module dynamics not only with raw EEG, but also with oscillatory EEG. Oscillatory activity has two advantages over event-related raw EEG activity: the mechanisms that produce field potential oscillations are well-understood, and because many phenomena get lost in averaging, time-frequency de-

Table 2: Cross-correlations for the various ACT-R module time courses, averaged across participants. The activation time course of the imaginal module is most different from the other modules, as would be predicted from the module time courses in Figure 1.

	Imaginal	Production	Retrieval	Visual
Imaginal	1.0000	0.3922	0.3057	0.3880
Production	0.3922	1.0000	0.2138	0.5046
Retrieval	0.3057	0.2138	1.0000	0.5238
Visual	0.3880	0.5046	0.5238	1.0000

composition allows for a more comprehensive and in-depth picture of the data (Cohen, Wilmes, & van de Vijver, 2011). It is well possible that different modules are associated with different oscillatory frequency bands. Table 1 shows my hypotheses about correlations between ACT-R modules and components of EEG activity based on the EEG literature.

Results

The basic behavioural and EEG data for this task are reported in Martens et al. (2006), who showed a classic attentional blink effect (dip in accuracy for the second target letter when it followed the first target letter with only 1 or 2 items in-between). This was accompanied by an increased P3 EEG component for blinked compared to non-blinked trials.

Figure 1 shows an example of average ACT-R module activation on a single trial for the lag 3/correct condition. These activation time courses were used to make regressors that could be used to extract corresponding patterns from our EEG data. I correlated these regressors with both the raw EEG data and oscillatory data in the different frequency bands. Figure 2 shows the resulting correlations between each module and EEG activity for all participants who showed evidence of an attentional blink in the task. Note that the activations of the visual, retrieval and production modules are highly correlated (Table 2; see also Borst, Taatgen, and Rijn (2011, for a discussion)) and are therefore expected to have very similar neural correlates. The highest correlations occur with activity
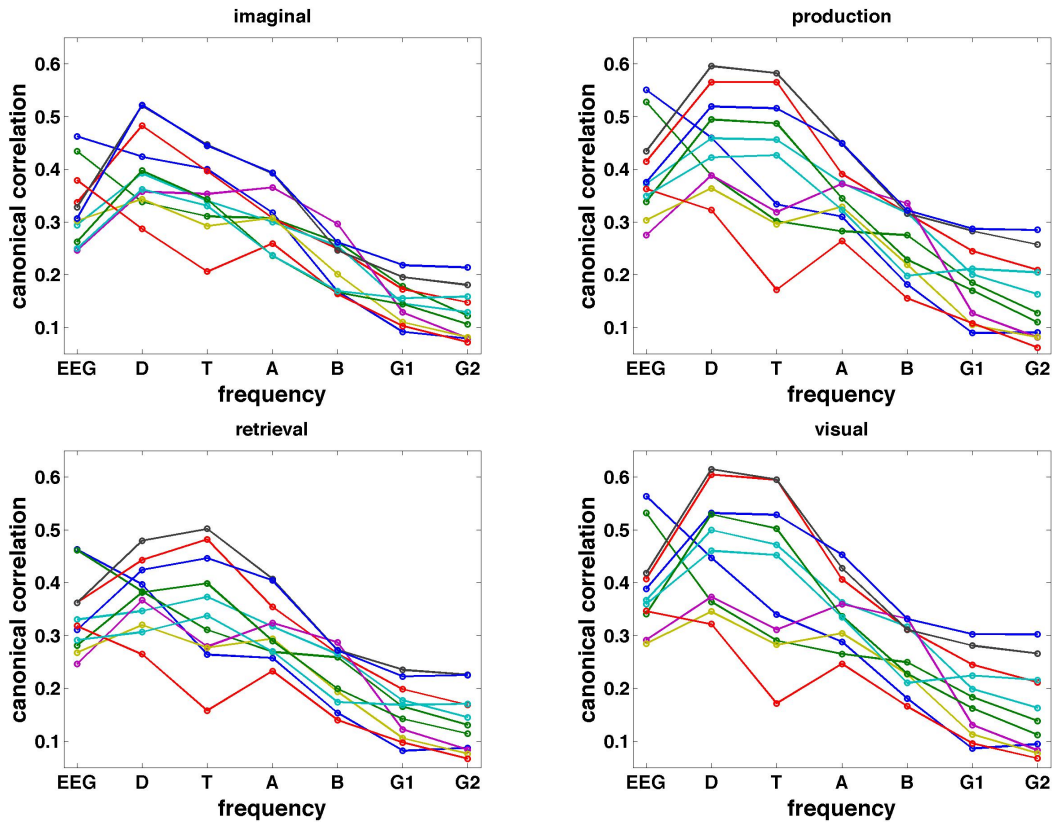
Figure 2: Canonical correlation between ACT-R module time courses and EEG time courses for raw EEG ("EEG"), 2–4 Hz delta ("D"), 4–9 Hz theta ("T"), 9–14 Hz alpha ("A"), 14–28 Hz beta ("B"), 28–48 Hz low gamma ("G1") and 48–90 Hz high gamma ("G2") activity. Different lines reflect different participants. As expected from the correlations between module activations, the imaginal module shows a different pattern (most prominent correlation with EEG activity in the delta band) from the other modules (most prominent correlation with EEG activity in the theta band).

in the 4–9 Hz theta band for the visual, retrieval, and production modules. For the imaginal module, we additionally see a fairly high correlation with activity in the 2–4 Hz theta band as well. Overall canonical correlations are lowest for the retrieval module. Interestingly, there are three participants for whom raw EEG activity shows the highest canonical correlation, while for all others oscillatory EEG shows the highest canonical correlation. In contrast to our expectations, the different modules in this experiment do not exhibit correlations with distinct frequency bands. Part of this may be due to the relatively high correlations between module activation time courses.

I then examined the topographies associated with the magnitudes of ACT-R–EEG correlations in the different frequency bands. While the correlations look quite similar across modules and frequency bands, the topographies in Figure 3 show more variation. In this graph, I chose for each module a frequency band based on either the magnitude of the correlation of the EEG with the module activation in Figure 2 or based on the hypotheses in Table 1. I found that the imaginal module correlate in the 4–9 Hz theta band was primarily associated with right-lateral activation that could be

consistent with a parietal source as I expected. The production module correlate in the 4–9 Hz theta band was found predominantly in superior channels that are in the same location as where Lateralized Readiness Potentials are observed. The retrieval module correlate in the 2–4 Hz delta band showed a negative correlation in frontal channels, consistent with a correlate of the retrieval module in frontal cortex (but unlike my prediction of hippocampal theta oscillations, although those are virtually impossible to observe on the scalp). It also showed a positive correlation with right-lateral channels. Finally, the visual module correlate in the 28–48 Hz gamma band had a central topography, which is quite different from the occipital locus I expected for this module. The gamma band correlate was also much weaker than the correlation in the delta and theta bands, which may therefore be much more likely correlates of this buffer.

Discussion

I have proposed a new method to find the electrophysiological correlates of ACT-R module activations, and shown that different buffers show different patterns of correlation with EEG data. Not all the predictions in Table 1 have been ver-

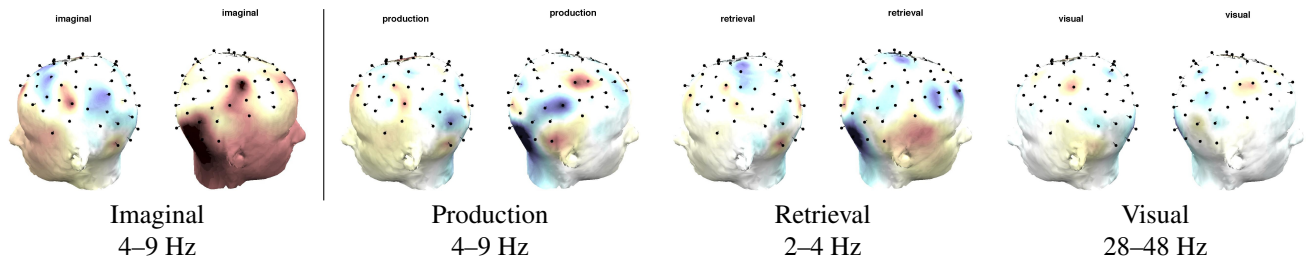| Imaginal	Production	Retrieval	Visual
4–9 Hz	4–9 Hz	2–4 Hz	28–48 Hz

Figure 3: Topographical representation of electrode weights from the canonical correlation analysis for the different modules. The choice of frequency band was guided by the canonical correlation observed in Figure 2. Plotted are the magnitudes of the canonical correlation weights across the brain for the canonical correlation at the respective frequency and with the respective module. Positive weights are red and negative weights are blue.

ified. The data reported here support the idea that the imaginal module is associated with parietal theta oscillations, but do not support the mappings for visual and retrieval modules. The visual module seems to instead be correlated more strongly with frontal gamma oscillations, just like the retrieval module. Interestingly, this mapping of the retrieval module is fairly consistent with fMRI-based localization. The association of the production module with central theta oscillations could be consistent with an association with Lateralized Readiness Potentials, but this should be tested more explicitly. What further remains to be done is applying the same methods to different tasks and to optimize artifact detection methods. In particular, it is important to look at tasks in which the production, retrieval and visual module are not as highly correlated as they are in the attentional blink task discussed here, such that their neural correlates can be pulled apart. Only this will allow us to make claims about the electrophysiological correlates of ACT-R's various resources.

Areas that warrant further investigation are modeling individual differences and examination of the neural correlates of module interaction. Individual differences could solidify our confidence in the mapping between modules and EEG activity. If individual differences are modeled in ACT-R (e.g., Lovett, Daily, & Reder, 2000) and if those individual differences correlate with individual differences in those participants' electrophysiology, then this makes the EEG–ACT-R-module relation more specific (see van Vugt et al., 2011, for an application of this approach to perceptual decision making). In other words, if individual differences covary with dynamics of the neural ACT-R modules, that could greatly increase our confidence in the accuracy of our mapping of ACT-R to electrophysiological brain activity.

Once the electrophysiological correlates of ACT-R have been determined, a large area of new research is opened up. An advantage of the fact that ACT-R consists of multiple modules is that their interaction provides a principled way to look for patterns of synchronization in EEG activity. If synchronization reflects information transfer between the modules (Buzsáki, 2006; Singer, 1993), then increases of synchronization should occur in specific frequency bands and between specific sets of electrodes that correspond to the re-
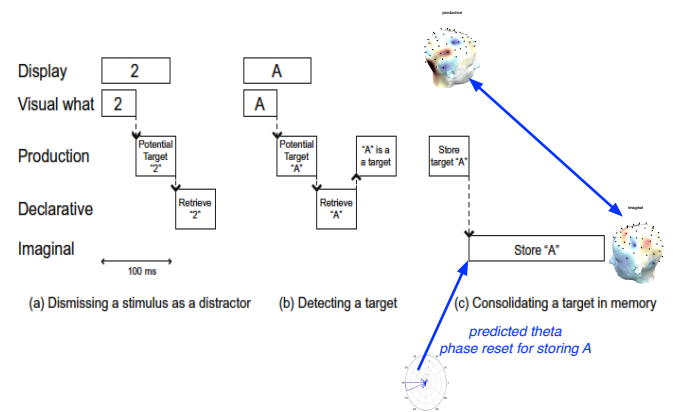

Figure 4: Predictions for neural correlates of interaction between modules in the Attentional Blink task (adapted from Taatgen et al., 2009). At the time of storing the letter "A" in the imaginal buffer, there should be phase reset (observable in the phase histogram) and subsequent synchronization in the theta/alpha band between the channels corresponding to the production and imaginal modules. These sets of channels were based on Figure 3.

spective modules. For the modules in Figure 4, for example, I predict that after every stimulus presentation, there should first be increased synchronization between the neural correlates of the visual module and the production module, and then between the production and declarative module (panel a). When a target is stored successfully, but not when it is not, there should be increased synchronization between the production and the imaginal module (panel c).

I believe that relating ACT-R to EEG activity is a fruitful endeavor that could eventually also have interesting implications for ACT-R modeling. For example, there might be subtle differences in module activation that may not lead to observable differences in behavior. If we could observe module activation time courses in EEG, this could potentially al-

low us to distinguish between these different ACT-R models. While fMRI has had tremendous success in defining brain regions associated with different ACT-R modules, it does not have the temporal resolution on millisecond-scale to compare different time courses of module activation. EEG could fill this gap. Moreover, EEG is much better suited to capture brief interactions between ACT-R modules, which would be too short for fMRI to detect.

In conclusion, I have outlined methods to study the neural correlates of ACT-R in electrophysiological data. I have also shown how they work in the case of an attentional blink task, and how different ACT-R modules can be associated with specific frequency bands and topographies observable in EEG data. I argue that these methods can lead to a wealth of understanding on how the time courses of ACT-R models develop over time. Moreover, they could provide neuroscientists with directly-testable hypotheses about interaction between different neural populations.

Acknowledgments

References

Anderson, J. R., Fincham, J. M., Qin, Y., & Stocco, A. (2008). A central circuit of the mind. *Trends in Cognitive Sciences*, *12*(4), 136–143.

Borst, J., Taatgen, N., & Rijn, H. van. (2011). Using a symbolic process model as input for model-based fmri analysis: Locating the neural correlates of problem state replacements. *NeuroImage*, *58*(1), 137-147.

Buzsáki, G. (2006). *Rhythms of the brain*. New York: Oxford University Press.

Cohen, M. X., Wilmes, K. A., & van de Vijver, I. (2011). Cortical electrophysiological network dynamics of feedback learning. *Trends in Cognitive Sciences*, *15*(12), 585-566.

Dehaene, S., Sergent, C., & Changeux, J.-P. (2003). A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proc. Nat. Acad. Sci., USA*, *100*(14), 8520-8525.

Gaillard, R., Dehaene, S., Adam, C., Clémenceau, S., Hasboun, D., Baulac, M., et al. (2009). Converging intracranial markers of conscious access. *PLoS Biology*, *7*(3), e1000061.

Lovett, M. C., Daily, L. Z., & Reder, L. M. (2000). A source activation theory of working memory: cross-task prediction of performance in ACT-R. *Cognitive Systems Research*, *1*(2), 99-118.

Luck, S. J., Vogel, E. K., & Shapiro, K. L. (1996). Word meanings can be accessed but not reported during the attentional blink. *Nature*, *383*, 616-618.

Martens, S., Munneke, J., Smid, H., & Johnson, A. (2006). Quick minds dont blink: Electrophysiological correlates of individual differences in attentional selection. *Journal of Cognitive Neuroscience*, *18*(9), 1423-1438.

Simen, P., van Vugt, M. K., Balci, F., Freestone, D., & Polk, T. (2010). Toward an analog neural substrate for production systems. In *Proceedings of the international conference in cognitive modeling*.

Singer, W. (1993, March). Synchronization of cortical activity and its putative role in information processing and learning. *Annu Rev Physiol*, *55*, 349-374.

Slagter, H. A., Lutz, A., Greischar, L. L., Francis, A. D., Nieuwenhuis, S., Davis, J. M., et al. (2007). Mental training affects distribution of limited brain resources. *PLoS Biology*, *5*(6), e138.

Taatgen, N. A., Juvina, I., Schipper, M., Borst, J. P., & Martens, S. (2009). Too much control can hurt: A threaded cognition model of the attentional blink. *Cognitive Psychology*, *59*(1), 1–29.

van Vugt, M. K., Schulze-Bonhage, A., Litt, B., Brandt, A., & Kahana, M. J. (2010). Hippocampal gamma oscillations increase with working memory load. *Journal of Neuroscience*, *30*, 2694–2699.

van Vugt, M. K., Sederberg, P. B., & Kahana, M. J. (2007). Comparison of spectral analysis methods for characterizing brain oscillations. *Journal of Neuroscience Methods*, *162*(1-2), 49–63.

van Vugt, M. K., Simen, P., & Cohen, J. D. (2011). Finding neural correlates of drift diffusion processes in eeg oscillations. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the cognitive science society* (p. 2439-2444). Austin, TX: Cognitive Science Society.

van Vugt, M. K., Simen, P., Nystrom, L. E., Holmes, P., & Cohen, J. D. (submitted). Lateralized readiness potentials reveal properties of a neural mechanism for crossing a decision threshold. *submitted*.

Varela, F., Lachaux, J. P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, *2*(4), 229-239.

Zylberberg, A., Dehaene, S., Roelfsema, P. R., & Sigman, M. (2011). The human Turing machine: a neural framework for mental programs. *Trends in Cognitive Sciences*, *15*(7), 293–300.