During goal-directed behavior, humans purportedly form and retrieve so-called event files, conjunctive representations that link context-specific information about stimuli, their associated actions, and the expected action outcomes. The automatic formation, and later retrieval, of such conjunctive representations can substantially facilitate efficient action selection. However, recent behavioral work suggests that these event files may also adversely affect future behavior, especially when action requirements have changed between successive instances of the same task context (e.g., during task switching). Here, we directly tested this hypothesis with a recently developed method for measuring the strength of the neural representations of context-specific stimulus-action conjunctions (i.e., event files). Thirty-five male and female adult humans performed a task switching paradigm while undergoing EEG recordings. Replicating previous behavioral work, we found that changes in action requirements between two spaced repetitions of the same task incurred a significant reaction time cost. By combining multivariate pattern analysis and representational similarity analysis of the EEG recordings with linear mixed-effects modeling of trial-to-trial behavior, we then found that the magnitude of this behavioral cost was directly proportional to the strength of the conjunctive representation formed during the most recent previous exposure to the same task, that is, the most recent event file. This confirms that the formation of conjunctive representations of specific task contexts, stimuli, and actions in the brain can indeed adversely affect future behavior. Moreover, these findings demonstrate the potential of neural decoding of complex task set representations toward the prediction of behavior beyond the current trial.

Key words: cognitive control; electroencephalography; repetition cost; representational similarity analysis; task switching

Significance Statement

Understanding how the human brain organizes individual components of complex tasks is paramount for understanding higher-order cognition. During complex tasks, the brain forms conjunctive representations that link individual task features (contexts, stimuli, actions), which aids future performance of the same task. However, this can have adverse effects when the required sequence of actions within a task changes. We decoded conjunctive representations from electroencephalographic recordings during a task that included frequent changes to the rules determining the response. Indeed, stronger initial conjunctive representations predicted significant future response-time costs when task contexts repeated with changed response requirements. Showing that the formation of conjunctive task representations can have negative future effects generates novel insights into complex behavior and cognition, including task switching, planning, and problem solving.

Introduction

According to influential theories of human cognition (Hommel et al., 2001; Frings et al., 2020), humans automatically store and retrieve episodic representations of their interactions with the world to guide future behavior. These representations contain context-specific combinations of stimuli, actions, and outcomes, and are also known as event files (Hommel et al., 2001; Hommel, 2009). For instance, in the example of opening a safe, an event file would contain a conjunction of the specific context (e.g., the room with the safe), a stimulus (the combination lock), an action (entering the combination), and an outcome (the opening of the safe). The formation and retrieval of such conjunctive event files...
can greatly facilitate behavior by making repeated actions more efficient.

However, it is also observed that these event files can have adverse effects on behavior (Hommel, 2004; Hommel and Colzato, 2004; Frings et al., 2020). For example, when the correct response or its outcome contingencies change between successive exposures to the same context (e.g., if the combination of the safe has been recently changed), preexisting event files may contain outdated representations and prime an inappropriate action (Neill, 1997; Waszak et al., 2003; Hommel et al., 2014). In the laboratory, these effects can be measured in the N-2 Repetition Cost (N2RC) paradigm (Mayr and Keele, 2000; Mayr, 2002; Grange et al., 2013; Kowalczyk and Grange, 2016). In such experiments, triplets of successive trials made in separate task contexts (C-B-A) are compared with triplets of trials that include an N-2 repetition of task A (A-B-A). The N2RC effect is a reaction time cost on the third trial of an ABA compared with a CBA triplet when the response on third trial differs from the response on the first trial (Fig. 1).

However, there is ongoing debate over the source of the N2RC and whether it indeed relates to the formation of event files (Grange et al., 2017; Grange, 2018; Kessler, 2018; Scheil and Kleinsorge, 2019, 2021; Kowalczyk and Grange, 2020). For example, an influential alternative account attributes N2RC effects to the inhibition of task set A by the intervening switch to task B. This may impair the return to the previous, now-inhibited task context (Mayr, 2002; Philipp and Koch, 2006; Koch et al., 2010; Schuch and Grange, 2015; Sexton and Cooper, 2017).

In the current study, we address this question from a neuroscientific perspective, using a recently developed method for the quantification of the type of conjunctive representation predicted by the event-file framework (Kikumoto and Mayr, 2020). Single-neuron recordings and fMRIs have already suggested that such conjunctive representations may exist in both nonhuman (Rigotti et al., 2013; Parthasarathy et al., 2017) and human brains (Kühn et al., 2011). Kikumoto and Mayr’s (2020) method uses multivariate pattern analysis (MVPA) and representational similarity analysis (RSA) to noninvasively track the emergence of such representations with millisecond precision from human EEG recordings. This method also allows for a direct test of the relationship between the strength of these representations and behavior via linear mixed-effects modeling of trial-to-trial reaction time. Kikumoto and Mayr (2020) demonstrated that conjunctive representations decoded from EEG recordings can be used to predict reaction time on the current trial; stronger conjunctions between task context, stimulus information, and response activity accompany faster reaction times.

Here, we leverage this technique to investigate whether conjunctive representations can also account for N2RC effects, that is, whether the formation of a conjunctive event file on one trial has adverse effects on future trials, as predicted by the event-file account of the N2RC effect. Using an N2RC paradigm, we first aimed to replicate Kikumoto and Mayr’s (2020) finding that stronger conjunctions of task context, stimulus, and response accompany faster reaction times on the same trial. We then hypothesized that within ABA triplets, the strength of the conjunctive representation formed on the first instance of task A would be detrimental to reaction time on the second instance of task A if a different response was required on the second instance. This would provide direct evidence for the N2RC effect being partially attributable to previously formed, yet now outdated, event files. More broadly, it would indicate that although event-file formation is typically beneficial to immediate behavior, it can also be detrimental when the behavioral requirements associated with a specific context subsequently change.

Materials and Methods

Participants. Thirty-five healthy young adults participated in the experiment [age mean (SD), 19.6 (3.0), 1 left handed, 19 females]. This sample size was based on previous work using a similar behavioral paradigm (Grange et al., 2017, their Appendix B). Participants were either paid $15 per hour or received course credit for their participation in the study. All participants had normal or corrected-to-normal vision. The experiment was approved by the ethics committee at the University of Iowa (Institutional Review Board #201511709).

Materials and procedure. Experimental stimuli were presented via an Ubuntu Linux computer, running Psychtoolbox-3 software (Brainard 1997) under MATLAB 2015b (MathWorks). Participants sat upright with their arms and the computer keyboard resting on a supportive platform placed on the arm rests of the chair.

Experimental paradigm. Our paradigm was adapted from Grange et al., (2017; experiment 2; Fig. 2). A large hollow square outline (the frame) remained on the screen across trials. Trials began with the presentation of the task cue (the outline of one of three shapes) inside the center of the frame, followed by a 500 ms cue-target interval. The target (filled black dot) would then appear in one of the four corners of the frame.

Based on the combination of cue shape and target location, participants had 2000 ms to respond using their right index finger. The cue shapes Hexagon, Triangle, and Square instructed one of three different response mappings (tasks), which corresponded with Vertical, Horizontal, and Diagonal movement of the target to one of the unoccupied corners of the frame. For example, should the target appear in the top-left corner of the frame, a triangle cue would imply that the target should move horizontally to the top right, hence cuing the response associated with that location. Participants responded using their right index finger and the 4, 5, 1, and 2 keys on the keyboard number pad, which corresponded to the top-left, top-right, bottom-left, and bottom-right corners of the frame, respectfully. Participants were asked to return their index finger to the center of the response keys after each trial.

Responses were followed by a 200 ms blank period where only the frame remained on screen. If participants failed to respond within 2000 ms, or made an incorrect response, the usual blank period was replaced by a 500 ms presentation of the words Faster or Error in red text. The experiment consisted of 800 trials (eight blocks). Cue identity and target location were pseudorandomized to prevent immediate cue (and task) repetitions and an equal number to task conditions (50% ABA triplets, 50% CBA triplets).

The first two trials of each block were removed from the analysis, as they could not be considered part of a triplet. The behavioral data were then screened for response errors/omissions, RTs lower than 150 ms (three trials overall), and RTs outside 2.5 SDs of the mean (2.33% overall). These trials were removed from further analysis, along with the two subsequent trials. There were three possible task triplet conditions of interest (ABA-S, ABA-R, CBA) with the -S and -R denoting either a switch or repeat in the required response between trials N-2 and N of an ABA triplet. (In principle, CBA-S and CBA-R also exist, but as there is no repetition of task cues within these triplets, the overlap of event files should be zero, regardless of response requirement.) Our main comparison of interest was between ABA-S trials and CBA trials. ABA-R trials were of secondary interest as there was only a limited amount of those triplets per subject (as target locations were drawn from a uniform distribution to not introduce a bias, resulting in only one ABA-R triplet for every three ABA-S triplets). The dependent variables of interest were RT and accuracy on the final trial of ABA-S and CBA triplets, which were compared using paired-samples t tests.

EEG recording. EEG data were recorded using a 64-channel active electrode cap connected to an actiChamp amplifier (Brain Products) at a rate of 500 Hz (10 s time-constant high-pass and 1000 Hz low-pass
hardware filters). Electrodes Pz and Fz served as the reference and ground, respectively.

EEG preprocessing. Raw EEG data were preprocessed using custom MATLAB scripts and EEGLAB (RRID:SCR_007292) toolbox functions (Delorme and Makeig, 2004). Individual participant data were imported, resampled to 250 Hz, and filtered using Hamming windowed sinc FIR filters (pop_eegfiltnew.m; high-pass, 0.01 Hz; low-pass, 50 Hz). The resulting data were then epoched into two datasets, one target locked [-600:800 ms] and the other response locked [-700:200 ms], which were used to investigate the emergence of task-feature representations relative to either event. Epochs containing nonstereotypical artifacts were then removed from further analysis (joint probability and joint kurtosis; 5.5 SD cutoffs; compare Delorme et al., 2007). Data were referenced to a common average before independent component analysis decomposition (Infomax; Bell and Sejnowski, 1995) with extension to sub-Gaussian sources (Lee et al., 1999). Components representing eye movement and electrode artifacts were identified using outlier statistics and removed from the data. Unlike Kikumoto and Mayr (2020), EEG data did not undergo time-frequency analysis, and the following analyses were performed on the raw EEG channel data (i.e., the event-related potential).

Linear discriminant analysis. Linear discriminant analysis (LDA) was run on one subject at a time. Each subject had 12 randomly selected trials held out as test data, whereas their remaining trial data were used for training. Training trial data were then sorted into 12 classes, one for each unique combination of cue, target, and response (i.e., the event file or constellation in Kikumoto and Mayr, 2020). In the case of unequal numbers of trials across these classes, trials were randomly excluded to achieve equal numbers. The excluded trials were randomly paired with the remaining trials in their respective classes and averaged. This 1–1 pairing and averaging allowed for all data to be included in each training and testing loop while also Figure 1. The event-file hypothesis of N-2 repetition costs. A, Diagram of how individual task features are bound as event files and then retrieved at a later time. B, Generic task diagram for ABA and CBA trial triplets. Binding of the event file in trial N-2 occurs for both tasks A and C, but retrieval of N-2 event file only occurs in trial N of an ABA triplet, not a CBA triplet.

Figure 2. Experimental paradigm. A, Diagram for a single trial. The shape cue is presented for 500 ms before the target (black dot) onset. The target will be displayed for 2000 ms or until a response is made. Subjects responded using the number pad on a full keyboard. Response is immediately followed by a 200 ms blank inter-trial interval before the next cue. B, The three possible shape cues and their meanings for the task. C, Examples of ABA-S and CBA trial triplets.
ensuring an LDA with balanced classes. LDA was conducted using the MATLAB fitdiscr function. Each training-testing loop applied the LDA to each time point, with all 63 EEG channels included as features. Posterior probabilities for the hold-out data (12 trials) were saved for each time point. This process was repeated with a new set of 12 hold-out trials until each trial had been tested. This entire process was then repeated for 10 iterations, resulting in each trial being included in the testing set 10 times.

RSA. The posterior probabilities for all trials and iterations were smoothed using an overlapping moving median window of 20 ms. As in Kikumoto and Mayr (2020), the posterior probabilities of each trial for each time point were regressed onto model vectors representing the idealized classification probabilities for the true class identity of the trial (Fig. 3). Specifically, four vectors were used as predictors in the regression, representing the cue, target, response, and the conjunction of all three (i.e., the event file). The t values of these predictors were saved as a measure of similarity. A fifth vector of z-scored/averaged reaction times for each class was also included to account for variations in RT.

Mixed-model analysis. As in Kikumoto and Mayr (2020), the representational strengths of each of the four factors of interest (cue, target, response, conjunction) in the neural data were related to behavior (reaction time) using a mixed-effects model. The t values from the RSA were baseline corrected (first 100 ms) for each subject. Values outside ±5 × SDs the mean were excluded, and the data were then smoothed using an overlapping 80 ms moving mean window. Subject RTs were then log transformed (ln) and detrended to remove any linear and/or quadratic trends. Subjects’ data were first modeled with the RSA values for cue, target, response, and conjunction added as fixed effects predicting the RT on the same trial. Two models were created to predict trial N RT, one using trial N-2 predictors (Eq. 2) and the other, N-1 predictors (Eq. 3). In both cases the N-1 or N-2 RSA values for cue, target, response, and conjunction were added as fixed effects, as well as the trial N triplet (ABA-S or CBA). Interactions for all RSA values with the trial N conditions were also included. Subject was included as the only random effect. This themselves did not explain their choice between the t statistic and the raw betas.)

Mixed-model analysis. As in Kikumoto and Mayr (2020), the representational strengths of each of the four factors of interest (cue, target, response, conjunction) in the neural data were related to behavior (reaction time) using a mixed-effects model. The t values from the RSA were baseline corrected (first 100 ms) for each subject. Values outside ±5 × SDs the mean were excluded, and the data were then smoothed using an overlapping 80 ms moving mean window. Subject RTs were then log transformed (ln) and detrended to remove any linear and/or quadratic trends. Subjects’ data were first modeled with the RSA values for cue, target, response, and conjunction added as fixed effects predicting the RT on the same trial. Two models were created to predict trial N RT, one using trial N-2 predictors (Eq. 2) and the other, N-1 predictors (Eq. 3). In both cases the N-1 or N-2 RSA values for cue, target, response, and conjunction were added as fixed effects, as well as the trial N triplet (ABA-S or CBA). Interactions for all RSA values with the trial N conditions were also included. Subject was included as the only random effect. This
analysis was applied to all time points of the trials. All mixed-model analyses were conducted in R (version 4.0.4, packages car, lme4) as follows:

\[
\ln(\text{rt}(N)) - 1 + \text{cue}(N) + \text{tar}(N) + \text{resp}(N) + \text{conj}(N) + (1|\text{subject})
\]

(1)

\[
\ln(\text{rt}(N)) - 1 + (\text{cue}(N-2) + \text{tar}(N-2) + \text{resp}(N-2) + \text{conj}(N-2)) + \text{triplet}_{\text{ABA-S\&CBA}} + (1|\text{subject})
\]

(2)

\[
\ln(\text{rt}(N)) - 1 + (\text{cue}(N-1) + \text{tar}(N-1) + \text{resp}(N-1) + \text{conj}(N-1)) + \text{triplet}_{\text{ABA-S\&CBA}} + (1|\text{subject})
\]

(3)

**Cross-correlation analyses.** To investigate the relationships between the strengths of the neural task representations across the trials of a triplet, the same response-locked data used in the final mixed model (Eqs. 2, 3) were then cross-correlated for each subject. The RSA values were first averaged down to 45 time points (900 ms in 20 ms steps) for the following analysis. The conjunction strength in trial N was predicted using four mixed models, one for each N-2 predictor (cue, target, response, conjunction). These models included both ABA-S and ABA-R triplets, with CBA triplets as the reference group and subject as the random effect (Extended Data Fig. 8-1). These models were run for each time point in trial N and N-2, producing \(45 \times 45\) correlation matrices of \(\beta\) coefficients. A similar analysis was performed using N-1 RSA values to predict trial N conjunction strength. However, these models differed in that they were performed for each triplet separately and included all trial N-1 task features as predictors (Extended Data Fig. 8-4). The purpose of this was to remove the comparisons between triplets, so the true pattern of N-1 to N correlation could be observed.

To obtain single values for each subject, trial triplets were sorted into three condition groups (ABA-S, ABA-R, CBA). The RSA values for all task features at every time point in trials N-2 and N-1 of each group were compared with the corresponding RSA values at every time point in trial N, using Pearson’s correlations. This produced a \(225 \times 225\) matrix of Pearson’s \(r\) values comparing the task features in trials N-2 and N-1 with those in trial N for each condition and subject. These matrices were Fisher’s \(z\) transformed before plotting and further testing (Fisher, 1915). The difference between the cross-correlation matrices comparing the conjunctions in trial N-2 and N for ABA-S and CBA triplets was tested using a one-sided permutation test for dependent measures (10,000 permutations; Gerber, 2022). A difference matrix (ABA-S – CBA) was produced for each subject and then averaged across both dimensions, postbaseline (first 100 ms), resulting in a single value per subject (Extended Data Fig. 8-2). This was correlated with each subject’s N2RC (Pearson’s).

**Results**

**Behavior**

Subjects failed to respond on 1.09% of trials and committed response errors on 4.05% of trials overall. As predicted, a comparison of ABA-S and CBA trials reveals N-2 Repetition Cost effects on both RT (ABA-S, mean = 748 ms, SE = 26.3; CBA, mean 707.4 ms, SE = 26.1; Fig. 4A) and accuracy (ABA-S, mean = 94.2%, SE = 0.82; CBA, mean 95.1%, SE = 0.74). Paired-samples \(t\) tests revealed that these differences were highly significant for both RT \(t(34) = 7.48, p < 0.001\) and accuracy \(t(34) = -3.74, p < 0.001;\) Fig. 4A).

**Decoding conjunctive representations on the current trial**

In a first step, we aimed to replicate the results from Kikumoto and Mayr (2020) to validate that we could identify the conjunctive representation of cue, target, and response (i.e., the event file) from the neural data. Following EEG epoch rejections, each subject had an average of 517.4 (±58.0) cue/target-locked trial epochs and 515.7 (±58.2) response-locked trial epochs. From the EEG data of each subject, we successfully produced significant point-by-point estimates of representational strength for all task features (cue, target, response) and their unique conjunction (event file). Single-subject RSA averages (\(r\) values from regression of models onto LDA output) were plotted as a grand average, time locked to both cue/target onset (Fig. 5A) and response onset (Fig. 5B). Notably, the representations for the task cue, target, and conjunction were significantly decoded across both time series. The response representation was only significantly decoded when the data were response locked, likely because of variance in RT (Fig. 4B).

Furthermore, the mixed-effects model approach produced results similar to those of Kikumoto and Mayr (2020; Fig. 6). Analysis of the response-locked RSA values (Fig. 5B) showed that increased strength of the cue, target, and conjunctive representations predicted faster RTs on the same trial. The response representation did not produce a significant relationship by itself. (However, it did in the stimulus data; Extended Data Fig. 6-1.)

**Main analysis: trial N-2 conjunctions explain RT cost on trial N**

We then tested our main hypothesis, that is, in ABA-S triplets, the neural representation of the conjunction (event file) on trial N-2 would negatively affect the reaction time on the current trial N. A mixed-effects model was used to investigate the relationship between the trial N-2 representational strength of each task factor (cue, target, response, conjunction) and the reaction time of trial N, using the response-locked RSA values. The \(\beta\) coefficients for the interaction terms of each model were plotted at each time point in Figure 7. As hypothesized, the strength of the conjunctive representation in trial N-2 significantly predicted increased trial N RTs. Significant time periods were found from \(-175\) to \(-83\) ms and from \(-63\) to \(-33\) ms, relative to the response, for ABA-S triplets compared with CBA triplets (FDR < 0.05; Fig. 7, left). In other words, the greater the strength of the conjunctive representation on trial N-2, the slower the RT of trial N in an ABA-S triplet. This was only true for the conjunctive representation and not for any individual factor (cue, target, response) by itself. Moreover, the conjunctive representation on the immediately preceding trial N-1 did not predict anything about trial N RT (Fig. 7, right). Beyond these interactions, the mixed-model analysis did not show any simple main effects that survived FDR correction \(p < 0.05\); Extended Data Fig. 7-1). These analyses were also repeated using the stimulus-locked data but yielded no significant interactions or main effect (Extended Data Fig. 7-2).

**Exploratory analysis: quantifying cross-trial interactions between conjunctive representations**

To further explore the cross-trial dynamics that give rise to the N2RC, we calculated cross-correlations between the representational strengths of the conjunctive representations on the first and third trials within ABA and CBA triplets (Fig. 8). The output of the mixed-model analysis shows a significant negative main effect (FDR < 0.05), which is likely driven by CBA triplets. For ABA-S triplets, the interaction shows a notable but not significant positive difference from CBA triplets, peaking after the response in trial N (Fig. 8A). This suggests that in CBA triplets, stronger conjunctions on trial N-2 corresponded to the formation of weaker conjunctions on trial N, indicating possible
interference between the different event files corresponding to tasks C and A. This effect is either absent or marginally reversed on ABA-S trials, whose event files partially overlap. These findings are also consistent with the Pearson’s correlation matrices (Extended Data Fig. 8-2). For ABA-R triplets, a large and significant positive difference from CBA triplets was observed. As expected, the identical event files of trials N-2 and N in these triplets are positively correlated. To investigate the critical contrast underlying the N2RC, we took the difference between the N2-N Pearson’s correlation matrices of ABA-S and CBA triplets and averaged them to obtain a single value per subject (Extended Data Fig. 8-2). These differences in correlation directly related to the size of the N2RC across subjects; that is, subjects with greater differences in the Pearson’s cross-correlations between CBA and ABA-S trials showed greater N-2 repetition costs (Fig. 8B). Control analyses revealed that the same was not true for the correlations between nonconjunctive representations (cue, target, response). As would be expected, this effect appears to be driven by an increase in the RTs of ABA-S triplets, as opposed to a decrease in CBA triplet RT (Extended Data Fig. 8-3).

To buttress this analysis, we explored whether the correlation patterns of other trial combinations matched the above-mentioned interpretation. Regardless of triplet condition, trial N-1 never fully overlaps with the conjunctions of trial N-2 or N. As such, the mixed-model output matrices for trial N-1 conjunctions predicting trial N conjunction strength all show a notable negative correlation consistent with nonmatching event files (Fig. 8C). This suggests that when there is a mismatch in task features as early as the cue (i.e., a task switch), a stronger conjunctive representation of a past task feature combination will impair the formation of the new conjunction on the current trial. Together, these correlation patterns suggest that the formation of a neural representation of the task feature conjunction on a past trial (N-2) benefits the formation of the conjunctive neural representation on the present trial (N) if the task sets match and is detrimental to the formation of a different conjunction if they mismatch. Moreover, the cross-subject Pearson’s correlations suggest that the strength of these processes is directly proportional to the size of the N2RC effect. In contrast to our main analysis, these exploratory analyses were performed post hoc and therefore must be interpreted with caution. However, they generate further insights into the neural dynamics of how the trial N-2 task set may affect behavior on trial N to affect the N2RC.

**Discussion**

In the current study, we used a recently established neurophysiological method for tracking conjunctive representations between specific task contexts, stimuli, and responses (i.e., event files; Hommel et al., 2001) to address an ongoing debate about the role of such event files in task switching. Based on the methodological groundwork of Kikumoto and Mayr (2020), we leveraged the combination of RSA and MVPA of EEG data to investigate...
the trial-to-trial relationship between the strength of conjunctive event-file representations and behavior. However, although Kikumoto and Mayr’s (2020) investigation evaluated immediate task repeats/switches, we assessed whether the formation of an event file on the current trial can influence behavior two trials later. Using the N-2 Repetition Cost as a behavioral model, we found that the strength of the conjunctive representation of cue, target, and response on a previous trial indeed adversely influences performance on the current trial, specifically, in cases in which response requirements changed between two instances of the same task. In this way, we provide a direct test of a critical component of the event file hypothesis, that is, that conjunctive representations are encoded and retrieved for the performance of future actions (Hommel et al., 2001; Hommel, 2019; Frings et al., 2020).

Kikumoto and Mayr (2020) demonstrated that partial-repetition costs can be partially explained by the strength of the conjunctive representation formed in the immediately preceding trial. The more strongly encoded the conjunction (event file) is on trial N-1, the more difficult it would be to unbind the constituent task features and reintegrate them into a new conjunction, leading to greater partial-repetition costs. By replicating this effect across trial triplets (ABA and CBA), we were able to show that conjunctive representations formed on trial N-2 can also affect behavior on trial N, despite there being an interleaving task switch (N-1). These findings further support the Theory of Event Coding (Hommel et al., 2001), which proposed that event files are created during each trial of a task and can also be automatically retrieved and reintegrated across extended periods of task performance.

Furthermore, the exploration of correlation patterns between the strength of conjunctive representations on different trials within triplet conditions generated further insight into the dynamics by which this adverse influence of past event files may manifest itself. First, the comparison of nonmatching task contexts (trial N-1 vs trial N comparisons) revealed a stereotypic pattern of decorrelation between the strength of the respective conjunctive representations of successive, different task contexts (Fig. 8C). In other words, the stronger the representational strength of feature conjunctions on the current trial, the

Figure 5. Decoding current trial representations of cue, target, response, and their conjunction from the neural data. A, Grand average of the cue/target-locked RSA output. B, Grand average of response-locked RSA output.
weaker the representational strength of feature conjunctions of subsequent nonmatching trials. This suggests that once-formed conjunctive representations may generally interfere with the formation of feature conjunctions on subsequent trials, a phenomenon that may underpin task-switch costs more generally. Moreover, as we observed in the main effect of our mixed-model output comparing the conjunctive strength between trials N-2 and N (Fig. 8A), the comparison of task C and task A conjunctions of CBA triplets revealed that this pattern extended beyond the immediately subsequent trial; that is,
stronger representations of feature conjunctions on a given trial still impair the formation of feature conjunctions two trials later. Most interestingly, the comparison of ABA-R triplets to CBA triplets provided a possible explanation on why this takes place, as it showed the opposite pattern between the N-2 and trial N conjunctions (i.e., between two spaced trials with matching event files). In such triplets, stronger conjunctive representations on the first instance of task set A (trial N-2) correlated with stronger conjunctive representations on the second instance (trial N). In line with the event-file hypothesis, this suggests that the adaptive function of this apparent forward propagation of a task set is to benefit behavior if the task set repeats/remains the same. However, if the task set switches, the anticorrelation pattern ensues. Such an automated storage-and-retrieval process makes sense in everyday scenarios the brain has evolved to handle, in which task sets do not typically change with the same regularity as in laboratory experiments such as the current one. ABA-S triplets then form a middle ground between the mismatch scenarios provided by CBA triplets, on the one hand, and the full-match scenarios provided by ABA-R trials, on the other; the decorrelation observed on mismatch trials does not take place as the strength of the N-2 conjunction does not adversely influence

**Figure 8.** Exploration of the representational cross-correlation matrices. A, The three correlation matrices represent the relationship between the conjunction strength of trials N-2 and N at each time point. The main effect and two interactions (left to right) were produced using a mixed model (Extended Data Fig. 8-1). Regions of significance following FDR correction \((p < 0.05)\) are denoted by black lines. B, Scatter plots show the relationship between subjects’ averaged difference in ABA-S and CBA Pearson’s cross-correlation matrices (Extended Data Fig. 8-2, N-2 and N) and their individual N2RC for all task features. Only the conjunction shows a significant relationship \((r = 0.37, p = 0.03)\). This finding remains true even when comparing ABA-S and CBA trials separately (Extended Data Fig. 8-3). C, These example matrices show the distinct patterns of nonoverlapping (N-1 and N) and conjunctive representations for all three triplet conditions. The mixed model also included the individual task features from trial N-1 as predictors, which similarly showed an absent or negative correlation with trial N conjunctive strength (Extended Data Fig. 8-4). For ABA-S triplets, the N-2 conjunction was unable to be decoded on trial N (Extended Data Fig. 8-5).
the formation of the task N conjunction (Fig. 8). However, as the trial unfolds and the target indicates that response requirements have changed compared with the most recent presentation of the cue, conflict ensues and results in the behavioral cost expressed in the N2RC effect. Indeed, as seen in Figure 8B, the difference in correlation between N-2 to N trial conjunctions in ABA-S triplets compared with CBA triplets (i.e., the main contrast of interest in the current study) was directly correlated with the N2RC effect across subjects. This correlation is very much in line with the results of our main analysis, that is, that the strength of N-2 conjunctive representations predicted longer reaction times in trial N of an ABA-S triplet. Indeed, subjects for whom the conjunctive trial N-2 representation was stronger—and therefore, purportedly interfered more strongly with the formation of the conjunctive trial N representation within a CBA triplet (or primed the response contained in the trial N-2 event file to a greater extent on trial N within an ABA-S triplet)—showed a larger N2RC effect.

Regarding our main finding—the inverse relationship between the strength of the conjunctive representation on trial N-2 and RT on trial N of an ABA-S triplet—two observations are key to rule out alternative explanations. First, the conjunctive representation of the intervening trial N-1 in an ABA-S triplet did not predict RT on the current trial. This confirms that the N-2 Repetition Cost is indeed specific to a precise match between task contexts. Second, neither the cue nor target representations of trial N-2 by themselves predicted behavior on trial N. This shows that regardless of how strongly the target or cue was encoded on trial N-2, only the strength of their binding with the task context in a conjunctive representation affected future behavior. As such, our results are consistent with previous work, suggesting that the episodic retrieval event files significantly contribute to the N2RC effect (Neill, 1997; Hommel, 1998; Spapé and Hommel, 2010; Kühn et al., 2011; Grange et al., 2017; Grange, 2018; Frings et al., 2020). Additionally, we are the first to demonstrate that a direct measure of task representational strength from neural data can be used to predict future RT costs.

As such, we successfully used the decoding of conjunctive representations from neural recordings to not only show that greater task-feature binding can both be beneficial to current behavior and detrimental to future behavior but we also demonstrated how neurophysiological methods can provide crucial insights into behavioral phenomena that have hitherto mainly been investigated using experimental psychology. Moreover, our results provide a powerful demonstration of the event-file framework in explaining behavior and linking it to neural activity. Indeed, conjunctive representations of stimuli, actions, and outcomes appear to be a fundamental code that the human brain uses to guide future behavior. Although early conceptualizations of event or object files focused on the conjunction of visual stimulus representations (Kahneman et al., 1992; Treisman, 1998), the idea that actions and their outcomes can be commonly linked with perceptual information is hence a key insight into human adaptive behavior (Hommel, 1998; Hommel et al., 2001; Spapé and Hommel, 2010). This proposal sheds light on how event files allow us to do more than simply recognize a familiar sensory stimulus but also contextualize that stimulus by its expected effects, our own actions, and the surrounding environment.

One open question that remains is the role of backward inhibition in the N2RC effect. Past studies have shown that ABA triplets can show N-2 Repetition Costs even when the current response matches the past response (Grange et al., 2017; Grange, 2018), which is a counterintuitive finding under the event-file framework. The magnitude of these costs tends to be smaller than those in ABA-S trials, suggesting that they may represent a pure form of backward inhibition that can only be properly examined when the effects of episodic retrieval are controlled for (Grange, 2018; Kowalczyk and Grange, 2020). Although ABA-R trials were not the primary interest of the current study (as the experiment only contained a limited number of ABA-R trials and used a cue-target interval that was too long to expect substantial RT effects on ABA-R trials), the framework presented here could be adapted to specifically test this hypothesis. Indeed, neurophysiological signatures of inhibitory processes are well established in both the motor and cognitive domains (Anderson et al., 2016; Wessel and Aron, 2017), and it has long been proposed that task switch cues activate inhibitory networks (Jamadar et al., 2010; Koch et al., 2010). Some EEG studies examining the N2RC as a function of backward inhibition have produced contrary results, with differences in event-related potentials appearing over centroparietal and parietal sites (Sinai et al., 2007; Zhang et al., 2016) rather than over frontocentral electrodes, as would be expected. However, these studies did not properly account for episodic retrieval effects and were also testing additional task conditions that would obscure the results. The combination of the current approach with direct measurements of inhibition could provide insight into whether there is indeed an additional role for inhibition in the N2RC effect. Moreover, the analytical method developed by Kikimoto and Mayr (2020) could be adapted to identify signatures of the latent conjunctive representations established in trial N-2 before its retrieval on trial N (Rose et al., 2016) and to investigate whether task switching cues on the intervening trial N-1 directly affect the strength of the latent conjunctive representation.

Indeed, to perform an additional test of the retrieval hypothesis suggested by a reviewer, we attempted to decode the conjunctions formed during trial N-2 of ABA-S triplets during trial N as a post hoc analysis. In particular, it was suggested that the representation of the trial N-2 conjunction may be explicitly decodable from the EEG data of trial N and that the strength of this recollection of trial N RT, just like we found for its initial strength on trial N-2. This turned out not to be the case as no significant N-2 conjunction was able to be decoded during trial N (Extended Data Fig. 8-5). There are a few possible explanations for this finding. It is possible that the N-2 conjunction may be held as an activity-silent representation in working memory after being retrieved (Rose et al., 2016). Alternatively, the N-2 conjunction may simply lack the representational strength to be decoded from a sample of our size. Future investigations should focus on this aspect of the event-file hypothesis as it may be able to provide more direct evidence of how and when past event files are retrieved and reintegrated into working memory.

In summary, we here used the recently developed methodological approach for the measurement of conjunctive representations of task context, stimuli, and responses (Kikimoto and Mayr, 2020) to demonstrate that such representations, once formed, can adversely affect future behavior, here, specifically in the form of the N2RC effect. This not only highlights the usefulness of this analytical approach for addressing questions relating to cognitive control processes but it contributes to the extant literature on the role of episodic retrieval in N-2 Repetition Costs. Hence, this work provides fundamentally novel insights into a long-standing debate in the behavioral literature and illustrates the potential of this neuroscientific approach to tackling questions of adaptive and flexible behavior.
References


