Neural mechanisms underlying the abstract rule learning and its cross-modality generalization

**Abstract**

This study investigated the spatial temporal dynamics underlying multi-level abstract rule learning (local vs. global) and across multi-modality stimuli (auditory vs. tactile). Electroencephalography (EEG) signals of participants were recorded while they performed auditory or tactile local-global oddball paradigm. Local rule deviant led to mismatch negative response peaking at around 100-250ms relative to deviant onset time and localized at frontal-central brain area. In contrast, global rule deviant led to a late P300 response broadly ranging from 220ms to 900ms and reached maximum at central-parietal region. The high spatial correlations among topography of global effect across auditory and tactile modality indicated global rule learning can generalize among different modalities. However, local rule learning failed to generalize in this way. In total, these results provided systematical and comprehensive evidence for neural mechanisms underlying multi-level abstract rule learning and test their ability of generalization across auditory and tactile modality.

**Keywords:** abstract rule learning; local-global rule violation paradigm; auditory modality; tactile modality; cross-modality; MMN; P3b; pupillary dilation

**Introduction**

Abstract rule learning is characterized as the ability to extract the abstract rule from currently existing sensory sequences to predict the upcoming stimulus and the ability to generalize across different sensory modalities. Abstract rule learning plays a significant role in human cognition. However, neural mechanisms underlying abstract rule learning are still under debate. Thus, there are great need to systematically investigate neural activity correlates with abstract rule learning and its cross-modality ability.

The brain response to sensory violation rule comprises two distinct EEG components: an early mismatch negativity (MMN) and a late P300 (P3b). MMN can be elicited in an auditory oddball paradigm in which a sequence of identical standard stimuli establishes a regularity rule that is violated by a novel/deviant stimulus (Näätänen et al., 2010). MMN is usually found around 100-200ms relative to the onset of deviant stimulus, and has largest amplitude over frontal-central scalp area. Within the framework from predictive coding model, MMN is thought to reflect the prediction error signal generated by comparing bottom-up sensory information with top-down prediction based on priors over a short time scale. In contrast, P3b is considered to elicited by infrequently occurring auditory stimuli which are attended and task-relevant, broadly ranging between 300-800ms after stimuli onset, and dominant over central-parietal region. P3b is thought to be associated with updating and regulation of working memory (Polich, 2007).

Apart from auditory novelty, auditory omission is also a kind of rule deviant situation, however, brain response to auditory omission is still under debates and need to be systematically investigated. Also, it is important to note that brain responds to deviant stimuli are also studied in visual (Stefanics et al., 2011) , olfactory (Pause & Krauel, 2000) and somatosensory (Naejie et al., 2018) modality. However, the relationships among these rule deviant responses in different sensory modalities are less investigated. By studying the similarity among rule deviant responses in various sensory modalities, we can uncover the unified neural mechanism shared by abstract rule learning in difference modalities.

In this study, we use auditory or tactile odd-ball task with local-global paradigm (Bekinschtein et al., 2009) to disentangle multiple levels of violation rule, such as local-, global- and omission- violation rule. Through exposing participants to series of sensory sequence in auditory (experimrnt 1, 2) and tactile (experiment 3) modality to learn these abstract rules in different sensory modalities while EEG was recording, this study aims to investigate neural mechanisms underlying the various kinds of abstract rule learning and their cross-modality generalization abilities between auditory and tactile modality.

**Experiment 1**

**Materials and Methods**

**Participants**

Healthy adults with no neurological and psychiatric pathology were recruited in this study. All of them had normal hearing, normal or corrected-to-norm vision. The experiment protocol was approved by the Local Research Ethics Committee. All participants signed written informed consent prior to experiment and were paid for their participation. Eighteen participants (mean ± SD age, ± years; 12 females) took part in experiment 1.

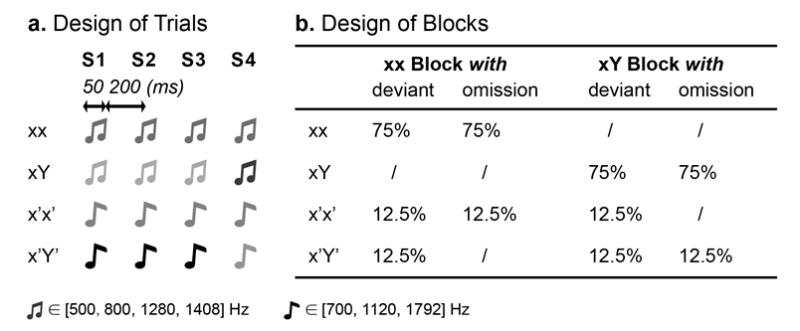
**Trial design and experimental Paradigm**

For experiment 1, each trial was either a 4-sound sequence which consisted of four tones of 50ms long with gaps of 200ms between each tone, or a 3-sound sequence which contained three 50-ms long tones with 200ms gaps. Intervals between trials were random selected between 1.5 second to 2 second. Trials with four identical tones were referred as local standard trials. Trials with three identical tones/tactile stimuli followed by a tone of different frequency were classified as local deviant trials. Trials only consisted of three identical tones were named omission trials.

All participants were seated approximately 100 cm in frontal of the computer monitor and instructed to remain still during the task. After clarification of the instructions of the task, the participants were monitored for the practice block of each experiment condition to ensure proper understanding of the task. Participants were required to fix their gaze on a cross in the central of the screen during the task. The sound was produced through speakers on the two sides of the screen at a comfortable loudness level.

For experiment 1, there were two kinds of tasks, XXXY task and XXXX task. Each task contained 4 blocks. Each block began with habituation period to establish global regularity rule, followed by violation period. There were 24 trials in habituation period and 96 trials in violation period. In XXXY task, the abstract rule of all trials in habituation period was XXXY. Meanwhile, 75% trials in violation period were the same as habituation trials, both in global abstract rule and tone frequency, named standard trial XXXY. The tone frequencies of these two kinds of trials were selected from 500, 800, 1280 and 2048 Hz. The abstract rule of the rest 12.5% trials was the same as that of habituation trials, but tone frequencies of these trials were selected from another frequency dataset (700, 1120 and 1792 Hz), these trials were called new frequency trial X’X’X’Y’. Another rest 12.5% trials differed from habituation trials both in abstract rule and tone frequency, named rule deviant trial X’X’X’X’, in which the fourth tone was identical as the first three tones; or named omission Y trial X’X’X’\_, in which the fourth tone was omission. These two kinds of violation happened in different blocks. XXXX task was similar with XXXY task, the differences were that habituation trial and standard trial were XXXX, new frequency trial was X’X’X’X’, rule deviant trial was X’X’X’Y’ in deviant block, and omission X trial was X’X’X’\_ in omission block. In other words, global standard and deviant trials were reversed between XXXY and XXXX task. Each of above four blocks repeated twice and presented in a random order, so there were totals of eight blocks in experiment 1.

In summary, there were two kinds of block in XXXX/XXXY task: deviant block and omission block. In XXXY task, global standard (frequent stimuli) trials were also local deviant. In contrast, global standard trials were local standard in XXXX task. This orthogonal design enabled two level of regularities, standard versus deviant and local versus global.



**Figure 1. Illustrations of trial design and block design for experiment 1.** The upper panel (A) illustrates each trial is either a 4-sound sequence which consists of four tones of 50ms long with gaps of 200ms between each tone, or a 3-sound sequence which contains three 50-ms long tones with 200ms long gaps. Intervals between trials are random selected between 1.5 second to 2 second. The down panel (B) shows the experiment involves both XXXX and XXXY task. XXXX task with deviant block is consisted of 75% standard trials (XXXX), 12.5% new frequency trials (X’X’X’X’ and 12.5% rule deviant trials (X’X’X’Y’); XXXX task with omission block is consisted of 75% standard trials (XXXX), 12.5% new frequency trials (X’X’X’X’ and 12.5% omission X trials (X’X’X’). In contrast, XXXY task with deviant block involves 75% standard trials (XXXY), 12.5% new frequency trials (X’X’X’Y’ and 12.5% rule deviant trials (X’X’X’X’); XXXX task with omission block involves 75% standard trials (XXXY), 12.5% new frequency trials (X’X’X’Y’) and 12.5% omission Y trials (X’X’X’). The orthogonal design of these experiments enable two level of regularities, standard versus deviant and local versus global.

**EEG recordings**

EEG signals were recorded with 65 active Ag/AgCl electrodes placed according to the international 10-20 standard system (antiCAP and BrainAmp, Brain Products, Germany), in which FCz and AFz were used as reference and ground electrode, respectively. Two additional active Ag/AgCl electrodes were placed below and extender of the right eyes to record horizontal electrooculograms (HEOG) and vertical electrooculograms (VEOG). The impedance between electrode and scalp interface was kept below 5 kΩ. EEG signals were band-pass filtered between 0.1 and 250 Hz. The sampling rate was set at 1000 Hz.

**EEG data pre-processing**

The EEG signals were preprocessed using Fieldtrip toolbox with MATLAB. EEG signals were band-pass filtered between 0.1- 40 Hz and resampled at 200 Hz. EEG signals contain muscle or movement artifacts were excluded from further analysis. Independent components analysis (ICA) as implemented in runica algorithm was used to remove components corresponding EOG artifacts. Channels that were removed by artifact rejection were interpolated from the remaining neighbor channels, and the EEG signals were subsequently re-referenced to the common average of all channels.

As regard to auditory experiment, EEG signals were segmented from 950ms before to 1500ms after the onset of the fourth sound in deviant block, EEG signals were segmented from 700ms before to 1750ms after the onset of the third sound in omission block. First 200ms of the epoch was used for baseline correction. As for tactile experiment, EEG signals were segmented from 1280ms before to 1180ms after the onset of the fourth sound in deviant block, EEG signals were segmented from 920ms before to 1540ms after the onset of the third sound in omission block. First 200ms of the epoch was used for baseline correction.

**ERP analysis**

We defined local deviant trials as X’X’X’Y’ in both XXXY task and XXXX task. Local standard trials referred to X’X’X’X’ in both XXXY task and XXXX task. Global deviant trials equaled to X’X’X’X’ in XXXY task and X’X’X’Y’ in XXXX task. Global standard trials mean X’X’X’X’ in XXXX task and X’X’X’Y’ in XXXY task.

Local effect was defined as the difference between local deviant trials versus local standard trails, suggesting a violation of rule within trial; Global effect was calculated by the difference between global deviant trials versus global standard trials, denoting violation to regulation rule across trials; Omission Y effect was characterized by the difference between omission Y trials X’X’X’\_ in XXXY task with X’X’X’Y’ in XXXY task, while omission X effect was calculated by the difference between omission X trials X’X’X’\_ in XXXX task with X’X’X’X’ in XXXX task.

ERP analysis was performed by averaging EEG segments of the same condition. For local condition comparison, pair t-tests were performed at each sensor and each time sample, no correction for multiple comparisons was applied and the significance level was set at 0.005. For all other comparisons, global standard versus global deviant and non-omission versus omission, correction for multiple comparisons was applied using cluster-based permutation test, and significance level was set at 0.05.The cluster-based permutation test used Monte Carlo method implemented in Fieldtrip toolbox for estimating significant spatial temporal clusters against surrogate null distribution. The null distribution was obtained by permutating the condition labels across trials by number of times, in this study 1000 times. The region of interests (ROIs) were the electrodes where the paired t-test of each condition comparison between ERPs are the maximal in brain middle line.

Topographic analyses were also performed to ground-averaged ERPs to identify the scalp distribution of each condition comparison within certain time window. The spatial correlation analysis among these topographic patterns was further conducted to identify the spatial similarity of ERP patterns among local-, global- and omission effect across auditory and tactile modality. High spatial correlation coefficient indicated high similarity.

**Decoding analysis**

We examined whether multi-channel patterns of ERPs can be used to reveal neural representation of local-, global- or omission effect. Single-trial decoding method (Grootswagers et al., 2017; Bae & Luck, 2018) was applied. The classifier was based on linear support vector machine (SVM) with parameter of BoxConstraint seted to 1 using MATLAB fitcsvm() function, which is included in MATLAB Statistics and Machine Learning Toolbox. The SVM classifier was trained using the time domain feature vector from the extracted epochs of all channels. As for single-trial decoding, a three-folds cross validation was adopted at each time samples. Specially, trials were divided into three subgroups. Two of them were used to train the classifier, and then the performance of the classifier was assessed with remaining one subgroup. This procedure was repeated three times until all the subgroup served as the testing dataset. The cross-validation procedure described above was iterated 10 times, each time with a new random assignment of trials into three subgroups. This iteration could help to minimize idiosyncrasies associated with trial assignments, and thus yield a more stable result. The chance level was 50%. We tested whether the group-level decoding accuracy at each time sample was above chance level by performing cluster-based permutation test.

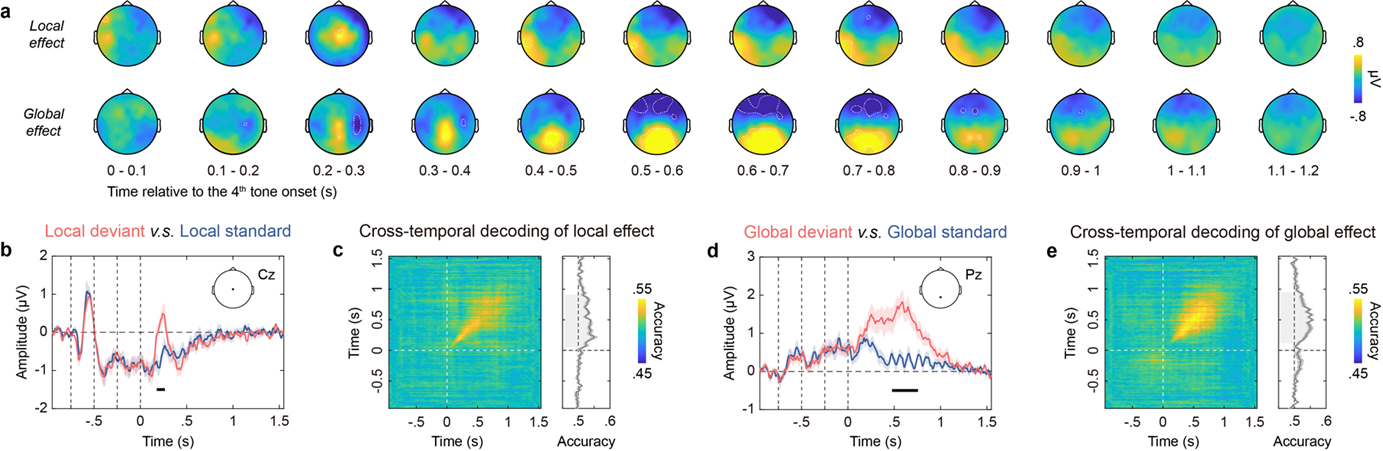
In addition, we investigated how a pattern classifier trained at time t1 generalizes to data from another time point t2. This approach results in a temporal generalization matrix that contains considerable information about the dynamics of neural codes across all time point combinations. If a given neural code is active at times t1 and t2, then a classifier trained at time t1 should be able to decode information, or generalize to t2. In other words, the shape of the temporal generalization matrix can provide valuable information about whether brain activity patterns at one time point are sustained at another time point.

**Results**

**ERP results: Local and global effect**

As for experiment 1, upper left picture showed mean amplitude and standard error of ERPs at Cz electrode evoked by local standard and local deviant condition. Upper middle picture revealed that significant difference of ERP between local standard and local deviant condition was found in time window ranging from 190-250ms after the onset of deviant tone, indicating that the early presence of MMN was an index for local rule violation. For local effect, the scalp topographic maps of t value in upper right picture confirmed that MMN component during 190-250ms dominant in central area was related to local rule violation.

Bottom left picture showed group averaged time course of ERPs and their standard error at Pz electrode for global standard and global deviant condition in experiment 1. The cluster-based permutation test in bottom middle picture showed significant difference from 220ms to 900ms. For global effect, the scalp topographic maps of statistical difference in bottom right picture showed significant P3b component maximal around parietal cortex ranged from 220ms to 900ms.

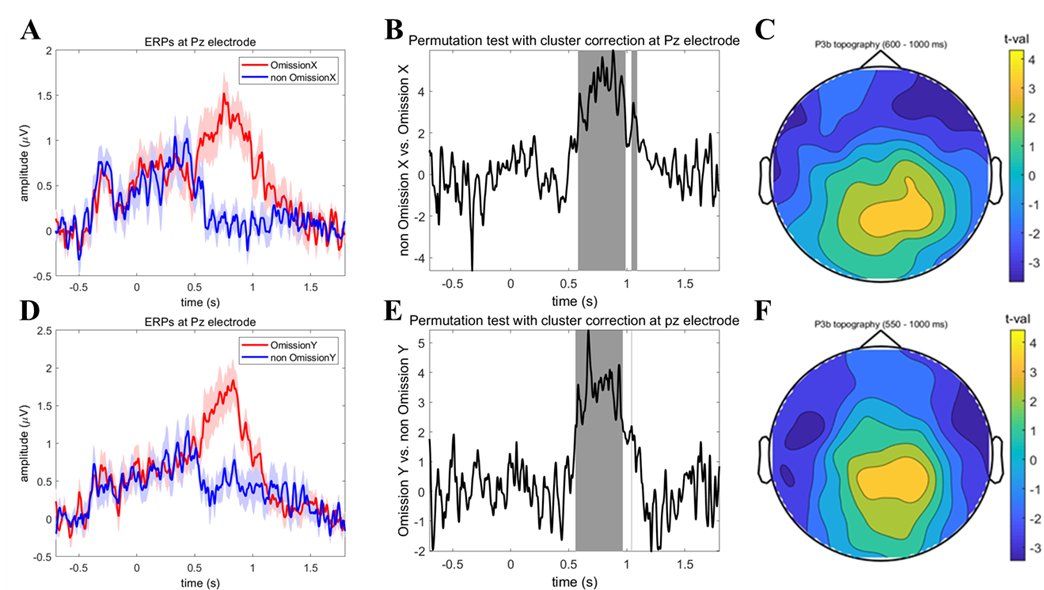


**Figure 2. Grand-averaged ERPs elicited by local and global comparison in experiment 1.** (A): Local deviant trials elicit mismatch negativity (MMN) response compared local standard trials at Cz electrode. Global deviant trials elicit P3b response compared global standard trials at Pz electrode. (B) indicates statistically significant time window (190-250ms) for this MMN response. The scalp topographic maps of t value upper right panel (C) confirms that MMN component during 190-250ms dominant in central area is related to local rule violation. The bottom left panel (D): Global deviant trials elicit P3b response compared global standard trials at Pz electrode. Gray-shaded area in the bottom middle panel (E) indicates statistically significant time window (220-900ms) for this P3b response. The scalp topographic maps of statistical difference in down right panel (F) shows significant P3b component related to global rule violation maximal around parietal cortex ranged from 220ms to 900ms.

**ERP results: Omission X and omission Y effect**

As for experiment 1, upper left picture showed the grand averaged ERPs and their standard error at Pz electrode for omission X condition and non-omission X condition. Statistical analysis in upper middle picture revealed that significant difference between these two conditions ranged from 600ms to 1000ms. For omission X effect, the scalp topographic maps of statistical difference in upper right picture showed significant P3b component maximal around parietal cortex ranged from 600ms to 1000ms.

Bottom left picture showed group averaged time course of ERPs and their standard error at Pz electrode for omission Y condition and non-omission Y condition in experiment 1. The cluster-based permutation test in bottom middle picture showed significant difference from 550ms to 1000ms. For omission Y effect, the scalp topographic maps of statistical difference in bottom right picture showed significant P3b component maximal around parietal cortex ranged from 550ms to 100ms.

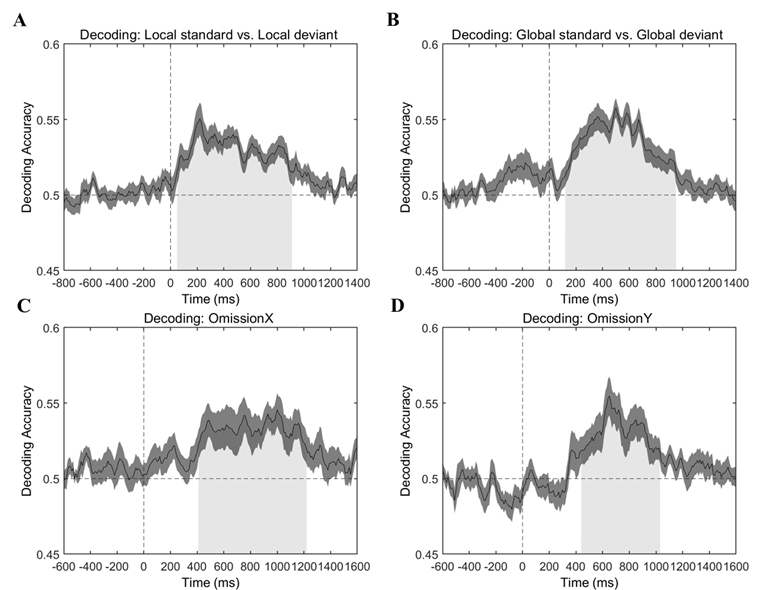


**Figure 3. Grand-averaged ERPs elicited by two kinds of omission comparisons in experiment 1.** The upper left panel (A): Omission X trials elicit P3b response compared non-omission X trials at Pz electrode. Gray-shaded area in the upper middle panel (B) indicates statistically significant time window (600-1000ms) for this P3b response. The scalp topographic maps of t value upper right panel (C) confirms that P3b component during 600-1000ms dominant in parietal area. The bottom left panel (D): Omission Y trials elicit P3b response compared non-omission Y trials at Pz electrode. Gray-shaded area in the bottom middle panel (E) indicates statistically significant time window (550-1000ms) for this P3b response. The scalp topographic maps of statistical difference in down right panel (F) shows significant P3b component maximal around parietal cortex ranged from 550ms to 1000ms.

**Decoding results**

As for decoding, we should know two points. The first is that, the classifier’s performance is significantly above chance, this indicates that EEG activity patterns contain class specific information, and we can conclude that the class can be decoded from the EEG data; the second is that, there are two key index for time-resolving decoding, one is the onset time of significant decoding performance; the another is the time when the peak in decoding performance occur.

The pictures were about the results of single-trial decoding in experiment 1. As we can see from picture, the onset time of significant decoding performance was earlier in local comparison compared to that of global comparison, indicated that local effect information was earlier available than global effect, which was consistent with ERP results, what’s more, the time when the peak in decoding performance occur also showed local effect was precede global effect. The interest observation was that the onset time of significant decoding performance in Omission X and Omission Y comparison were all later than that of local effect, but similar with global effect, which suggested that brain response to auditory omission may be not just predictive error, because they shared some common characteristics with global effect.



**Figure 4. Single-trial decoding results for experiment 1.** The onset time of significant decoding performance was earlier in local comparison (A) compared to that of global comparison (B), indicated that local effect information was earlier available than global effect. Furthermore, the time when the peak in decoding performance occur also showed local effect was precede global effect. The onset time of significant decoding performance in Omission X comparison (C) and Omission Y comparison (D) were all later than that of local effect, but similar with global effect, which suggested that brain response to auditory omission may be not just predictive error, because they shared some common characteristics with global effect.

**Experiment 2**

**Materials and Methods**

**Participants**

Healthy adults with no neurological and psychiatric pathology were recruited in this study. All of them had normal hearing, normal or corrected-to-norm vision. The experiment protocol was approved by the Local Research Ethics Committee. All participants signed written informed consent prior to experiment and were paid for their participation. Sixteen participants (mean ± SD age, ± years; 11 females) joined in experiment 2.

**Trial design and experimental Paradigm**

Experiment 2 also used the auditory local-global oddball paradigm which was similar with experiment 1. However, two kinds of additional pure omission block were presented following deviant block in XXXX task and XXXY task respectively. All the trials in the pure omission block were comprised of 3-sound sequence, in other words, the fourth tone of trials in pure omission block was missed. In addition, the total of 8 tone frequencies was divided into two subgroups, the one was high-frequency sound group (800, 1100, 1500 and 1900 Hz), the another was low-frequency sound group (400, 500, 700 and 1000 Hz). Participants were then randomly assigned to two subgroups, the one received training in high frequency sound and testing in low frequency sound, and the another vice versa. This design was to rule out the impact of tone frequency on abstract rule learning.

Additionally, eye movements were monitored and pupil size was recorded by Desktop-Mount eyetracker (Eyelink 1000, SR Research) during the task. Participants’ right eye was tracking with the sampling rate of 500 Hz. Calibration and validation of eye fixation were conducted before each block started.

**Pupil preprocessing**

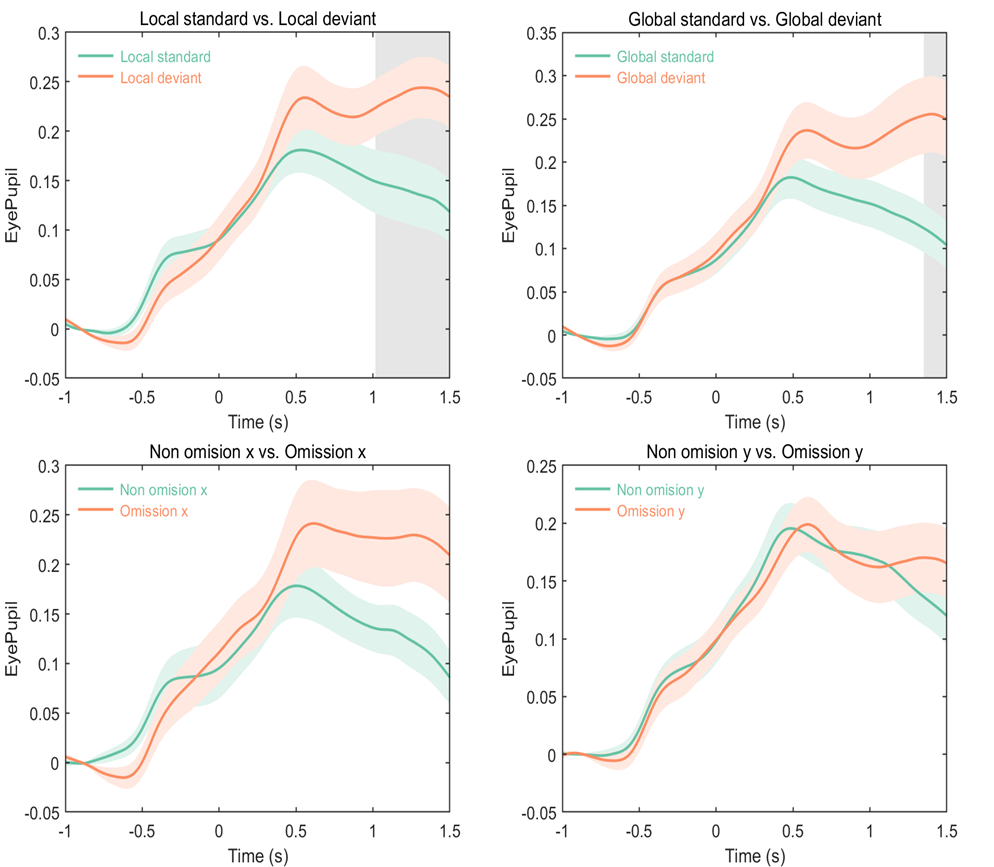
Eye blinks were identified based on EyeLink defined standard criteria. These blink events were linearly interpolated from 150ms before the starting point until 150ms after the endpoint of the blink. After interpolation, pupil recordings were band-pass filtered between 0.02 and 4 Hz. Then, data were demeaned per condition and resampled to 100 Hz. Lastly, data were segmented like ERP analysis and plotted event-related pupil responses (Urai et al., 2017).

It was suggested that the size of global response (the amplitude of P3b) is influenced by attention while local response (the amplitude of MMN) seems stable no matter which level of attention engagement (Näätänen et al., 2001), and pupil size change is an index marker for attention (Van der Wel & van Steenbergen, 2018). Thus, the cross-correlation relationship between ERP difference with pupil size change was also calculated.

**Results**

**Pupil size results**

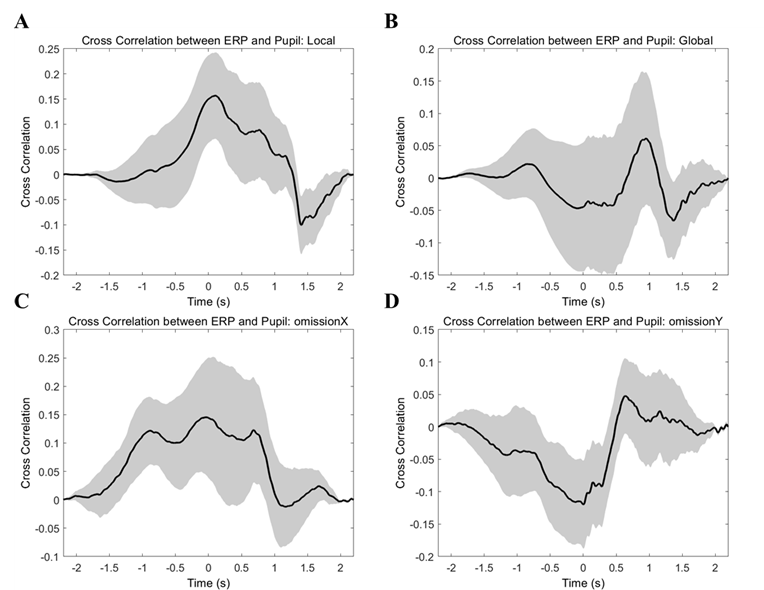
Significant difference of pupil size was observed between Local deviant versus Local standard, and Global deviant versus Global standard. The modulation of local effect on pupil size seemed relatively earlier than global effect. However, we observed different pupil response for omission x effect and omission y effect. As for omission X, it showed obvious difference compared to no omission x trials. But there was no significant difference between Omission y trials and No omission y trials, suggesting that these two kinds of omission response may reflect different process.



**Figure 6. Pupil dilation as a marker for four kinds of deviant detections in experiment 2.** Significant difference of pupil size was observed between Local deviant versus Local standard (A) and Global deviant versus Global standard (B). The modulation of local effect on pupil size seemed relatively earlier than global effect. However, we observed different pupil response for omission x effect (C) and omission y effect (D). As for omission X, it showed obvious difference compared to no omission x trials. But there was no significant difference between Omission y trials and No omission y trials, suggesting that these two kinds of omission response may reflect different process.

**Results of cross Correlation between ERPs and Pupil size**

The cross correlation was calculated between ERPs and pupil size in four condition comparisons corresponding to local-, global-, omission X- and omission Y rule violation effect.



**Figure 7. Cross Correlation between ERP difference and Pupil size difference in four condition comparisons in experiment 2.**

**Experiment 3**

**Materials and Methods**

**Participants**

Healthy adults with no neurological and psychiatric pathology were recruited in this study. All of them had normal hearing, normal or corrected-to-norm vision. The experiment protocol was approved by the Local Research Ethics Committee. All participants signed written informed consent prior to experiment and were paid for their participation. Twenty participants (mean ± SD age, 23.2 ± 2.2 years; 10 females) took part in experiment 3.

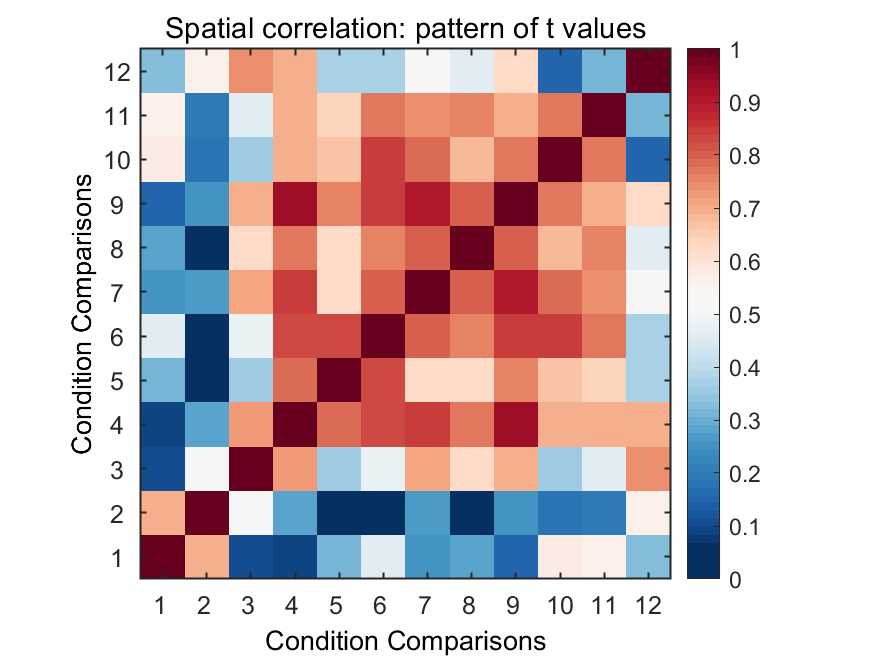
**Trial design and experimental Paradigm**

For experiment 3, each trial was composed of either four 150-ms tactile stimuli and inter-stimulus interval (ISI) was 210ms, or three 150-ms tactile stimuli with ISI of 210ms. Inter trial intervals were also obtained between 1.5 second to 2 second randomly. Meanwhile, the frequency of the fourth tactile stimulus was either same or different with that of preceding three tactile stimuli.

Experiment 3 was the same as experiment 1 apart from changing auditory stimulus into tactile stimulus. The tactile frequencies of trials in habituation phase were selected from 20, 80, 180 and 300 Hz. Meanwhile, the tactile frequencies of trials in violation phase were selected from 40, 140 and 280 Hz.

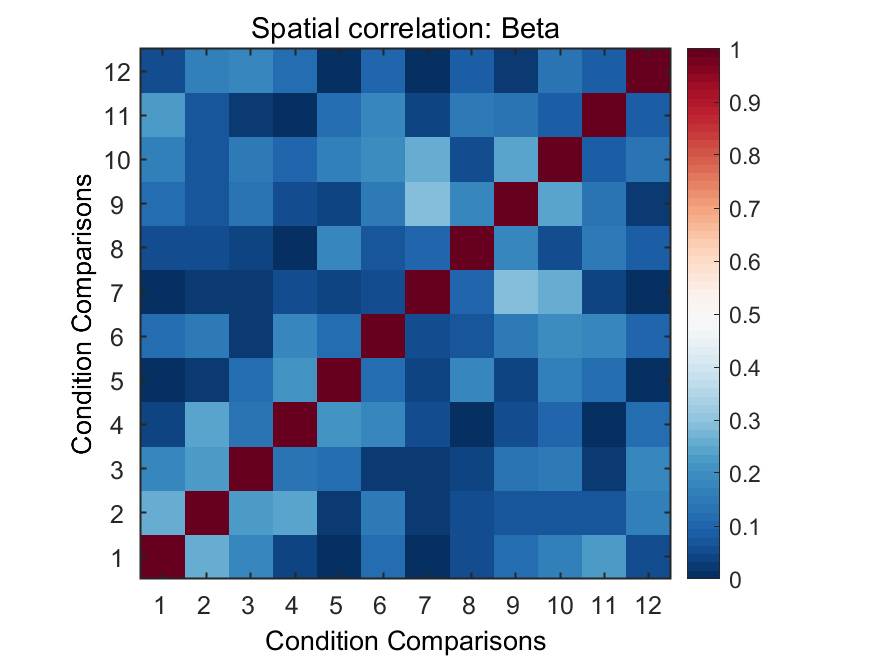
**Results of cross-modal generalization of Patterns of ERPs**

The high spatial coefficients among topography of global rule violation across auditory and tactile modality indicated global rule learning can generalize among different modalities. However, local rule learning failed to generalize in this way.



**Figure 8. Spatial correlation matrix of t values for four condition comparisons across three experiments.** Conditions (1-3) correspond to local comparison in experiment 1, 2 and 3 respectively; Conditions (4-6) correspond to global comparison in experiment 1, 2 and 3 respectively; Conditions (7-9) correspond to omission X comparison in experiment 1, 2 and 3 respectively; Conditions (10-12) correspond to omission Y comparison in experiment 1, 2 and 3 respectively. High correlations of patterns of t values are found within and between global and omission X condition comparison across three experiments. However, high correlations of patterns of t values are only found within omission Y condition comparison across two auditory experiments.

**Cross task decoding results**



**Figure 9. Spatial correlation matrix of beta values for four condition decoding across three experiments.**

**Discussion**

Arguably, there are two subcomponents within MMN: an early negative component around frontal electrodes (ie. Fz), and a subsequent positive component appearing at central electrodes (ie. Cz). In this study, we observe only positive subcomponent in both auditory and tactile local violation rule condition. Electrodes for most significant MMN component differ between auditory and tactile modality, central electrodes in auditory experiment and frontal electrodes in tactile modality.

The interesting observation is that omission Y effect in XXXY task is larger than omission X effect in XXXX task in auditory experiment, suggesting that omission response is larger when there are both violation to the expectation of deviance and violation to the expectation that a tone would be presented. In some previous studies, the omission response is similar with tradition MMN and occurs at approximately the same time as the MMN (Chennu et al., 2016). However, within the framework of predictive coding model, the omission response reflects pure top-down process including top-down prediction and subsequent bottom-up prediction error, because there is no incoming information from bottom-up processing in the absence of sensory input (Bendixen et al., 2013). In our current study, the omission response is more like p3b component and differs between different global standard contents, suggesting that omission response is influenced by kinds of global regularity and reflects a kind of global abstract rule learning in this experiment.

**References**

Bae, G. Y., & Luck, S. J. (2018). Dissociable decoding of spatial attention and working memory from EEG oscillations and sustained potentials. *Journal of Neuroscience*, *38*(2), 409-422.

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, *106*(5), 1672-1677.

Chennu, S., Noreika, V., Gueorguiev, D., Shtyrov, Y., Bekinschtein, T. A., & Henson, R. (2016). Silent expectations: dynamic causal modeling of cortical prediction and attention to sounds that weren't. *Journal of Neuroscience*, *36*(32), 8305-8316.

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time series neuroimaging data. *Journal of cognitive neuroscience*, *29*(4), 677-697.

SanMiguel, I., Widmann, A., Bendixen, A., Trujillo-Barreto, N., & Schröger, E. (2013). Hearing silences: human auditory processing relies on preactivation of sound-specific brain activity patterns. *Journal of Neuroscience*, *33*(20), 8633-8639.

Näätänen, R., Astikainen, P., Ruusuvirta, T., & Huotilainen, M. (2010). Automatic auditory intelligence: An expression of the sensory–cognitive core of cognitive processes. *Brain research reviews*, *64*(1), 123-136.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). ‘Primitive intelligence’in the auditory cortex. *Trends in neurosciences*, *24*(5), 283-288.

Naeije, G., Vaulet, T., Wens, V., Marty, B., Goldman, S., & De Tiege, X. (2018). Neural basis of early somatosensory change detection: a magnetoencephalography study. *Brain topography*, *31*(2), 242-256.

Pause, B. M., & Krauel, K. (2000). Chemosensory event-related potentials (CSERP) as a key to the psychology of odors. *International Journal of Psychophysiology*, *36*(2), 105-122.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology*, *118*(10), 2128-2148.

Stefanics, G., Kimura, M., & Czigler, I. (2011). Visual mismatch negativity reveals automatic detection of sequential regularity violation. *Frontiers in Human Neuroscience*, *5*, 46.

Urai, A. E., Braun, A., & Donner, T. H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nature communications*, *8*(1), 1-11.

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic bulletin & review*, *25*(6), 2005-2015.

**Supporting Information**