Cross-Level Effects Between Neurophysiology and Communication During Team Training

Jamie C. Gorman, Georgia Institute of Technology, Atlanta, **Melanie J. Martin**, California State University–Stanislaus, Turlock, **Terri A. Dunbar**, Georgia Institute of Technology, Atlanta, **Ronald H. Stevens**, **Trysha L. Galloway**, The Learning Chameleon, Los Angeles, California, **Polemnia G. Amazeen**, and **Aaron D. Likens**, Arizona State University, Tempe

Objective: We investigated *cross-level effects*, which are concurrent changes across neural and cognitive-behavioral levels of analysis as teams interact, between neurophysiology and team communication variables under variations in team training.

Background: When people work together as a team, they develop neural, cognitive, and behavioral patterns that they would not develop individually. It is currently unknown whether these patterns are associated with each other in the form of cross-level effects.

Method: Team-level neurophysiology and latent semantic analysis communication data were collected from submarine teams in a training simulation. We analyzed whether (a) both neural and communication variables change together in response to changes in training segments (briefing, scenario, or debriefing), (b) neural and communication variables mutually discriminate teams of different experience levels, and (c) peak cross-correlations between neural and communication variables identify how the levels are linked.

Results: Changes in training segment led to changes in both neural and communication variables, neural and communication variables mutually discriminated between teams of different experience levels, and peak cross-correlations indicated that changes in communication precede changes in neural patterns in more experienced teams.

Conclusion: Cross-level effects suggest that teamwork is not reducible to a fundamental level of analysis and that training effects are spread out across neural and cognitive-behavioral levels of analysis. Cross-level effects are important to consider for theories of team performance and practical aspects of team training.

Application: Cross-level effects suggest that measurements could be taken at one level (e.g., neural) to assess team experience (or skill) on another level (e.g., cognitive-behavioral).

Keywords: communication, coordination, cross-correlation, latent semantic analysis, neurophysiology, teams

Address correspondence to Jamie C. Gorman, School of Psychology, Georgia Institute of Technology, 654 Cherry Street, Atlanta, GA 30332, USA; e-mail: jamie.gorman@gatech.edu.

HUMAN FACTORS
Vol. XX, No. X, Month XXXX, pp. 1–19
DOI: 10.1177/0018720815602575

INTRODUCTION

There is general agreement that individual neural, cognitive, and behavioral processes change and develop as people interact with the world around them, but are some of these changes correlated across individuals when they work together as a team? In prior research, we have found that when people work together as a team, neurophysiological (Stevens, Gorman, Amazeen, Likens, & Galloway, 2012; Likens, Amazeen, Stevens, Galloway, & Gorman, 2014), perceptual-motor (Gorman & Crites, 2013, 2015), and cognitive-behavioral (e.g., communication; Gorman & Cooke, 2011) patterns emerge between team members that would not otherwise develop individually. However, it is unknown how these different patterns are linked across levels of analysis (e.g., neural and cognitive-behavioral levels) as teams interact.

With this in mind, we investigate *cross-level effects*, which are concurrent changes across neural and cognitive-behavioral levels of analysis as teams interact. Specifically, we investigate whether cross-level effects exist between brainwaves measured through electroencephalography (EEG) and team communication under variations in team training and amount of team experience. In this paper, we attempt to address theoretical and practical implications of cross-level effects. If cross-level effects exist, then new forms of team skill assessment may be revealed, and the relative contributions of neural and cognitive-behavioral processes to team development may be clarified.

Theoretical Underpinnings of Cross-Level Effects

Most theories that are relevant to understanding team performance either implicitly or explicitly suggest a hierarchical arrangement of neural and cognitive-behavioral levels. Some suggest that team performance is fundamentally caused by inner mental processes and structures in the individual that are subsequently coordinated through communication (either overtly or implicitly). For example, joint action (Sebanz, Bekkering, & Knoblich, 2006), shared cognition (Cannon-Bowers & Salas, 2001), and social neuroscience (Frith & Wolpert, 2004) posit fundamental neural processes and knowledge structures as the basis for interpersonal coordination and team performance. On the other hand, others suggest that emergent team-level constraints, such as communication patterns, better explain team coordination and performance variance, where changes in inner mental processes and knowledge structures are meaningful only in light of interpersonal interactions (e.g., interactive team cognition; Cooke, Gorman, Myers, & Duran, 2013; team coordination dynamics; Gorman, 2014). These different theoretical perspectives link neural and cognitive-behavioral levels but introduce the question of which should come first, the neural or cognitive-behavioral level, in explaining team performance and skill development.

To fully address this question is beyond the scope of this paper, but we will explore it by examining how lead-lag cross-correlations between neurophysiological and communication variables develop by comparing cross-level effects in more versus less experienced (skilled) teams. Although we think that it is obvious that one cannot have team communication patterns without neural patterns (or vice versa), by examining lead-lag cross-correlations between these two levels of analysis, we will test which tends to come first, the neural or communication level, in the establishing of cross-level effects.

Practical Implications of Cross-Level Effects

From a practical standpoint, evidence for cross-level effects could suggest new ways to measure team skill development. For example, when a team shows improvement on one level of analysis (e.g., communicating effectively), then it would be changing on the other, less well-understood level of analysis (e.g., team neural pattern). Insofar as neural and communication variables exhibit cross-level effects, then one level could be used to better understand the development of effective teamwork along the other. For example, communication indicators of effectiveness that can be observed by instructors could be transformed into neural metrics that could be automatically monitored by machines.

Measuring Cross-Level Effects

To measure cross-level effects, we chose neural and communication variables based on their previous success in capturing performance variability and skill level during team training. The team neural variable is measured using an EEG-based neurodynamic approach (Stevens, Galloway, Wang, & Berka, 2012), and communication variables are measured using latent semantic analysis (LSA; Landauer & Dumais, 1997) team communication metrics (Gorman, Foltz, Kiekel, Martin, & Cooke, 2003).

Team neurophysiology: The neurodynamic approach. Neurophysiological processes reflect variability in mental state that can be measured using EEG. Specifically, EEG produces oscillatory brainwave signals related to brain functioning that can be mapped onto mental state variables, such as neurophysiological engagement. Engagement is essentially a measure of attention (Berka et al., 2007), and the neurodynamic approach introduces a new, team-level variable that captures the distribution of engagement states across team members. For example, whether all team members have high engagement, some have high and some low engagement, or all have low engagement are all different team neurophysiological distributions that map onto qualitatively different team mental states in the team neurodynamic approach.

In this approach, qualitative change of the team neurophysiological distribution is measured over time and is then transformed into a *neurodynamic entropy* time series (this procedure is described in detail later). Neurodynamic entropy is a continuously varying index of how much the team neurophysiological distribution is changing. Low entropy means that the team neurophysiological distribution is changing less, and high entropy means that it is changing more; low entropy can be interpreted as a relatively fixed team mental state, and high entropy can be interpreted as a more flexible team mental state.

It is critical to note that low entropy does not mean that all team members simultaneously share a high joint focus of engagement (attention) but means simply that the distribution of engagement across team members, whatever that distribution looks like, is relatively fixed over time. Conversely, high entropy is thought to be associated with team flexibility and the ability to adapt to dynamic changes in the task environment or team structure because the team neurophysiological distribution is highly responsive to changes in the team or task environment (Stevens et al., 2012).

Through the lens of interactive team cognition and team coordination dynamics (Cooke et al., 2013; Gorman, 2014), we would expect teams to change not only their communication patterns but also their neural patterns to match changes in task dynamics. For example, in a relatively predictable and stable team task, we would expect the team neurophysiological distribution to be relatively fixed and stable (low entropy); whereas in a highly dynamic, unpredictable team task, we would expect the team neurophysiological distribution to be more variable and flexible (high entropy) if the team is responding appropriately to those different task dynamics. Our study of cross-level effects is motivated by our initial observations that experienced submarine piloting and navigation (SPAN) teams become more neurally flexible during more dynamic training segments and more neurally fixed during more predictable and stable training segments in accordance with task demands (Stevens et al., 2012) and, moreover, that those neural changes appear to be correlated with changes in communication content patterns between those training segments.

For example, during the scenario training segment in Figure 1, entropy fluctuates around a relatively high value throughout the segment as team members communicate new and evolving information (a highly dynamic task). In contrast, during debriefing, entropy precipitously drops to a low value, indicating a more fixed team neurophysiological distribution, as team members communicate to reach consensus about information they previously encountered during the

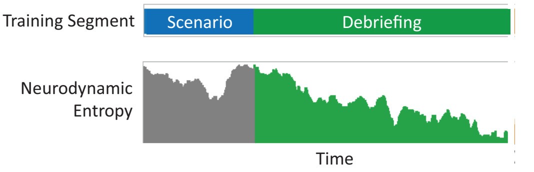

Figure 1. Neurodynamic entropy pattern as a function of scenario and debriefing training segments from prior submarine piloting and navigation (SPAN) research (Stevens, Gorman, Amazeen, Likens, & Galloway, 2012).

scenario (a comparatively predictable and stable task). On the basis of the communication that goes on during training segments such as these, we predicted that changes in entropy should correspond to changes in communication pattern in response to the different task dynamics encountered in SPAN training segments.

Communication content analysis: The LSA approach. Whereas neurodynamic entropy provides a gauge on neurophysiological change across team members, communication content analysis provides a gauge on the interactive expression of domain-relevant knowledge as team members interact (Cooke & Gorman, 2009). We analyzed communication content using LSA (Landauer, Foltz, & Laham, 1998). LSA is a mathematical/statistical method for representing and analyzing semantic knowledge in a particular work domain and is based on the theory that knowledge is reflected in how words group into contexts within meaningful discourse (Landauer & Dumais, 1997).

LSA has many applications for knowledge assessment, from education and testing (Islam & Latiful Hoque, 2010) to communication analysis in real-time work domains (Dong, 2005). LSA has been primarily used in team performance domains to study the relationship between communication content and team skill level (Foltz & Martin, 2009; Gorman et al., 2003). This objective is accomplished by plotting transcribed communications (e.g., utterances) in a "semantic space," where the semantic space is a factor-analytic model of the domain of discourse. As we describe later, plotting utterances in the semantic space provides quantitative measures of the amount of domain-specific content contained in utterances (semantic content) and how correlated those utterances are with each other (semantic similarity).

We have previously used semantic similarity to discriminate between more and less skilled teams using a military unmanned air vehicle (UAV) semantic space (Cooke & Shope, 2005; Martin & Foltz, 2004), where the communications of more skilled teams were more similar to each other than to less skilled teams. In this way, LSA can be used to link communication content to team skill level, which is relevant to distinguishing between more versus less experienced SPAN teams in the current study. Also, we have previously used semantic content of utterances to distinguish between different UAV task dynamics (e.g., low workload versus high workload; Cooke, Gorman, Kiekel, Foltz, & Martin, 2005), which is relevant to distinguishing between different SPAN training segments (e.g., scenario vs. debriefing; Figure 1). These prior results are important because they demonstrate that LSA communication metrics (semantic similarity, semantic content) capture the cognitive-behavioral variability required to identify cross-level effects across changes in training segments and differences in team experience level in SPAN teams.

Research Questions for Testing Cross-Level Effects

Do both neural and communication variables change in response to changes in training segments? We hypothesized that if cross-level effects are present, then they should be reflected in concomitant changes across both neural and communication variables in response to variations in training segments, such as those in Figure 1. We refer to concomitant changes across neural and communication variables as *concomitancy*. Alternatively, if cross-level effects are not present, then changes in one level should not correspond to changes in the other level as training segments vary. To test for concomitancy, we calculated LSA communication metrics and neurodynamic entropy from SPAN teams to determine whether both variables change together as training segments (briefing, scenario, debriefing) vary.

Do neural and communication variables mutually discriminate between teams of different experience (skill) levels? We have previously found that when teams are of different skill levels, it is reflected in differences in the semantic similarity of their communication (Foltz & Martin, 2009; Gorman et al., 2003). Similar to the idea of the development of "common ground" in communication (Clark & Brennan, 1991), it is thought that the communication of more skilled teams is more similar compared to less skilled teams because they have interacted longer in a particular task domain. In this study, we extend that logic to more versus less experienced SPAN teams. We predicted that more experienced SPAN teams could be discriminated from less experienced SPAN teams based on the similarity of their communications. In accordance with our general aim of identifying cross-level effects, we carried out a parallel neurodynamic analysis to determine whether entropy similarly discriminates between more and less experienced SPAN teams. If both neural and communication variables discriminate between teams of different experience levels, then we refer to it as *mutual discrimination* and take it as an indicator that cross-level effects correspond to differences in team skill development across both neural and cognitive-behavioral levels of analysis. In that case, either or both variables could be used to discriminate between teams with different experience (skill) levels.

Assuming cross-level effects occur, how are levels linked? Communication and neural processes can span different timescales, from long conversations in the former case to fleeting synaptic processes in the latter. Hence, these levels might become linked in temporally complex ways as teams develop. Therefore, we used lead-lag cross-correlations (Box, Jenkins, & Reinsel, 1994) to identify linkages between neurodynamic entropy and semantic content across multiple temporal alignments of these two variables.

Cross-correlation functions measure the direction of correlation between two variables as a function of the temporal offset ("lag") between them. By examining the direction of the "peak" (maximum) cross-correlation (i.e., positive or negative correlation), we expect to learn how team neurodynamics and communication are related. For example, if neural flexibility (high entropy) is associated with terse domain-specific communication (low semantic content), then the correlation will be negative; if neural flexibility (high entropy) is associated with lengthier, more open-ended domain-specific communication (high semantic content), then the correlation will be positive. By lining up the variables in different temporal alignments (e.g., we could line up current communication values at time t with future neural values at time $t + 1$ [Lag 1] and compute the correlation, which would be one possible temporal alignment) and finding the lag where the peak cross-correlation occurs, we expect to learn whether the neural or communication process tends to lead as cross-level effects develop.

In this light, differences in the lead-lag nature of peak cross-correlations across more and less experienced teams may provide insight into the theoretical question of which tends to come first, the neural or communication level, as teams develop. The view that team performance is fundamentally caused by mental and neural activity in the individual that is subsequently expressed as cognitive-behavioral (e.g., communication) variability across team members suggests that neural should be leading and communication lagging in these peak cross-correlations. Alternatively, the view that cognitive-behavioral constraints, such as emergent communication patterns, constrain neural and mental changes in the individual suggests that communication should be leading and neural lagging in these peak cross-correlations.

The Current Study

In addressing these research questions, we tested three predictions in the context of SPAN that should be met with positive results if cross-level effects are present: (a) concomitancy of training segment effects on neurodynamic entropy and communication metrics, (b) that semantic relatedness and neurodynamic entropy mutually discriminate between teams of different experience levels, and (c) that the development of cross-level effects can be observed through changes in lead-lag peak cross-correlations between neurodynamic entropy and communication across more and less experienced teams.

METHOD

Participants

Neurophysiology and communication data were obtained from junior officer navigation teams enrolled in the Submarine Officer Advanced Candidacy class at the U.S. Navy Submarine School in Groton, Connecticut. These SPAN teams consisted of six crew members: quartermaster on watch, navigator, officer on deck, assistant navigator, contact coordinator, and radar. (Other people were "satellite" team members but were not directly involved in the team processes analyzed here.) We analyzed seven SPAN training sessions, four from more experienced teams (teams that had recently returned to port) and three from less experienced teams (candidates training to become ship's drivers and navigators). It is important to note that more experienced teams were more experienced both with SPAN in general and with working together as a team. Those two factors are not teased apart in the current study, and either could play a role in the cross-level effects described later. We use a between-subjects variable, experience, to index more versus less experienced SPAN teams in the analyses.

SPAN Training

SPAN training focused on instruction and assessment of four levels of team resilience and five team practices (Stevens, Galloway, & Lamb, 2014). The four team resilience levels included unstressed battle rhythm, leader-dependent battle rhythm, team-based resilience, and advanced team resilience. The five team practices included quality of dialogue, decision making, critical thinking, bench strength, and problem-solving capacity. Definitions and examples of each of these aspects of SPAN training were contained in a submarine team behaviors instructor manual, and instruction and assessments were performed by a high-ranking submarine commander.

For each team, SPAN training consisted of performing three training segments: briefing, scenario, and debriefing. Overall goals of the scenario, team member responsibilities, and task coordination were planned out during briefing. Scenario required teams to steer and change course or speed while identifying landmarks and other ships that factor into SPAN; scenario was the most dynamic training segment, wherein team members coordinated novel and evolving information to navigate a submarine through a high-fidelity simulation environment. Debriefing was essentially an after-action review, during which the team members discussed what worked and different actions that could have been taken based on events that unfolded during the scenario. We use a within-subjects variable, training segment, to index briefing, scenario, or debriefing in the analyses.

Measures

Neurodynamic entropy. The team neurophysiology measure (neurodynamic entropy) is derived from the EEG-based neurophysiological symbol (NS) method (Stevens, Galloway, et al., 2012; Stevens, Gorman, et al., 2012). B-Alert® X10 headsets from Advanced Brain Monitoring, Inc., were used for EEG data collection. These wireless headsets included electrocardiography and nine referential EEG channels located at F3, F4, C3, C4, P3, P4, Fz, Cz, and POz in a monopolar configuration referenced to linked mastoids. Eye blinks and electromyography artifacts were decontaminated using proprietary Advanced Brain Monitoring, Inc., algorithms (Berka et al., 2004). Neurocognitive tasks were first used to time and record presentation and responses to stimuli in order to generate individual models of engagement prior to team performance (for validation of these metrics and task details, see Johnson et al., 2011).

The neurocognitive tasks used to build the individual models were presented using acquisition software proprietary to Advanced Brain Monitoring, Inc. This software contains algorithms that were trained using EEG data collected during the Osler maintenance-of-wakefulness task (Krieger, Ayappa, Norman, Rapoport, & Walsleben, 2004), eyes-closed passive vigilance, eyes-open passive vigilance, and three-choice active vigilance tasks to define the classes of sleep onset, distraction/relaxed wakefulness, and low and high engagement, respectively. The purpose of generating individual models of engagement using neurocognitive tasks was to ensure that we were able to accurately track the engagement of each team member before feeding their data into the team neurodynamic analysis. Using the NS method, the six EEG streams collected individually from each team member are processed to generate a sequence of discrete NS states sampled at a fixed time interval (1 Hz).

The EEG-to-NS mapping is such that each discrete NS state identifies a different neurophysiological distribution of engagement across team members (e.g., Figure 2a): As the task varies, the team neurophysiological distribution changes, and those changes are indexed over time using discrete NS states. The set of NS states for SPAN was determined using an artificial neural network approach, which resulted in a state space of 25 discrete NS states (Figure 2b; Stevens, Galloway, et al., 2012; Stevens, Gorman, et al., 2012). Figure 2c lists the 25 NS states on the left vertical axis, and a bar is plotted whenever each NS state is expressed over time (the horizontal axis). In this way, Figure 2c shows the time series expression of all 25 NS states during a SPAN performance.

Although the number of NS states is fixed, there is no inherent numerical ordering among the states because they are nominal and discrete (e.g., NS State 2 is qualitatively different than NS State 1, but it is not numerically different or larger). Therefore, to quantify change in the team neurophysiological distribution, we calculated entropy (Shannon & Weaver, 1949; Stevens & Galloway, 2014) across the discrete NS time series (Equation 1).

$$\text{NS entropy} = -\sum_{i=1}^{25} p_i \cdot \log p_i. \tag{1}$$

In Equation 1, p_i is the relative frequency of NS state i, where i indexes each of the 25 possible NS states, over a 100-s window. Specifically, entropy is repeatedly calculated as a 100-s window is slid across the NS time series, resulting in a continuous entropy time series (Figure 2c; right axis). Using this technique, for an input NS time series of length N, the output is a continuous univariate entropy time series of length $N - 99$. In this way, we use the first 100 samples to calculate the first entropy value at time $t = 100$, Samples 2 through 101 to calculate the next entropy value at $t = 101$, and

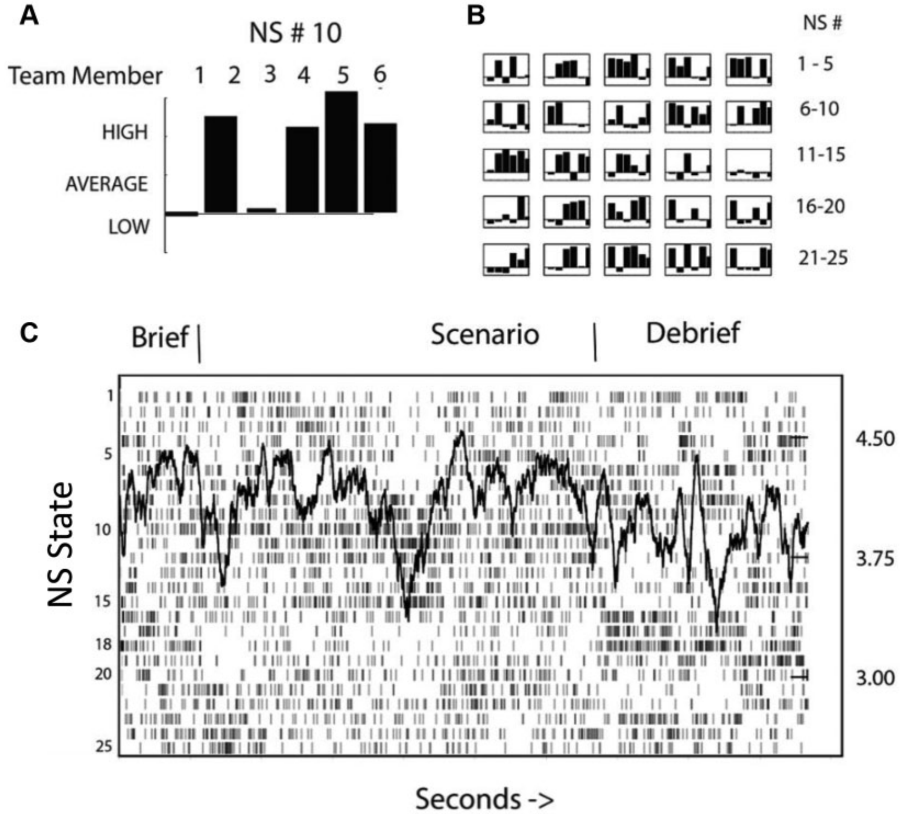

Figure 2. Steps involved in calculating neurodynamic entropy from a six-member submarine piloting and navigation crew. (a) This neurophysiological symbol (NS state) represents times when Team Members 1 and 3 had below-average electroencephalography engagement and the remaining team members (i.e., 2, 4, 5, and 6) had above-average engagement. (b) The 25 NS states for these submarine teams (the numbers on the right assigned to the rows are identifiers used in the left vertical axis of Panel C; the NS state in Panel A corresponds to NS State 10 in Panel B). (c) Each row represents the sequential expression of each of the 25 different NS states over time and is overlaid with a continuous trace of the neurodynamic entropy calculated over the discrete NS states over time. Time is on the horizontal axis, and the fluctuations in the entropy of the distribution of symbol expression across the 25 NS states over time can be viewed by tracking the entropy signal from left to right: When the team has lower entropy, the distribution of NS states is relatively fixed; when the team has higher entropy, the distribution is changing.

so forth. Using a window smaller than 100 s has been found to increase the potential for artefactual spikes in NS entropy time series (Likens et al., 2014; Stevens, Gorman, et al., 2012).

LSA metrics. In constructing a semantic space, LSA takes as input a body of text (e.g., training manuals and transcripts) and starts by representing the body of text as a matrix of frequency co-occurrence of unique words by unique paragraphs. The SPAN semantic space took as input a body of text containing submarine phraseologies, the International Rules of the Road (COLREGS), the Doctrine on Submarine Interior communications, and the seven SPAN training transcripts (including transcripts in the corpus is standard practice; e.g., Foltz, Martin, Abdelali, Rosenstein, & Oberbreckling, 2006). The input dimensions were 6,846 unique words by 5,904 unique paragraphs (124,326 total words).

For the next step, LSA assumes that low-dimensional (latent) semantic factors underlie the observed co-occurrence frequencies between words and paragraphs in the input matrix. These latent factors are uncovered using singular value decomposition, which is similar to the procedure used for principal components factor analysis, where larger singular values (cf. eigenvalues) correspond to more salient factors. The optimal number of factors (dimensions) is chosen such that relationships between words and paragraphs correspond to correct inductions. The optimal number of factors (dimensions) is found by optimizing the semantic space's performance on tests of synonym matching and missing word replacement (Foltz et al., 2006; Landauer et al., 1998). The resulting SPAN semantic space had 314 dimensions, which is consistent with semantic spaces created in other domains (Landauer et al., 1998).

Two metrics derived from the LSA model of communication content are (a) the *vector length* of a piece of discourse and (b) the *cosine* between two pieces of discourse. We use these metrics to analyze (a) the semantic content contained in an utterance and (b) the semantic similarity between different pieces of discourse.

The vector length of an utterance (e.g., "Recommend steering course 178 to regain track.") measures the amount of speech weighted by the domain-specific content the discourse contains. It is calculated as the Euclidean norm of a vector of words (e.g., an utterance) plotted in the semantic space.

The cosine between any two pieces of discourse (e.g., any two utterances, any two training segments, any two transcripts) is the vector dot product between two word vectors containing the discourse plotted in the semantic space. The correlation between two pieces of discourse can be shown to be the cosine of their joining angle when their vectors are plotted in the semantic space (e.g., independent, perpendicular vectors have $\cos[90°] = 0$ and are completely uncorrelated). In other words, cosine measures the degree of semantic similarity, or correlation, between any two pieces of discourse.

RESULTS

Concomitancy

Regarding our first research question, we predicted that if cross-level effects are present, then both neural and communication variables should change together as training segment (briefing, scenario, debriefing) is varied.

To examine the effects of training segment and experience on neurodynamic entropy, we computed mean entropy for each team at each training segment and analyzed those values using a 3 (training segment) × 2 (experience) mixed ANOVA. All ANOVA assumptions were tested and upheld. Only the training segment effect was significant, $F(2, 10) = 19.77$, $p < .001$, $\eta^2 = .80$. A Tukey test on training segment ($\alpha_{FW} = .05$) revealed that debriefing entropy was significantly lower than either briefing or scenario entropy (Figure 3a). This result indicates that the neurophysiological distribution across team members was more flexible during briefing and scenario but more fixed during debriefing.

To determine whether LSA communication metrics were similarly affected, we computed mean vector length (semantic content) across utterances for each team at each training segment and analyzed those values using a 3 (training Segment) × 2 (Experience) mixed ANOVA. All ANOVA assumptions were tested and upheld. As with neurodynamic entropy, only the training segment effect was significant, $F(2, 10) = 15.78$, $p < .001$, $\eta^2 = .76$. A Tukey test on training segment ($\alpha_{FW} = .05$) similarly revealed that vector lengths were significantly different during debriefing compared to briefing and scenario (Figure 3b). This result indicates that communication content was terser and domain specific during briefing and scenario and lengthier and domain specific during debriefing.

Together, these results suggest that training segments (briefing, scenario) that lead to a more flexible team neurophysiological distribution (higher entropy) also resulted in terser domain-specific communication (smaller vector lengths), which we take as evidence for concomitancy.

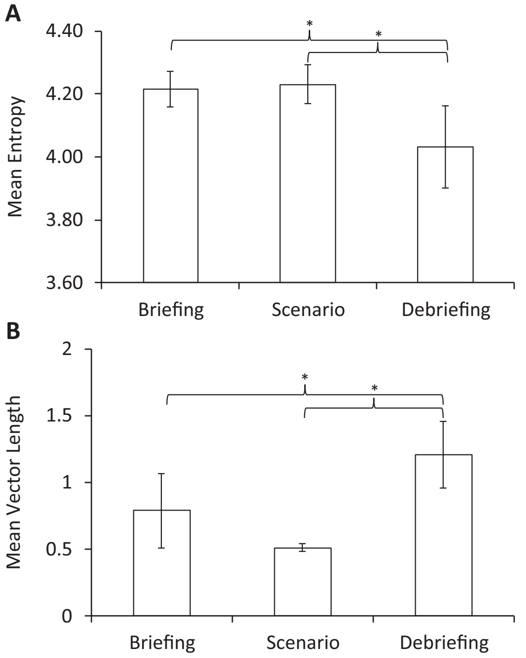

Figure 3. The pattern of significant differences between training segments was the same for both neural and communication variables: (a) mean neurodynamic entropy as a function of training segment and (b) mean vector length as a function of training segment. Error bars are 95% confidence intervals; asterisks indicate significant difference with $\alpha_{FW} = .05$.

Mutual Discrimination

In addressing our second research question, we predicted that if cross-level effects are present, then neural and communication variables should mutually discriminate between more and less experienced teams.

First, to discriminate between more versus less experienced teams using communication, we calculated the LSA cosine (semantic similarity) between all possible pairs of transcripts as a function of training segment and experience. We report the cosine matrix (correlation matrix) for scenario in Table 1 because it best discriminated between more and less experienced teams. If semantic relatedness discriminates between more and less experienced teams, then the bold values in Table 1 should be larger than the underlined values, as these correspond to within- versus between-group correlations. We analyzed these groupings using cluster analysis and multidimensional scaling (MDS).

Hierarchical clustering of the scenario cosine matrix (Table 1) revealed that more and less experienced teams clustered based on their communication differences (Figure 4a). The two-dimensional MDS solution (Stress = .86; R^2 = .96) similarly revealed that more and less experienced teams were positioned at opposite ends of an "experience" dimension (Figure 4b). The second MDS dimension discriminated teams along an as-yet-unidentified factor.

Because entropy time series depended on the exact amount of time teams performed in each training segment, the entropy time series for each team were of unequal lengths, which ruled out computing a correlation or distance matrix between all teams for cluster analysis and MDS. Therefore, to test whether neurodynamic entropy also discriminated between more and less experienced teams, we used discriminant function analysis to predict experience level (group membership) using mean entropy at briefing, scenario, debriefing, and overall (i.e., entropy over all training segments) as the predictors (i.e., discriminators). To find the ideal set of discriminators, we conducted a stepwise analysis.

The discriminant function with scenario and overall entropy as predictors was optimal, Λ = .15, $\chi^2(2) = 7.61$, $p = .022$, $\varphi^2 = .54$, such that all teams were correctly discriminated as more versus less experienced. The more and less experienced team bivariate means (scenario entropy, overall entropy) were 4.25, 4.24, and 4.20, 4.17, respectively, and their group centroids were 1.75 and −2.33, respectively. Hence, more experienced teams scored higher on this "amount-of-entropy" discriminant function than less experienced teams. This finding indicates that more experienced teams were more neurodynamically flexible than less experienced teams during scenario performance and overall, and that more and less experienced teams were correctly discriminated based on that difference in 100% of the cases.

Together, these results indicate that semantic similarity and neurodynamic entropy both discriminate between more and less experienced teams (mutual discrimination). Furthermore,

TABLE 1: Semantic Similarity (Cosine/Correlation) Matrix Computed Between All Pairs of Team Transcript During Scenario

	More 1	More 2	More 3	More 4	Less 1	Less 2	Less 3
More 1	—						
More 2	**.91**	—					
More 3	**.81**	**.85**	—				
More 4	**.85**	**.87**	**.82**	—			
Less 1	.82	.85	.78	.81	—		
Less 2	.76	.77	.72	.77	**.88**	—	
Less 3	.81	.81	.79	.84	**.85**	**.83**	—

Note. More = more experienced team; Less = less experienced team. Bold values indicate within-group cosines (correlations), and underlined values indicate between-group cosines (correlations).

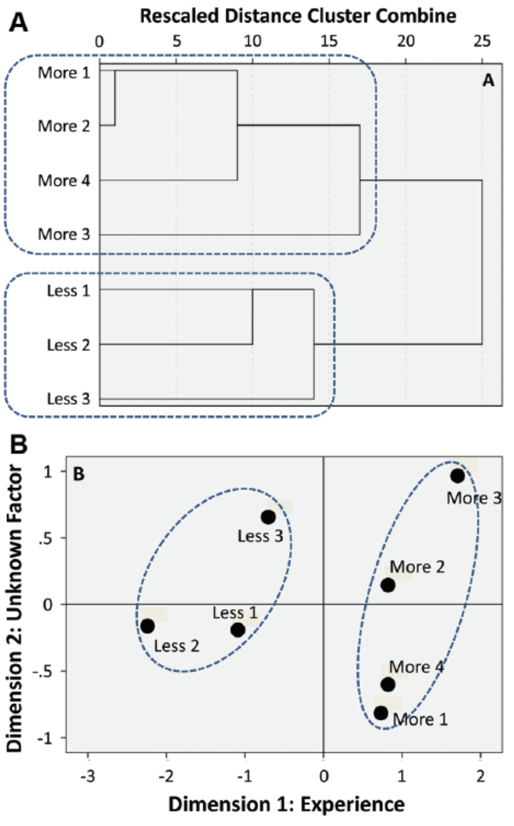

Figure 4. (a) Hierarchical cluster analysis of the latent semantic analysis semantic similarity (cosine) matrix between teams from the scenario training segment and (b) multidimensional scaling of this matrix for the scenario training segment. More = more experienced team; Less = less experienced team; dashed outlines indicate those groupings.

because mutual discrimination was specific only to the scenario segment, it suggests that mutual discrimination may be specific to more dynamic, real-time aspects of team performance, such as scenario performance.

Cross-Correlations

Our third research question was whether the development of cross-level effects can be observed through differences in lead-lag cross-correlations between neural and communication variables across more and less experienced teams.

We calculated lagged cross-correlation functions between LSA vector length of each utterance (Variable 1) and mean entropy during each utterance (Variable 2) for each transcript using the MatLab crosscorr function (Box et al., 1994; Figure 5). Which measure was assigned to be Variable 1 or Variable 2 was arbitrary; however, with the variable assignment that we used, a significant peak cross-correlation at a positive lag means that communication is leading and neurodynamics is lagging, whereas a significant peak cross-correlation at a negative lag means that neurodynamics is leading and communication is lagging. We calculated separate cross-correlation functions for each combination of training segment and experience. Number of utterances (N) determined the number of lags that could be analyzed ($N-1$) in each cross-correlation function. Because scenario contained the most utterances for all teams, there were more lags

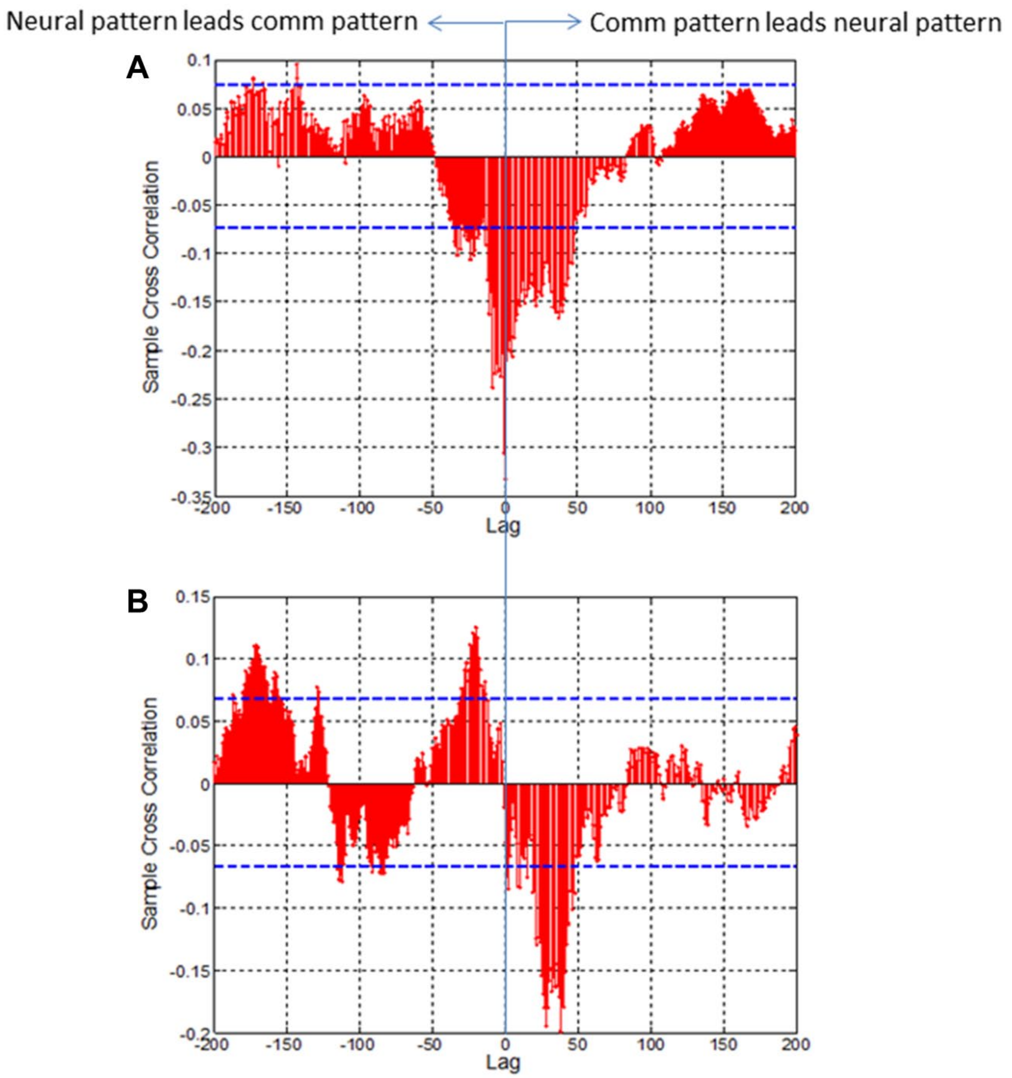

Figure 5. Cross-correlation functions for (a) a less experienced team and (b) a more experienced team computed over all training segments. The bold dashed lines represent 95% confidence intervals for zero correlation; if the correlation lies outside of those lines, then the correlation is significant at α = .05. As described in the text, peak cross-correlations at negative lags indicate that neurodynamics is leading, and peak cross-correlations at positive lags indicate that communication is leading.

for analyzing the scenario training segment. Peak cross-correlation was identified as the largest absolute correlation over all possible lags.

A peak cross-correlation that is significantly negative at a negative lag would mean that a more fixed neurophysiological distribution (low entropy) across team members tends to precede an increase in the amount of domain-specific content of utterances (large vector lengths). Conversely, a peak cross-correlation that is significantly positive at a positive lag would mean that increases in the amount of domain-specific content of utterances (large vector lengths) tend to precede a more flexible neurophysiological distribution (high entropy). A significant peak

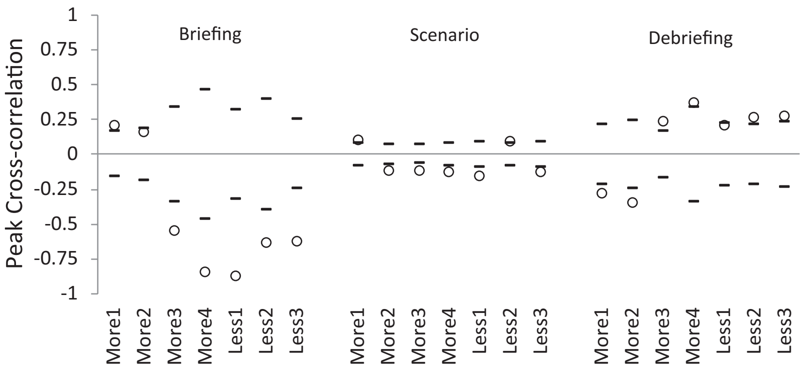

Figure 6. Peak cross-correlations (open circles) for all teams at each level of training segment. Hash marks indicate 95% confidence intervals for zero correlation; if the correlation lies outside the hash marks, then the correlation is significant at α = .05. More = more experienced team; Less = less experienced team.

cross-correlation at Lag 0 would mean that changes in neurophysiology and communication do not tend to precede or follow each other in time; it means that neurophysiology and communication are correlated only in the present. The meanings of other combinations of peak cross-correlation direction and lag can be inferred from these examples.

Figure 6 shows the direction and significance of peak cross-correlations at each level of training segment for more and less experienced teams. This figure indicates that cross-level effects were prevalent across all training segments and experience levels. We analyzed the absolute values (cf. effect size) and lag (i.e., whether neurophysiology or communication was leading) of these peak cross-correlations separately using 3 (training segment) × 2 (experience) mixed ANOVAs to determine the training segments where the strongest cross-level effects occurred and whether one level tended to lead the other as a function of experience.

Because peak cross-correlations could be either positive or negative (Figure 6), we analyzed absolute values to identify differences in the strength of peak cross-correlation between neural and communication variables. Only the training segment effect was significant, $F(2, 10) = 17.09$, $p = .001$, $\eta^2 = .77$. A Tukey test on training segment (α_{FW} = .05) revealed that the strongest peak cross-correlations occurred during briefing, followed by debriefing, and then scenario (Figure 7a). Although peak cross-correlations were at their strongest during briefing, peak cross-correlations and, hence, cross-level effects were found to be significant across all training segments (Figure 6).

Peak cross-correlation lag was analyzed to determine whether changes in communication pattern preceded changes in neurophysiological pattern or vice versa. The experience effect, $F(1, 5) = 6.83$, $p = .048$, $\eta^2 = .58$, and the Training Segment × Experience interaction, $F(1.05, 5.25) = 8.42$ (Greenhouse-Geisser correction used), $p = .031$, $\eta^2 = .63$, were both significant. The experience effect indicated that the peak cross-correlations of more experienced teams were positively lagged (communication leading), whereas the peak cross-correlations of less experienced teams were essentially Lag 0 (neither level leading). However, the interaction indicates that the experience effect depended on training segment.

As shown in Figure 7b, the experience effect was significant only during the scenario training segment, $F(1, 5) = 8.45$, $p = .033$, $\eta^2 = .63$. More experienced teams were significantly greater than Lag 0, $t(3) = 2.94$, $p = .03$ (one tailed), $d = 1.47$, and less experienced teams did not significantly differ from Lag 0, $t(2) = -1.29$, $p = .16$ (one tailed), $d = -0.74$, during the scenario training segment.

These results indicate that cross-level effects become "temporally extended," with change in

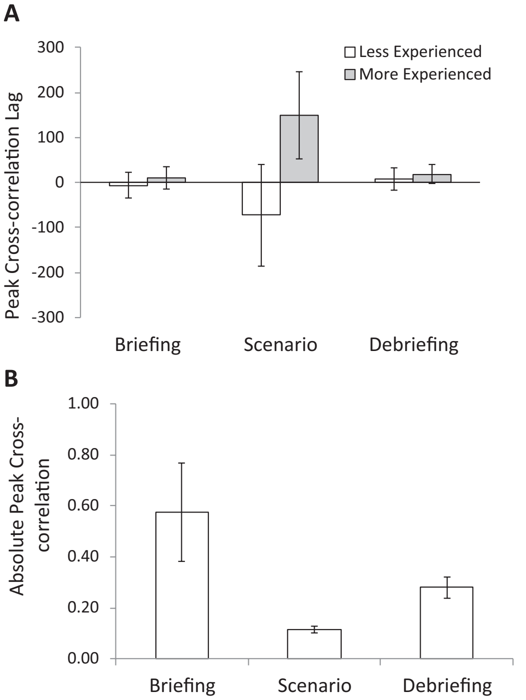

Figure 7. (a) Absolute strength of cross-level effects at each level of training segment and (b) the peak cross-correlation lag, which shows whether change in communication pattern precedes change in neurophysiological pattern (+ lag) or change in neurophysiological pattern precedes change in communication pattern (– lag) for more and less experienced teams at each level of training segment. Error bars are 95% confidence intervals.

communication pattern preceding change in neurophysiological pattern, in more experienced teams (see also Figure 5), but that effect is apparent only during the dynamic scenario training segment. By contrast, the peak cross-correlation for less experienced teams appears to be "temporally local" (i.e., correlated only in the present), such that although variations across neural and cognitive-behavioral levels are linked, neither level tends to lead or lag the other.

DISCUSSION

In this study, we found that concurrent changes across neural and communication variables as teams interact—cross-level effects—are subject to variations in team training and amount of team experience. In questioning whether cross-level effects exist and what drives them, we made predictions based on three research questions—concomitancy, mutual discrimination, and cross-correlation—which were each met with positive results.

Concomitancy

We hypothesized that if cross-level effects are present, then neural and communication variables should change together in response to changes in the training segments, which we termed *concomitancy*. Neurodynamic entropy was significantly lower and semantic content of utterances was significantly higher during debriefing compared to either the briefing or scenario training segment. Compared to scenario, debriefing is a relatively predictable and stable training segment that requires more open-ended discussions of previously encountered task elements and consensus-reaching processes. In contrast, scenario is more dynamic, requiring teams to coordinate new and evolving information as the situation unfolds. What these results show is that teams concomitantly alter their neural dynamics from more variable to more fixed and their communication content from terser and domain specific to lengthier and domain specific as they move from the scenario segment to the debriefing segment. We think that teams also changed their neural and communication patterns to match the dynamics of the briefing segment, but it is unclear at this point why those neural and communication patterns were not significantly different from the more dynamic scenario training segment. We suspect that it might be because, unlike the debriefing segment, both the briefing and scenario segments demanded a stricter division of labor and coordination of novel information.

We did not find concomitancy of team experience effects on neurodynamic entropy and semantic content, however. What this result suggests is that concomitancy may be a task-dependent cross-level effect: It is about changing the team neural and communication dynamics to match task dynamics, and to a degree, one should see this effect for any team, regardless of experience or skill level. This finding is consistent with the thesis that teams instinctively attempt to match their coordination dynamics to task dynamics regardless of their skill level (Gorman, Amazeen, & Cooke, 2010), which can be leveraged for training effective teams.

From a practical standpoint, task-dependent concomitancy means that if the objective of team training is the acquisition of both flexible neural and cognitive-behavioral processes, then more dynamic, scenario-based training rather than more retrospective, consensus-based training (e.g., debriefing) should be used. This finding conforms to the theory that flexible and adaptive team processes are induced by practicing in dynamic, unpredictable environments rather than by training on rote procedures or achieving consensus on shared knowledge (Gorman, Cooke, & Amazeen, 2010; Schollhorn et al., 2006) but extends that theory across neural and cognitive-behavioral levels of analysis.

Mutual Discrimination

We defined *mutual discrimination* as the capability of both neural and communication variables to discriminate between more and less experienced teams. Although we found mutual discrimination, it was centered exclusively on differences that occurred during the dynamic scenario training segment. This result suggests that cross-level effects are also experience (and presumably skill) dependent and that those differences become most apparent during real-time, dynamic task performance.

During the scenario (and overall), less experienced teams had a more fixed neurophysiological distribution (lower entropy) compared to more experienced teams (higher entropy), which indicates that more experienced teams were more neurally flexible. Scenario is the most dynamic training segment, where patterns of neural, cognitive, and behavioral activity must be flexible to adapt to changes in the task environment. The finding that more experienced teams were more neurally flexible during this dynamic training segment is consistent with the thesis that although all teams attempt to adapt, more experienced, skilled teams are more responsive in adapting their coordination dynamics to keep pace with changing task dynamics (Gorman, Amazeen, et al., 2010; Stevens, Gorman, et al., 2012). Also during the scenario, more experienced teams' communication was more similar to each other than to less experienced teams, which replicates prior research that discriminated between skilled and unskilled UAV teams (e.g., Gorman et al., 2003; Martin & Foltz, 2004). From a communication perspective, this finding further demonstrates that LSA is an effective diagnostic tool that generally discriminates more experienced from less experienced teams based on how they communicate during task performance.

That mutual discrimination between more and less experienced teams was found only during the scenario is consistent with the theory that team processes that account for differences in team effectiveness are most apparent during dynamic task performance. Specifically, this finding is aligned with interactive team cognition theory, which claims that team cognition is not contained separately in the heads of team members but is directly embedded in their interactions during dynamic task performance (Cooke et al., 2013; see also De Jaegher, 2009). Whereas the concomitancy result demonstrates that changes in training segment from planning (i.e., briefing) to task performance (i.e., scenario) to after-action review (i.e., debriefing) concomitantly modulate neural and cognitive-behavioral patterns, the mutual-discrimination finding suggests that the way teams respond to more dynamic, real-time tasks, such as the scenario training segment, is what really separates experienced from inexperienced teams along neural and cognitive-behavioral dimensions of teamwork.

Discriminating teams in terms of experience (or skill) level in the context of real-time, dynamic team interaction is a critical need in work domains such as emergency medicine (Shapiro et al., 2008) and cybersecurity (Rajivan, Janssen, & Cooke, 2013). In such domains, there is a need to assess real-time team member interactions in order to not miss out on the team processes underlying team effectiveness (Wildman, Salas, & Scott, 2014). Communication content analysis is one approach for doing so, but it is resource-intensive and time-consuming and generally must be performed by communication analysis experts (Emmert & Barker, 1989). Our mutual-discrimination results suggest that neural and communications variables may provide interchangeable (or at least complementary) metrics for discriminating team skill level. Capitalizing on this characteristic of cross-level effects, eventually scenario-based training could utilize real-time neurophysiological observations as a process-based measure of team experience and skill level while reducing the need for extensive post hoc communication analysis. For example, automated neural metrics could be monitored by machines, and more in-depth communication analysis could be performed by human experts whenever anomalies or other critical events are detected in the neural signals.

Cross-Correlations

To better determine how levels become linked in more and less experienced teams, we analyzed peak cross-correlations between neurodynamic entropy and semantic content. Of the 21 training segments that we analyzed across the more and less experienced teams, 19 exhibited significant peak cross-correlations (Figure 6). Of these significant peak cross-correlations, 14 of 19 were negative (binomial p = .03), suggesting that in general, as teams become more neurally flexible (higher entropy), their utterances become terser and more efficiently packed with domain-specific content (smaller vector lengths). Shorter communication patterns and more efficient ("low-overhead") communication are associated with behavioral flexibility in adaptive teams (Gorman, Cooke, Amazeen, & Fouse, 2012; MacMillan, Entin, & Serfaty, 2004). Therefore, this result suggests that neural flexibility may be linked to cognitive-behavioral flexibility in adaptive teams.

Although there were significant peak cross-correlations in all training segments, the scenario segment resulted in the most compelling differences between more and less experienced teams. More experienced teams' peak cross-correlations were positively lagged, with change in communication pattern portending change in neurophysiological pattern, whereas less experienced teams' peak cross-correlations were essentially zero lagged (i.e., neither level leading). Therefore, more experienced teams' cross-level effects were temporally extended (extending into the future and past of team performance), whereas less experienced teams' cross-level effects were temporally local (present only in the "here and now"). With greater experience, neural and cognitive-behavioral levels appear to become temporally intertwined with one another in more complex ways.

Temporally extended effects present a unique challenge for modeling human performance. As people establish a history working together as a team, team processes may become more and more embedded in the social history of the team over hours, days, and weeks, moving beyond timescales of deliberate human action (i.e., milliseconds to hours; see Newell, 1990, for timescales of deliberate human action). Accordingly, the level of explanation must also move beyond individual-level cognitive constructs defined at the level of deliberate human action, such as mental models, scripts, and schemas (e.g., the individual-level "inputs" in linear input→processing→output models; Ilgen, Hollenbeck, Johnson, & Jundt, 2005) to capture temporally extended team processes.

A dynamical systems approach (Gorman, Amazeen, et al., 2010) may be more appropriate because the level of explanation focuses not on individual-level constructs defined at the level of deliberate human action but on emergent relations across all timescales of team interaction, including those extending over hours, days, and weeks. For example, we have successfully used dynamical systems approaches in the past to model temporally extended team effects separately at the neural and communication levels using multifractal analysis (Likens et al., 2014) and attractor reconstruction (Gorman, Amazeen, et al., 2010). However, this type of modeling has yet to be applied to the phenomenon of cross-level effects.

In this study, we explored potential causal directions between neural and communication variables by examining the lead-lag nature of peak cross-correlations. In contrast to the idea that team performance might follow a causal arrow from neural to cognitive-behavioral (neural→communications; Frith, 2007), our results indicate that changes in cognitive-behavioral constraints (such as communication patterns) tend to precede changes in neural patterns as teams gain experience (Fuchs & De Jaegher, 2009). We think that as teams gain history and experience, an interactive context of constraints emerges, compelling individual thoughts and actions to unfold in particular ways (Gorman, 2014). Extending this idea to cross-level effects, once cognitive-behavioral constraints are established, for example, through highly evolved communication patterns, they begin to structure neural patterns within and across team members. This effect is similar to the idea of the development of individual knowledge through conversation and dialog (e.g., common ground; Clark & Brennan, 1991; see also Bakhtin, 1986), but here communication patterns provide constraints under which individual neural patterns fluctuate and vary. A practical implication of this result is that by altering communication patterns and team interaction constraints during training, we may be able to drive changes at both the individual and team neural levels.

Limitations and Future Directions

A methodological challenge in studying cross-level effects is the need to synchronize measurements across different levels of analysis. Although we measured team neurophysiology and communication using established methods, those methods do not naturally share a common sampling interval (i.e., communication content was measured utterance by utterance, whereas entropy was measured second by second). We aligned our data post hoc by computing mean entropy on an utterance-by-utterance basis (i.e., by downsampling entropy), which artificially reduced the amount of variability (information) in the entropy measurements. Authors of future research should address this issue by developing more directly matched sampling intervals between neural and communication measurements to maximize the amount of overlapping information for testing cross-level effects.

In this study, we used amount of team experience as a surrogate for team skill level. However, more direct measures of team skill—rate of task performance, accuracy, and so on—are needed to further validate cross-level effects. In the future, using a synthetic task environment with built-in objective performance metrics (e.g., Cooke & Shope, 2005) to examine cross-level effects could help address this issue.

In this study, amount of team experience included both the amount of submarine navigation experience (task familiarity) and the amount of experience working together as a team (team member familiarity). Therefore, we were unable to determine how each of these two team experience factors uniquely contributed to cross-level effects. In future research, it will be important to disentangle these factors in order to determine whether task familiarity or team member familiarity contributes more to the development cross-level effects.

We examined neural and communication variables in SPAN teams, but cross-level effects could be investigated across other levels of analysis—such as physiological (e.g., respiratory effort), cognitive (e.g., mental models), and behavioral (e.g., kinematics)—in other work domains. For example, cross-level effects between cognition and kinematics could be important for understanding system performance in human–robot interaction (De Santis, Siciliano, De Luca, & Bicchi, 2007). Although our research with SPAN teams represents just one possibility, our results suggest that cross-level effects may be promising for understanding team skill development in other work domains.

The strong emergence of cross-level effects during dynamic scenario performance highlights the importance of real-time team interaction for understanding the neural and cognitive-behavioral underpinnings of team performance. However, we do not wish to discount the important role of debriefing and after-action review in simulation-based learning. Learning requires feedback, and debriefing is a form of feedback that is used in a variety of military, industrial, and medical settings (Fanning & Gabba, 2007). That cross-level effects were less apparent during debriefing suggests that cross-level effects may not be as diagnostic of learning during feedback phases of training but that they may be more useful for assessing team skill development during dynamic scenario performances.

Do changes in communication patterns really cause changes in neural patterns? Although we observed that change in communication pattern portended change in neural pattern in more experienced teams, our results are correlational, and more research is needed to understand the causal relationships underlying cross-level effects. Such an understanding could be gained by experimentally inducing change on one level (e.g., neural) and then observing change on the other level (e.g., communication) (neural→communication) and vice versa (communication→neural), such that causal direction would serve as an independent variable in the experiment. In this way, whether a "causal hierarchy" exists between levels, where effects in one direction are stronger than effects in the other direction, could be determined. Resolving this issue is important for determining whether one level (e.g., communication) is more causal fundamentally than the other (e.g., neural) during team development and, therefore, should be the focus of team training and assessment. Of course, if no causal hierarchy is found, then neural and cognitive-behavioral processes would develop in a reciprocal fashion, and team training and assessment should place equal emphasis on both levels of analysis.

Finally, this study revealed that cross-level effects are subject to variations in task dynamics, such as differences in training segments (e.g., scenario vs. debriefing) and amount of team experience; however, to provide a more complete picture of cross-level effects, other variables relevant to team performance, such as shared-mental-model emergence (DeChurch & Mesmer-Magnus, 2010; Kozlowski & Klein, 2000) should be examined. For example, shared mental models are thought to reduce the need to communicate (Entin & Serfaty, 1999; MacMillan et al., 2004). Therefore, given our current results, emergent shared mental models could moderate cross-level effects by reducing the amount of communication overhead, which should in turn affect neural flexibility.

CONCLUSION

Our results indicate that teamwork is not reducible to a fundamental level of analysis (e.g., neural or cognitive-behavioral) but that training effects are spread out across multiple levels and timescales of analysis and are manifested in cross-level effects. Cross-level effects suggest that neural and cognitive-behavioral processes might be different faces of a unitary coordination process rather than separable teamwork dimensions and that linkages between levels are established in temporally complex ways as teams gain experience.

In this light, cross-level effects suggest that different lines of team research that currently focus on a single level of analysis (e.g., neural processes, communication processes) might be focusing in on the same phenomenon, just using different methodological "lenses," and that one key to an integrated picture may be expanding the analysis of team performance to incorporate team processes operating across different levels and timescales of analysis. Ultimately, understanding how cross-level effects develop as teams gain experience may lead to new forms of team skill assessment and new theories about what develops during team skill acquisition.

ACKNOWLEDGMENTS

Preliminary findings of this research were reported at the 2013 Human–Computer Interaction international conference (Gorman, Martin, Dunbar, Stevens, & Galloway, 2013). This research was supported by Defense Advanced Projects Agency Contract W31P4Q-12-C-0166 and National Science Foundation SBIR Grant IIP 121215327. The study protocol was approved by the Naval Submarine Medical Research Laboratory Institutional Review Board in compliance with all applicable federal regulations governing the protection of human subjects.

KEY POINTS

- Concurrent changes across neural and communication levels of analysis, which we call *cross-level effects*, are subject to variations in team training and amount of team experience.
- Neural and communication variables change together in response to changes in training segments (briefing, scenario, debriefing), and neural and communication variables mutually discriminate between teams with different experience levels.
- Cross-level effects become more complex and temporally extended in more experienced teams, and we found evidence that changes in communication pattern portend changes in neural pattern in more experienced teams.
- Cross-level effects could provide multiple routes for assessing team training effectiveness and help consolidate theories that currently focus on

different (e.g., neural, cognitive-behavioral) levels of analysis.

REFERENCES

Bakhtin, M. M. (1986). *Speech genres and other late essays* (C. Emerson & M. Holquist, Eds.; V. W. McGee, Trans.). Austin: University of Texas Press.

Berka, C., Levendowski, D. J., Cvetinovic, M. M., Petrovic, M. M., Davis, G., Lumicao, M. N., Zivkovic, V. T., Popovic, M. V., & Olmstead, R. (2004). Real-time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset. *International Journal of Human–Computer Interaction*, *17*, 151–170.

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., Tremoulet, P. D., & Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine*, *78*, B231–B244.

Box, G., Jenkins, G. M., & Reinsel, G. (1994). *Time series analysis: Forecasting and control*. Upper Saddle River, NJ: Prentice Hall.

Cannon-Bowers, J. A., & Salas, E. (2001). Reflections on shared cognition. *Journal of Organizational Behavior*, *22*, 195–202.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on Socially Shared Cognition*, *13*, 127–149.

Cooke, N. J., & Gorman, J. C. (2009). Interaction-based measures of cognitive systems. *Journal of Cognitive Engineering and Decision Making*, *3*, 27–46.

Cooke, N. J., Gorman, J. C., Kiekel, P. A., Foltz, P., & Martin, M. (2005). *Using team communication to understand team cognition in distributed vs. co-located mission environments* (Technical Report for Office of Naval Research Grant N00014-03-1-0580). Arlington, VA: Office of Naval Research.

Cooke, N. J., Gorman, J. C., Meyers, C. W., & Duran, J. L. (2013). Interactive team cognition. *Cognitive Science*, *37*, 255–285.

Cooke, N. J., & Shope, S. M. (2005). Synthetic task environments for teams: CERTT's UAV-STE. In N. A. Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. W. Hendrick (Eds.), *Handbook on human factors and ergonomics methods* (pp. 46-1–46-6). Boca Raton, FL: CRC Press.

De Jaegher, H. (2009). Social understanding through direct perception? Yes, by interacting. *Consciousness and Cognition*, *18*, 535–542.

De Santis, A., Siciliano, B., De Luca, A., & Bicchi, A. (2007). An atlas of physical human–robot interaction. *Mechanism and Machine Theory*, *43*, 253–270.

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). Measuring shared team mental models: A meta-analysis. *Group Dynamics: Theory, Research, and Practice*, *14*, 1–14.

Dong, A. (2005). The latent semantic approach to studying design team communication. *Design Studies*, *26*, 445–461.

Emmert, P., & Barker, L. L. (1989). *Measurement of communication behavior*. White Plains, NY: Longman.

Entin, E. E., & Serfaty, D. (1999). Adaptive team coordination. *Human Factors*, *41*, 312–325.

Fanning, R. M., & Gabba, D. M. (2007). The role of debriefing in simulation-based learning. *Simulation in Healthcare*, *2*, 115–125.

Foltz, P. W., & Martin, M. J. (2009). Automated communication analysis of teams. In E. Salas, J. Goodwin, & C. S. Burke (Eds.), *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches* (pp. 411–431). SIOP Organizational Frontiers Series. New York, NY: Routledge.

Foltz, P. W., Martin, M. J., Abdelali, A., Rosenstein, M., & Oberbreckling, R. (2006). Automated team discourse modeling: Test of performance and generalization. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 1317–1322). Mahwah, NJ: Lawrence Erlbaum.

Frith, C. D. (2007). The social brain? *Philosophical Transactions of the Royal Society*, *362*, 671–678.

Frith, C. D., & Wolpert, D. M. (Eds.). (2004). *The neuroscience of social interaction: Decoding, imitating, and influencing the actions of others*. New York, NY: Oxford University Press.

Fuchs, T., & De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences*, *8*, 465–486.

Gorman, J. C. (2014). Team coordination and dynamics: Two central issues. *Current Directions in Psychological Science*, *23*, 355–360.

Gorman, J. C., Amazeen, P. G., & Cooke, N. J. (2010). Team coordination dynamics. *Nonlinear Dynamics, Psychology, & Life Sciences*, *14*, 265–289.

Gorman, J. C., & Cooke, N. J. (2011). Changes in team cognition after a retention interval: The benefits of mixing it up. *Journal of Experimental Psychology: Applied*, *17*, 303–319.

Gorman, J. C., Cooke, N. J., & Amazeen, P. G. (2010). Training adaptive teams. *Human Factors*, *52*, 295–307.

Gorman, J. C., Cooke, N. J., Amazeen, P. G., & Fouse, S. (2012). Measuring patterns in team interaction sequences using a discrete recurrence approach. *Human Factors*, *54*, 503–517.

Gorman, J. C., & Crites, M. J. (2013). Are two hands (from different people) better than one? Mode effects and differential transfer between manual coordination modes. *Human Factors*, *55*, 815–829.

Gorman, J. C., & Crites, M. J. (2015). Learning to tie well with others: Bimanual versus intermanual performance of a highly practised skill. *Ergonomics*, *58*, 680–697.

Gorman, J. C., Foltz, P. W., Kiekel, P. A., Martin, M. J., & Cooke, N. J. (2003). Evaluation of latent semantic analysis–based measures of team communications content. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting* (pp. 424–428). Santa Monica, CA: Human Factors and Ergonomics Society.

Gorman, J. C., Martin, M. J., Dunbar, T. A., Stevens, R. H., & Galloway, T. (2013). Analysis of semantic content and its relation to team neurophysiology during submarine crew training. *Foundations of Augmented Cognition: Lecture Notes in Computer Science*, *8027*, 143–152.

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology*, *56*, 517–543.

Islam, M. M., & Latiful Hoque, A. S. M. (2010). Automated essay scoring using generalized latent semantic analysis. In *Proceedings of the 13th International Conference on Computer and Information Technology* (pp. 358–363). Piscataway, NJ: IEEE.

Johnson, R. R., Popovic, D. P., Olmstead, R. E., Stikic, M., Levendowski, D. J., & Berka, C. (2011). Drowsiness/alertness algorithm development and validation using synchronized EEG and cognitive performance to individualize a generalized model. *Biological Psychology*, *87*, 241–250.

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent properties. In K. L. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 3–90). San Francisco, CA: Jossey-Bass.

Krieger, A. C., Ayappa, I., Norman, R. G., Rapoport, D. N., & Walsleben, J. (2004). Comparison of the maintenance of wakefulness test

(MWT) to a modified behavioral test (OSLER) in the evaluation of daytime sleepiness. *Journal of Sleep Research*, *13*, 407–411.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes*, *25*, 259–284.

Likens, A. D., Amazeen, P. G., Stevens, R., Galloway, T., & Gorman, J. C. (2014). Neural signatures of team coordination are revealed by multifractal analysis. *Social Neuroscience*, *9*, 219–234.

MacMillan, J., Entin, E. E., & Serfaty, D. (2004). Communication overhead: The hidden cost of team cognition. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance* (pp. 61–82). Washington, DC: American Psychological Association.

Martin, M. J., & Foltz, P. W. (2004). Automated team discourse annotation and performance prediction using LSA. In *Proceedings of the Human Language Technology and North American Association for Computational Linguistics Conference* (pp. 97–100). Boston, MA: North American Association for Computational Linguistics.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Rajivan, P., Janssen, M. A., & Cooke, N. J. (2013). Agent-based model of a cyber security defense analyst team. In *Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting* (pp. 314–318). Santa Monica, CA: Human Factors and Ergonomics Society.

Schollhorn, W. I., Beckmann, H., Michelbrink, M., Sechelmann, M., Trockel, M., & Davids, K. (2006). Does noise provide the basis for the unification of motor learning theories? *International Journal of Sport Psychology*, *37*, 186–206.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Science*, *10*, 70–76.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.

Shapiro, M. J., Gardner, R., Godwin, S. A., Jay, G. D., Lindquist, D. G., Salisbury, M. L., & Salas, E. (2008). Defining team performance for simulation-based training: Methodology, metrics, and opportunities for emergency medicine. *Academic Emergency Medicine*, *15*, 1088–1097.

Stevens, R. H., & Galloway, T. (2014). Toward a quantitative description of the neurodynamics organizations of teams. *Social Neuroscience*, *9*, 160–173.

Stevens, R. H., Galloway, T., & Lamb, C. (2014). Submarine navigation team resilience: Linking EEG and behavioral models. In *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting* (pp. 245–249). Santa Monica, CA: Human Factors and Ergonomics Society.

Stevens, R. H., Galloway, T., Wang, P., & Berka, C. (2012). Cognitive neurophysiologic synchronies: What can they contribute to the study of teamwork? *Human Factors*, *54*, 489–502.

Stevens, R. H., Gorman, J. C., Amazeen, P. G., Likens, A., & Galloway, T. (2012). The organizational neurodynamics of teams. *Nonlinear Dynamics, Psychology, & Life Sciences*, *17*, 67–86.

Wildman, J. L., Salas, E., & Scott, C. P. R. (2014). Measuring cognition in teams: A cross-domain review. *Human Factors*, *56*, 911–941.

Jamie C. Gorman received his PhD in psychology from New Mexico State University in 2006 and is an associate professor in engineering psychology at the Georgia Institute of Technology.

Melanie J. Martin received her PhD in computer science from New Mexico State University in 2005 and is an associate professor in the Department of Computer Science at California State University, Stanislaus.

Terri A. Dunbar received her BS in psychology from the University of Utah in 2012 and is a PhD student in engineering psychology at the Georgia Institute of Technology.

Ronald H. Stevens received his PhD in molecular genetics from Harvard University in 1971 and is a professor at University of California–Los Angeles School of Medicine and CEO of the Learning Chameleon, Inc.

Trysha L. Galloway received her CPFDA, EFDA, and CDA (including radiation health and safety [RSH], infection control [ICE], and national-level general chairside [GC]) from Oregon Health and Science University in 1995 and is the director of EEG research at the Learning Chameleon, Inc.

Polemnia G. Amazeen received her PhD in psychology from the University of Connecticut in 1996 and is an associate professor of psychology at Arizona State University.

Aaron D. Likens received his MA at the University of Central Oklahoma in 2010 and is a PhD student at Arizona State University.

Date received: September 2, 2014
Date accepted: July 28, 2015