Term Paper

Shu Hu

Georgia State University

PSYC 8630

Dr. Vince D Calhoun

March 9, 2026

Summary

The study conducted by Zacks and colleagues (2011) investigated the relationship between prediction error, event boundaries, and brain activity. This study is important in the field of event cognition, especially in the branch of Event Segmentation Theory (EST; Zacks & Swallow, 2007). Studies about EST before this research mainly focus on behavioral evidence and computational simulation, demonstrating that people can either consciously or unconsciously segment continuous input into discrete units. EST emphasizes that the segmentations are driven by prediction errors. This study is one of the early studies that introduced the fMRI method to the event segmentation studies and laid a foundation for such method in future event segmentation studies. Researchers hypothesized that people generate more prediction error across events than within events, and the activity of the midbrain phasic dopamine system (MDS) is related to prediction errors. In 4 experiments, participants watch a series of video clips of daily activity (e.g., washing a car, building a LEGO model) and are asked to make predictions about what will happen in 5 seconds at multiple stop points. The stop points are set at 2.5 seconds before the local maximum and local minimum locations of the probability density function of an event segmentation task on the video clips (probabilities were estimated using Gaussian kernel density). The local maximum corresponds to the cross-event locations, and the local minimum corresponds to within-event locations. 4 experiments were conducted with similar procedures but with nuanced differences in how the prediction task was performed. In experiment 1, a frame of the 5 seconds later of the clip and a frame of a similar alternative clip were shown, and participants were asked to select one. Feedback on correctness was given after selection. In experiment 2, only one frame was presented, and participants were asked to judge whether the frame would appear in 5 seconds, and immediate feedback were given. In experiment 3, participants see two frames as in experiment 1 but also need to rate their confidence through a 6-point Likert-scale, no feedback was given after selection. Experiment 4 involves fMRI scanning. In experiment 4, participants were presented with 2 frames as in experiment 1, but no explicit feedback was given (the frames didn't offset immediately after the participant selected). There was a jitter (2s to 10s) between the offset of the frames and the restarting of the video clip.

In experiment 4, 24 participants were involved (12 female, ages 19-34). A 3T scanning with 64 ms TR and 25 ms TE was used. The researchers used GLM analysis with meta analysis approach. A set of region of interest analyses were conducted using manually selected regions of substantia nigra (SN) and ventral tegmental area (VTA) for each participant, and a whole brain analysis was conducted. Two types of analysis used the same design matrix. Motion correction, normalization, and noises where decently handled. The predictor of the design matrix is well defined, which is the local maximum and local minimum of the probability dense function of event segmentation and video restarting phase of these two conditions. The predictors of the video restarting phase were set using Finite Impulse Response (FIR) on the first 10 TR of the restarting phase. Researchers used participants' reaction time as the duration of the train of impulses, which assumes prediction as a continuous process. In addition, the researchers added a jitter (2s to 10s) between the offset of the two frame alternatives and the restarting of the video to separate the brain response to the task from the subsequent video watching phase.

The researchers found a strong connection between prediction error and event boundary such that people predict worse and are less confident in across-event conditions compared to within-event conditions. The activity of the right SN in the across-event condition is significantly stronger compared to within-event condition during the response period. There was a trend effect on right caudate with the same direction. In the restarting phase, the activity in the left caudate and left putamen was significantly greater in the across-event condition than within-event condition. The whole brain analysis demonstrates that the juncture of the parietal, temporal, and occipital lobes, and regions of the right anterior temporal cortex show significant differences between across-event conditions compared to the within-event condition.

Critique of Analysis

The findings of the study are valuable. Demonstrated areas that are related features that are related to cross-event conditions. It is feasible to consider whether these regions are related to event boundary detection and might be related to mechanisms such as state shifting, prediction error, or surprisal detection. However, there are several points that are worth consideration and can be improved based on contemporary techniques.

First, using the entire RT of the prediction task as the train of impulses in the response phase analysis may oversimplify the cognitive processes involved. Given that the judgment process of such a task involves predictive inference, memory retrieval, visual processing, error detection, decision-making, motor control, and other potential cognitive processes. Convoluting the whole response period with the same HRF may blend the BOLD signals from different cognitive processes and include signals from motor planning or other processes when interpreting. However, this might not be the case, as MDS is not in charge of these processes.

Second, the region of interest analysis deployed a manually selected region for each participant. While this helps handle individual differences, the process of manually selecting regions may still lead to human error and rater bias. Such processes can also involve errors from the manual process and reduce reproducibility (Poldrack, 2007).

Third, while the two-level GLM handles the subject residuals, it did not handle the residuals from items. With the small number of video clips and observation point, the residuals from the items may bias the model and lead to incorrect inferences.

Fourth, while using the local maximum and local minimum of the probabilistic density function of event segmentation is a clever and sensitive way of evaluating the relation between prediction error and brain activity. This led to a small size of observations (4 local maximums and 4 local minimums), which may reduce the statistical power of the GLM analysis.

Fifth, the authors acknowledged that when adding image similarity (the algorithm was not provided) as a control variable in their supplementary analyses for the restarting phase, the effects on the left caudate and left putamen are no longer significant. This suggests that the effects found in ROI analysis may be related to surficial-level dynamics and visual saliency but not semantic or episodic levels. Suggesting that the current method is relatively low in sensitivity at detecting abstract cognitive processes.

Sixth, the paper is grounded in EST, which focuses more on low level features such as prediction error as the primary cue of event boundary detection. However, there are other perspectives such as mental model constructions/updating from discourse psychology, which focus more on the coherence of the situation model/event model built for the observed information (Zwaan et al., 1995). This kind of more abstract mental processes may not be captured by the GLM approach and requires other types of analysis on dynamic processes.

One minor issue is the variance handling. The assumption of GLM is that the variance is i.i.d. and the study did not explicitly describe how the temporal autocorrelation was handled. It might be the case that such issue is automatically addressed by the tool FIDL, however the document of the tool is not accessible today.

Alternate Approaches

Regarding the first issue, perhaps applying FIR to the response phase analysis could help, which is relatively data driven compared to assuming the same HRF for the whole phase. However, using FIR instead of convolving HRF may reduce the statistical power, given that the sample size and observation points of the current data is relatively limited. A potential improvement is using ICA to classify different component maps that may be responsible for different processes, which could help decode more information from the given data. This idea can also apply to the whole brain analysis.

Second, to address possible human error or rater bias in manual region selection. One possible alternative approach is normalizing the participants' brain to a standard space (e.g., MNI space), and then using a high resolution anatomic template to select the regions of interest. This could increase the reliability and reproducibility of the definition of ROIs. Most of the modern fMRI analysis tool such as SPM has the probabilistic atlases that are useful for identifying the ROIs automatically. In addition, data driven methods like PCA or ICA could be considered. With ICA, data-driven component seeking might be more flexible and involve less artificial error.

Third, the traditional two level GLM approach only handles the random effects from subjects but ignores those from items. Using models like the cross-classified random effects model (CCREM) may help. CCREM can control random effects from multiple types of clusters. In cognitive studies, observations are treated as level 1 and subjects and items (response and restarting phases or video clips) are treated as level 2 clusters (Leroux & Beretvas, 2022).

Fourth, directly using the possibility of event segmentation as a predictor column might be able to detect a linear relationship between brain regions and prediction error. It is worth consideration.

Fifth, with a relatively small number of participants and sessions, either applying modern anatomy template on the normalized voxel data or using ICA to directly seek for component maps might be able to reduce the issue. Given the year of the study, the study may have used low level color features of the adherent images to calculate the image anatomy. To better control for the low level effects, calculating image similarity using embeddings of neural network based visual models might provide a finer control variable.

Sixth, to better understand how high-level cognitive processes affect the behavioral responses on event segmentation tasks or event boundary detection, analyses that focus on the dynamics of brain activity might be valuable. For instance, Dynamic Functional Network Connectivity (dFNC) or Windowed ICA (SW-ICA) (Iraji et al., 2020). Different abstract levels may be detectable with windows at different temporal scales. This approach requires using group ICA to extract component maps for functional networks such as DMN or FPN. Then using specified windows (the size of windows can be decided based on theoretical demands) to detect the correlation between the signals of different functional networks. Which will provide a second order dynamic on how the correlations between networks change during the response phase and the restarting phase. Such approaches can access the dynamics of each participant individually and will therefore provide higher sensitivity to high level cognitive processes.

Regarding the minor issue of autocorrelation handling, using an autocorrelation variance matrix may help.

References

Iraji, A., Miller, R., Adali, T., & Calhoun, V. D. (2020). Space: A Missing Piece of the Dynamic Puzzle. *Trends in cognitive sciences*, *24*(2), 135–149. https://doi.org/10.1016/j.tics.2019.12.004

Leroux, A. J., & Beretvas, S. N. (2022). Cross-Classified Random-Effects Models. In A. A. O'Connell, D. B. McCoach, & B. A. Bell (Eds.), *Multilevel Modeling Methods with Introductory and Advanced Applications* (p. 0). Emerald Publishing Limited. https://doi.org/10.1108/978-1-64802-873-120251008

Poldrack R. A. (2007). Region of interest analysis for fMRI. *Social cognitive and affective neuroscience*, *2*(1), 67–70. https://doi.org/10.1093/scan/nsm006

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction error associated with the perceptual segmentation of naturalistic events. *Journal of cognitive neuroscience*, *23*(12), 4057–4066. https://doi.org/10.1162/jocn_a_00078

Zacks, J. M., & Swallow, K. M. (2007). EVENT SEGMENTATION. Current directions in psychological science, 16(2), 80–84. https://doi.org/10.1111/j.1467-8721.2007.00480.x

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The Construction of Situation Models in Narrative Comprehension: An Event-Indexing Model. *Psychological Science*, *6*(5), 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x