Assessing Submariners' Intuitive Decision-Making Skills Using Neurocognitive Methods

Lisa C. Lucia, Ph.D.
Jeffrey M. Beaubien, Ph.D.
Aptima, Inc., Woburn, MA
llucia@aptima.com,
jbeaubien@aptima.com

Natalie Steinhauser
Naval Air Warfare Center Training Systems Division
Orlando, FL
Natalie.Steinhauser@navy.mil

CAPT Ronald Steed (USN, ret.)
UpScope Consulting Group, LLC
Mystic, CT
ronaldsteed@upscopeconsulting.com

ABSTRACT

Military missions pose complex cognitive and perceptual challenges, such as detecting potential improvised explosive devices along the roadside, or detecting anomalous social cues in a crowded market that may suggest an impending attack. Situations such as these do not allow time for extensive deliberation; rather, the ability to make quick and accurate decisions is key to survival. Evans and Stanovich (2013) suggest that, unlike deliberate decision-making, this type of intuitive decision-making is extremely fast and requires little or no working memory. Research has shed light on the neural mechanisms underlying intuitive decision-making. For example, Luu and colleagues (2010) used electroencephalography (EEG)-based techniques to identify a neural signal of intuition during an object detection task. In their study, participants viewed fragmented line drawings and indicated whether each image contained a real object. Participants' event-related potentials (ERPs) differed between correctly identified real objects and correctly rejected non-objects after ~200 milliseconds (ms) and this difference persisted through ~500 ms. The purpose of the current study was to examine the generalizability of this neural signal with a military sample performing both everyday decision tasks (object detection) and military-relevant tasks (course safety decisions). Twenty-seven submariners participated in a rigorously-controlled, within-subjects experiment. Statistical analyses of the participants' brain activity confirmed that the neural signal identified by Luu and colleagues generalizes across tasks. Moreover, this signal reliably differentiates expert submariners from novices. This study is unique in that it validated the existence of a neural indicator of accurate intuitive decision-making across both samples and tasks.

ABOUT THE AUTHORS

Lisa C. Lucia, Ph.D. is a Scientist at Aptima, Inc., where she leads projects that involve the development of neuroscience-based training tools, decision support tools, and medical informatics systems. Her research interests span from the brain-bases of human perception and memory to visuospatial abilities and skill learning. She holds a Ph.D. in Cognitive Neuroscience from Tufts University, and a B.S. in Biological Psychology from Bates College.

Jeffrey M. Beaubien, Ph.D. is a Principal Scientist at Aptima, Inc., where he leads projects on training and human performance assessment. His research interests include team dynamics and decision-making. He holds a Ph.D. in Industrial and Organizational Psychology from George Mason University, a M.A. in Industrial and Organizational Psychology from the University of New Haven, and a B.A. in Psychology from the University of Rhode Island.

Natalie Steinhauser is a Senior Research Psychologist at the Naval Air Warfare Center Training Systems Division (NAWCTSD) Basic and Applied Training and Technology for Learning and Evaluation (BATTLE) laboratory. She has 11 years of Navy experience conducting training effectiveness evaluations, developing intelligent tutoring systems, and conducting decision-making research. She holds a M.S. in Modeling and Simulation from the University of Central Florida, and a B.S. in Human Factors/Applied Psychology from Embry-Riddle Aeronautical University.

CAPT Ronald Steed (USN, ret.) is the Principal at UpScopeConsulting Group, LLC where he serves on a variety of research and development projects involving command decision-making, submarine systems, team dynamics, and human factors. Mr. Steed transitioned from the U.S. Navy's Submarine Force in 2007. His last assignment was as the Commander of Submarine Squadron Two in Groton, CT. Captain Steed holds a M.S. in Engineering Management from Old Dominion University, and a B.S. in Electrical Engineering from The Citadel.

Assessing Submariners' Intuitive Decision-Making Skills Using Neurocognitive Methods

Lisa C. Lucia, Ph.D.
Jeffrey M. Beaubien, Ph.D.
Aptima, Inc., Woburn, MA
llucia@aptima.com,
jbeaubien@aptima.com

Natalie Steinhauser
Naval Air Warfare Center Training Systems Division
Orlando, FL
Natalie.Steinhauser@navy.mil

CAPT Ronald Steed (USN, ret.)
UpScope Consulting Group, LLC
Mystic, CT
ronaldsteed@upscopeconsulting.com

BACKGROUND

Military missions often involve uncertainty, which poses complex perceptual and cognitive challenges for the warfighter. For example, warfighters must continually update their situation awareness, rapidly detect the onset of critical cues (e.g., potential improvised explosive devices along the roadside, anomalous social cues in a crowd that might suggest an impending attack, or the presence of surface vessels that obstruct a submarine's path to periscope depth), and respond appropriately without hesitation. These situations do not allow time for extensive deliberation; as such, intuitive decision-making is key to survival. The need to make accurate intuitive decisions is not unique to the military. It is critical to effective performance in law enforcement, disaster response, emergency medicine, surgery, and aviation, among others. Thus, it is important to study and better understand this phenomenon.

The dual-process theory of decision-making (Evans & Stanovich, 2013) postulates that there are two brain processes which operate in parallel. One of these processes, "Type 1," operates at the unconscious level. This type of decision-making is extremely fast, makes minimal demands on working memory, and operates by automatically and associatively comparing (i.e., pattern matching) the current situation to one's corpus of accumulated prior experiences in long-term memory. All humans engage in a considerable amount of Type 1 processing in their day-to-day lives, for example, identifying everyday objects such as chairs and tables. In addition, experts often have well-developed domain-specific Type 1 skills. For example, a chess master can quickly look at a populated chess board, determine each player's strategy, and project the next two or three best moves – all within a matter of seconds. By comparison, "Type 2" processing operates the conscious level. It is much slower, requires active attention, places heavy demands on working memory, and makes decisions based on calculations and deliberation. It is akin to the slow and deliberate decision-making approach used by novices, as well as the slow and deliberate approach used by experts when facing novel problems or situations (Kahneman & Klein, 2009). For example, a novice chess player must take time to carefully study the two players' positions, deliberately reflect on their relevant strategies, and consciously recall the relevant decision rules before being able to predict the next move. Some researchers have suggested that Type 1 processes' speed advantage is rooted in the ability to subconsciously extract statistical covariation from the environment through experience (Evans, 2010; Reber, 2013). Specifically, experts become adept at leveraging information from long-term memory to selectively direct the focus of attention onto those environmental cues which are likely to be most informative or diagnostic.

The fields of perceptual gestalt psychology and cognitive neuroscience provide some clues about the timing of intuitive decision-making. The central principle of gestalt psychology is that the whole is conceptually understood as different from the collected interpretations of its component parts (Kofka, 1922). To test how this perceptual process may lead to intuitive decisions, researchers have studied the ability to perceive coherent objects from images that are composed of line segments (Luu et al., 2010). When some of the line segments are removed from the image, it becomes difficult to interpret. These images are called "fragmented real objects." For other images, the line segments are randomly scrambled. As a result, the image no longer depicts the original object. These images are called "fragmented non-objects." Random sets of images are then presented to the participant for extremely brief periods of time (e.g., 400 milliseconds [ms]) and the participant is encouraged to report perceptions of object wholeness, regardless of their ability to verbalize the object's name. This method forces individuals to make quick, intuitive judgments about what they just observed. In theory, participants with normal vision should perform well on this task because intuitive decision processes are heavily dependent on expertise, and sighted humans are experts in the visual identification of everyday objects such as chairs, tables, and cars.

Luu and colleagues (2010) used this method along with electroencephalography (EEG) recording techniques to generate event-related potentials (ERPs) – an average of neural activity across trials – to better understand the millisecond-by-millisecond timing of intuitive decision-making. Specifically, they compared the ERPs between correctly indicated real objects (for which meaningful gestalts may have been created) and correctly rejected non-objects (which likely did not lead to an internalized gestalt representation). Brain activity was similar between these two conditions until about 200 ms, and then again after 500-600 ms. However, during the intermediate time frame, neural activity was more negative in voltage for the non-objects. This difference indicates that either different underlying brain areas were activated, or similar brain areas were activated but to different extents. These activity differences also likely reflect differences in internalized representations (gestalt achieved vs. no gestalt achieved) during Type 1 processing.

Objective of Current Study

In the present study, we investigated the generalizability of these findings to a different participant sample and to a different task. First, to determine whether Luu et al.'s (2010) findings generalize to a different sample, we replicated their study with submariners. We sought to answer the following questions during the Object Block: *Can submariners make intuitive decisions about degraded line drawings with the same level of performance accuracy as college students?*, *Does performance vary according to submarine-relevant skills?*, and *Do neural activation patterns reflect the underlying cognitive processes that contribute to accurate intuition?*

Second, we sought to determine whether Luu et al.'s (2010) findings generalize to a different task, and whether the findings vary as a function of domain-specific expertise. In the submarine domain, it is important for submarines at periscope depth to remain a safe distance from other vessels. The submarine crew must maneuver the very slow-moving boat as soon as it is apparent that another vessel's track will likely intersect the submarine's track. Time is critical because almost all vessels move considerably faster than a submarine at periscope depth. Therefore, without immediate action, an unsafe situation can quickly develop. The periscope operator's rapid and accurate visual assessment of another vessel's direction and their simultaneous decision regarding the safety of the current course provides the first indication of a potential safety threat. Expert submariners have characterized this assessment and decision response as involving an immediate "feeling of knowing" (i.e., intuition). Using an experimental task which elicits the type of intuitive response common to the periscope operator, we sought to answer the following questions during the Periscope Block: *Can submariners make accurate intuitive decisions about course safety based on a quick glimpse through the periscope?*, *If so, how does their performance on this domain-specific task compare to that from the Object Block?*, *Does performance vary according to submarine-relevant skills?*, and *Do neural activation patterns reflect the underlying cognitive processes that contribute to accurate intuition?*

All 27 participants completed both task blocks, the order of which was randomized and counterbalanced. Based on their performance on a periscope operator skills test, participants were identified as being either more (expert) or less skilled (novice) with regard to periscope operations. We hypothesized that both groups would perform equally well on the Object Block, and would have similar performance to those of the college students in the Luu et al. (2010) study. In this way, the Object Block served as our manipulation check. We further hypothesized that experts would perform faster, more accurately, or both compared to novices on the submarine-relevant task. Regarding the neural substrate of intuition, we hypothesized that there would be no group differences during the timeframe of interest (~200-500 ms) on the Object Block. By contrast, on the Periscope Block, we hypothesized that there would be significant neural activation differences between expertise groups during the timeframe of interest. We further hypothesized that neural activity (between ~200-500 ms) would be more negative in voltage in the novice group.

METHOD

Participants

The participants (n = 27) included a mix of active duty and recently-retired submariners[1] from the New London, CT area. The active duty participants were students and instructors from the Naval Submarine School (NSS) Submarine

[1] The sample included a mix of officers, officers-in-training, and enlisted personnel with experience as a junior officer, department head, executive officer, commanding officer, or assistant navigator. All of these positions involve periscope operation as a part of their duties.

Officer Advanced Course (SOAC); they were recruited with assistance from the Submarine Learning Center (SLC). The retired participants were recruited from a local consulting firm. The sample ranged in age from 26-53 years (M = 38, SD = 8.9). All were right-handed, male, native English-speakers with normal or corrected-to-normal vision. They had no history of head injuries, nor were they taking medications that would otherwise affect their EEG recordings.

Procedure

Upon arrival at the study location, the investigator explained the tasks to the participants and answered all of their questions. Written informed consent was obtained using a protocol that was approved by the Office of Naval Research (ONR). Next, the participants completed a brief demographics questionnaire, were fitted with a Quick-20 Dry EEG Headset (Cognionics, Inc.), and were seated 65 cm in front of a 22" LCD computer monitor. The participants began each task block (Object Block, Periscope Block) with a small set of practice trials. None of the images used during practice were repeated in the actual task blocks. All participants then completed both the Object Block and the Periscope Block, with a brief rest in between each block. Block order was randomized and counterbalanced to minimize fatigue and task interference effects. Following completion of the second task block, the EEG headset was removed, and participants completed the periscope operator skills test. Finally, the participants completed a post-session survey. The entire session lasted approximately 90 minutes.

Periscope Operator Skills Test

To assign participants to expertise categories, all participants performed a ~10-minute assessment of their angle on the bow (AOB) judgment skill using the pre/post-test component of the Periscope Operator Adaptive Training tool (POAT; Landsberg et al., 2012). This test was especially useful in grouping participants based on their relevant skill because demographic variables (e.g., years of submariner training) did not reveal any performance differences in the Periscope Block. This test was considered a valid assessment tool because it is a subcomponent of the decision-making task performed in the Periscope Block, and; it measures a submarine-specific skill that is known to degrade quickly without practice. The test involves viewing short video scenarios of surface contacts in the open water, as seen through a periscope, and indicating: the side of the contact that is visible [port, starboard, or N/A], as well as the angle on the bow [a number between 0 and 180 degrees, with lower numbers indicating that the contact is headed in the direction of the viewer's submarine]. Participants had a maximum of one minute to respond on each of 30 trials, although most trials were completed well before the time limit was reached. Overall periscope skill scores were calculated as a combination of accuracy and speed, with higher representing better performance. Scores ranged from 0.46 to 0.85 (M = 0.70, SD = 0.08). Using a median split procedure, participants were divided into two groups: experts (more skilled) or novices (less skilled).

Object Block Stimuli & Task

Object images were taken from the Snodgrass and Vanderwart (1980) inventory of line-drawn objects at varying levels of fragmentation. Through pilot work, 140 object images were selected (the fragmented real object set) and their non-object variants were created through random scrambling (the fragmented non-object set). Sample stimuli from this task are presented in Figure 1A: the fragmented image on the left depicts a duck (correct answer = yes; real object present) while the image on the right shows the scrambled, non-object version of the image on the left (correct answer = no; no object present). The Object Block consisted of 140 trials, a randomized mix of 105 real objects and 35 non-objects. Four versions of the object task block were created to counterbalance real and non-object stimuli across conditions. Each participant was exposed to only one of the four versions of the object task block. As such, participants viewed either the real or the non-object version of each object image, never both versions, and the same image was never repeated. For each trial, participants were instructed to indicate (as soon as the image appeared on screen) by button press whether the image contained a possible object or not. It was emphasized to participants that they should rely on their impressions, and that they did not need to name the possible object. Trials were presented in two blocks of 70, each lasting 6-7 minutes, with a brief rest in between.

Periscope Block Stimuli & Task

Submarine-specific images were taken from the POAT. Images depicted a vessel in the open water as seen through a submarine's periscope. The images varied as a function of: vessel type, distance from the submarine, angle on the bow, sea state, and time of day. With input from a retired Navy submarine Captain, each periscope image was

classified as to whether or not it was safe to remain on the present course (heading north, with the periscope facing at 45 degrees) and speed. Seventy-six trials were classified as safe and 64 were unsafe. Sample stimuli from the periscope task are presented in Figure 1B: the image on the left shows a container ship's port side with a 40° AOB (correct answer = no; submarine course not safe) while the image on the right shows a cruise ship's starboard side with a 35° AOB (correct answer = yes**;** submarine on safe course).

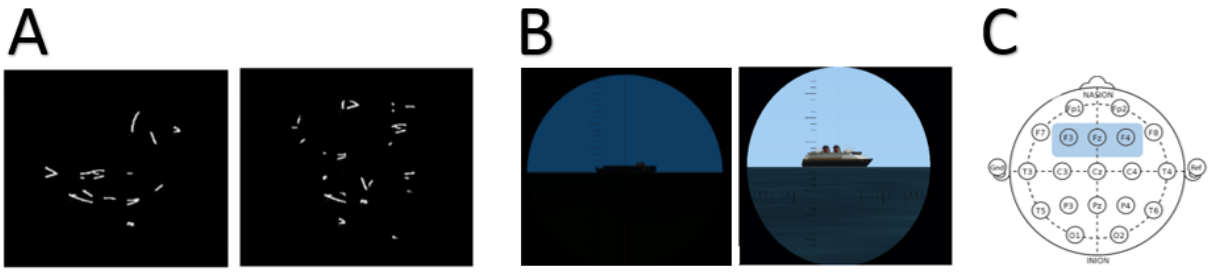

Figure 1. Stimuli & EEG Montage. Sample stimuli for the Object Block and the Periscope Block are displayed in (A) and (B), respectively. (C) shows the layout of electrode sites across the scalp (top-down view, nose on top) in the Quick-20 Dry EEG Headset; analyses focused on frontal sites labeled as F3, Fz, and F4 (as in Luu et al., 2010).

The Periscope Block also consisted of 140 trials. Participants viewed each periscope image once; periscope images were never repeated. For this task, participants were informed about their own submarine's current course and then answered by button press whether or not it was safe to remain on the present course, speed, and depth. The researcher emphasized that participants should rely on their impressions, and that they did not have to be able to indicate the action(s) they would take in response to this assessment (i.e., if any action needed to be taken, then it was not safe to remain on course). Trials were presented in two blocks of 70, each lasting 6-7 minutes, with a brief rest in between.

Task Timing

The trial timing was specifically designed to induce quick responses based on minimal perceptual input and to eliminate the opportunity for extensive deliberation. In other words, both tasks elicited intuitive decision-making responses. For both the Object Block and the Periscope Block, each trial began with a cue (an asterisk) that appeared on the screen for 500 ms. The cue, alerting participants to the start of a trial, was replaced by a fixation cross which also lasted for 500 ms. Following the fixation cross, an image was presented on screen for 400 ms; the image was then immediately masked by a blank screen. This blank screen remained for two seconds or until the participant made a binary response (object vs. not; safe vs. unsafe) with a button press. The entire response interval – from the onset of each image to the offset of the blank screen – was a maximum of 2.4 seconds. To prevent trial onset expectancy effects in the EEG, the inter-trial interval (a blank screen) lasted between 1.5 and 2.5 seconds, and was randomly generated.

Data Collection, Processing, & Analysis

Behavioral Response Processing

Mean response times (RTs) and accuracy scores (percent correct) were calculated per participant for each task. Per standard practice, trials with RTs faster than 200 ms or slower than 2.5 standard deviations above the mean were excluded from the analyses, and all RT analyses were based on correct trials (Luce, 1986).

Electroencephalography Recording

ERP processing and analyses were consistent with the guidelines set forth by Picton et al. (2000). The EEG recording was acquired using the Quick-20 Dry EEG Headset by Cognionics, Inc. (San Diego, CA) which covers the scalp with 19 Ag/AgCl electrodes arranged in the international 10-20 system configuration, as depicted in Figure 1C. During EEG recording, voltages at each electrode were recorded relative to voltages at a standardized reference site which, for this research, was positioned on the bone behind the right ear, the right mastoid. Lastly, recording procedures allowed EEG to be sampled at 500 Hz with a 24-bit analog-to-digital converter and band-pass filtered (0.1–100 Hz) to eliminate electrical noise from the environment.

Event-Related Potential Processing

During EEG recording, an electrode positioned on the left mastoid also recorded EEG. Offline, scalp EEG channels were re-referenced to the average of both mastoids to mitigate any laterality effects that resulted from the one-sided reference site (right mastoid) used during recording. The mastoids are beneficial placements for use during re-reference procedures because, while they are close in distance to the other electrodes (allowing them to pick up on similar artifacts, if any), they record less signal from the brain. Next, EEG was band-pass filtered between 0.1-30 Hz to focus on delta, theta, alpha, and beta brainwaves. Artifact rejection procedures eliminated trials from each channel's ERP average if EEG showed extreme shifts in voltage (i.e., >100 μV difference from highest to lowest points) within short time segments. This peak-to-peak artifact detection procedure was carried out in sliding windows of 200 ms. ERPs per condition were calculated by averaging artifact-free EEG segments together, time-locking segments to stimulus onset, retaining a 200 ms pre-stimulus baseline, and extending through perceptual and cognitive processing times to 800 ms.

To investigate the neural substrate, we investigated brain activity patterns that we expected were representative of perceptual, attention, and cognitive processes supporting intuitive decision-making. Specifically, ERP mean amplitudes and peak positive and negative latencies were calculated at frontal sites (F3, Fz, and F4), as in past work (Luu et al., 2010), to assess anterior effects during the time frames for two frontal components observed near our time period of interest (~200-500 ms): P200 (between 175-275 ms) and N300 (between 275-350 ms in the Object Block and between 275-400 ms in the Periscope Block). ERP components are traditionally named according to the valence of their peak (positive, "P" or negative, "N") along with their chronology. For example, the P200 is so named because it is a positive peaking component that occurs around 200 ms following the onset of a stimulus.

In the Object Block, four participants were removed due to too few artifact-free trials. As a result, there were 13 novices and 10 experts included in the Object Block data analysis set. In the Periscope Block, five participants were removed due to too few artifact-free trials. As a result, there were 12 novices and 10 experts included in the Periscope Block data analyses. Table 1 summarizes the demographic details per group per Block. As expected, the experts had significantly higher expertise test scores than the novices ($ps < .001$). However, there were no significant differences between groups as a function of age ($ps > .31$), years of submarine experience ($ps > .58$), or (for retirees only) years since leaving active duty ($ps > .13$), although trends were in the expected direction (experts were older, more recently retired, and had more years of Navy experience).

Table 1

Group Demographics per Block

Block	Group	Expertise Test Score M (SD)	Age (years) M (SD)	Navy Experience (years) M (SD)	Retirees per group (#)	For retirees, years since active duty M (SD)
Object Block	Novice (n=13)	63.3 (6.2)	37.2 (9.9)	14.1 (9.0)	5	15.6 (6.3)
	Expert (n=10)	75.8 (4.6)	40.3 (8.6)	15.3 (7.6)	7	9.7 (5.9)
Periscope Block	Novice (n=12)	63.3 (6.5)	36.2 (9.7)	13.3 (8.9)	4	15.5 (7.3)
	Expert (n=10)	75.8 (4.6)	40.3 (8.6)	15.3 (7.6)	7	9.7 (5.9)

Statistical Analyses

Accuracy analyses included all trials, while RT and EEG analyses focused on correct trials only. Statistical tests (*t*-tests and Analyses of Variance) included the between-subjects factor of expertise group and, for the EEG analyses, the within-subject factor of channel location (F3, Fz, and F4).

RESULTS

Performance (Accuracy & Response Time)

In the Object Block, participants were on average 75.51% accurate ($SD = 9.48$). Contrary to our hypothesis, accuracy differed between the expertise groups, $t(21) = -2.68$, $p = .02$. Specifically, experts were significantly more accurate than novices, achieving a mean of 79.70% correct ($SD = 7.90$) compared to the novices' mean of 70.53% ($SD = 9.47$)

(Figure 2, left). Focusing on correct trials only, participants responded on average within approximately one second (*M* = 1008.64 ms, *SD* = 243.40). There were no differences in response speed between the expertise groups, $t(21) = -0.10$, $p = .92$ (experts: *M* = 1014.36 ms, *SD* = 220.20; novices: *M* = 1004.24 ms, *SD* = 268.70). In summary, experts and novices took approximately the same amount of time to respond but, nevertheless, experts were significantly more accurate than novices.

In the Periscope Block, participants were on average 60.75% accurate (*SD* = 5.79), which is substantially lower than in the Object Block condition, but still significantly better than chance. As hypothesized, accuracy differed between the expertise groups, $t(20) = -2.37$, $p = .03$. Experts were significantly more accurate than novices, achieving a mean of 63.59% correct (*SD* = 4.45), compared to the novices' mean of 58.38% (*SD* = 5.86) (Figure 2, right). Focusing on correct trials only, participants responded within approximately 1¼ seconds (*M* = 1254.84 ms, *SD* = 276.67), which is similar to performance in the Object Block condition. There were no differences in speed between the expertise groups, $t(20) = -0.26$, $p = .80$ (experts: *M* = 1271.99 ms, *SD* = 271.90; novices: *M* = 1240.55 ms, *SD* = 291.82). Again, similar to the Object Block condition, experts and novices took the same amount of time to respond but experts were significantly more accurate than novices.

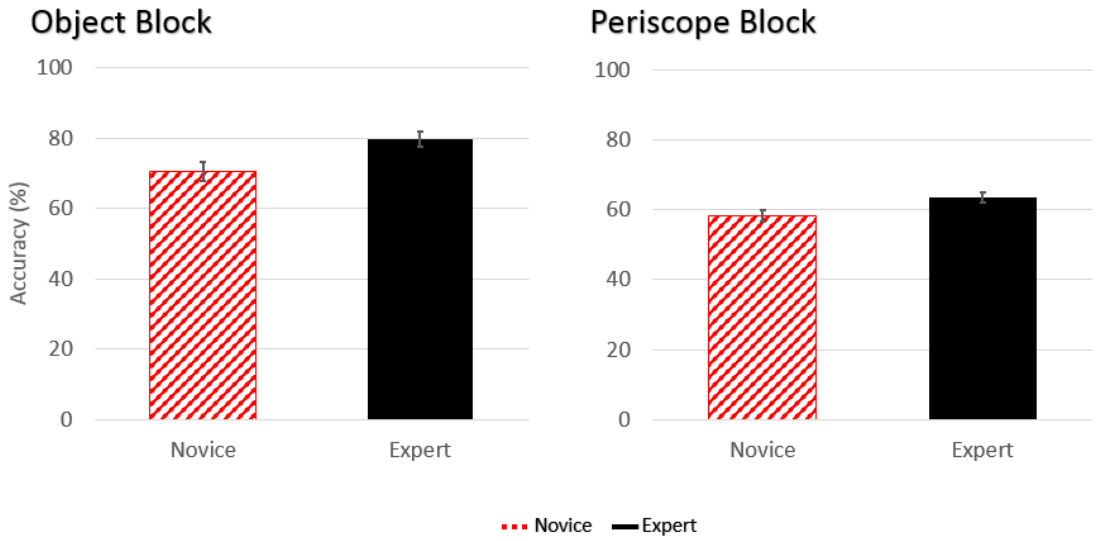

Figure 2. Performance Results. This figure shows performance accuracy in the Object Block (left) and in the Periscope Block (right) for both novices (red/white) and experts (black). In both cases, experts were significantly more accurate than novices; there were no differences between groups in terms of response speed.

Event-Related Potentials (ERPs)

In the Object Block, at frontal sites, ERPs began to differ between the expertise groups at the P200 ERP component (between 175-275 ms), such that amplitude was more positive and slightly later peaking in experts compared to novices. ANOVAs comparing P200 mean amplitude and P200 peak latency across channel locations (F3, Fz, F4) and group (novice, expert) showed main effects of Group: $F(1,63) = 7.32$, $p = .01$ and $F(1,63) = 4.55$, $p = .04$, respectively. Following the P200, N300 amplitude and latency appeared similar between groups (between 275-350 ms); as such, a second set of ANOVAs showed no significant effects of Group, $ps > .58$ at the N300. Figure 3 (left) shows ERPs per group on this object task.

In the Periscope Block, at frontal sites, ERPs began to differ between groups at the P200 ERP component (between 175-275 ms) such that amplitude was more positive in experts compared to novices. ANOVAs comparing P200 mean amplitude and P200 peak latency across channel locations and group factors showed a main effect of Group for amplitude, $F(1,60) = 3.93$, $p = .05$, but not for latency, $p = .13$. Following the P200, the N300 appeared to peak earlier in experts compared to novices (between 275-400 ms); accordingly, ANOVAs showed a main effect of Group for N300 latency, $F(1,60) = 3.95$, $p = .05$, but not for amplitude, $p = .17$. Figure 3 (right) shows ERPs per group on this submarine-relevant task.

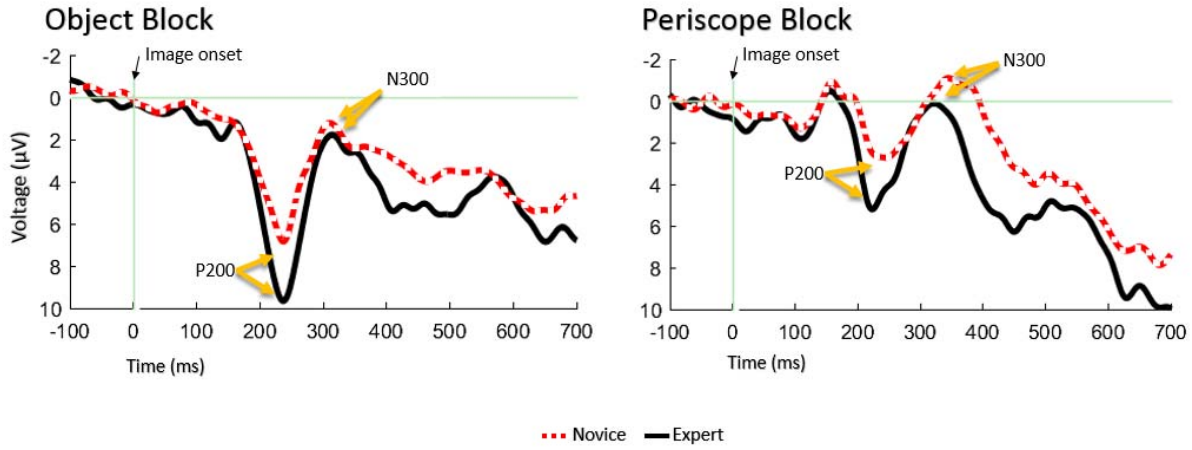

Figure 3. Electrophysiological Results. This figure shows average neural activation at frontal scalp sites (F3, Fz, and F4 averaged together), with negative plotted up (as is convention), from 100 ms prior to the stimulus onset, through image presentation (0-400 ms), to 300 ms after image offset. The novice group is depicted in red-dashed lines and the expert group is depicted in black. Object Block ERPs are on the left while Periscope Block ERPs are on the right.

DISCUSSION

These two tasks required participants to engage in intuitive decision-making. In both the everyday object detection task (Object Block) and the submarine-relevant task (Periscope Block), participants were encouraged to make quick decisions based on briefly-presented visual stimuli, many of which were visually degraded. The participants were instructed to respond even if they weren't 100% sure of the accuracy of their response. In essence, they were instructed to respond based on their intuitive hunches or gut-feelings. Furthermore, since accurate intuition is heavily dependent upon one's accumulated experiences and pattern-matching skills (Kahneman & Klein, 2009; Reber, 2013), we can confidently expect that everyday object recognition expertise (occurring naturally) and periscope expertise (learned through training and deployed experiences) contributed to the participants' performance.

Generalizability of Object Recognition Intuition across Samples

The results from the Object Block offer several interesting findings regarding our research questions about the generalizability of intuitive decision-making. To begin with, we examined whether the findings of Luu et al. (2010) with college participants would generalize to a different sample of interest. Experimental procedures were nearly identical between our two studies. Overall performance accuracy (~75%) was similar to that which was observed in past work (~70%), suggesting that the task does generalize to other populations. Regarding reaction time, on average, submariners took about one second to log their correct responses. Their responses were slightly slower than the mean response time observed among college students (~778 ms, or ¾ of a second) in the study conducted by Luu et al. (2010). Stimulus-response speed slows with age (Woods, Wyma, Yund, Herron, & Reed, 2015), as do spatial abilities (Techentin, Boyer, & Voyer, 2014). As this perceptually-driven task involves both spatial processing and stimulus-dependent responses, this difference between samples makes sense given that the mean age of our submariners was 38, while the mean age of the college student sample was 22.

Unexpectedly, our statistical analyses suggested a significant performance advantage for expert submariners: experts were approximately 10% more accurate at intuiting real objects than were novices. This performance enhancement was not due to a speed accuracy trade-off, since there were no significant group differences detected in response times. The results suggest that our group of expert submariners possess a special skill that allows them to perform more accurately on quick perceptually-driven decision-making tasks. However, the exact nature of this skill – whether it involves better selective attention to relevant features, superior perceptual processing, and/or enhanced application of stored knowledge – cannot be determined from performance data alone.

An analysis of the neural activation differences between the two expertise groups helps shed some light on the performance findings. ERPs, as seen in Figure 3, show the average voltages of brain activity over time per group. Starting from the time when fragmented stimuli were visible on screen (at 0 ms), brain activity was similar between novice and expert submariners until about 200 ms (see Figure 3, left), suggesting that initial perceptual processing did not differ between groups. However, shortly after that at the ERP component called a P200, brain activity (between 175-275 ms) was much more positive for experts (~10 μV, black line) than for novices (~7 μV, red-dashed line). Based on these findings, it seems plausible that brain activity (and concurrent processes or mental representations) indexed by P200 amplitude likely contributed to the observed between-group difference in performance. In other words, when P200 amplitude is more positive, it may reflect better selective attention, superior perceptual processing, and/or enhanced application of stored knowledge. More broadly, the P200 is elicited as part of the normal response to visual stimuli. It has been studied in relation to visual search and attention (e.g., target detection, vigilance tasks), feature analysis, and memory and repetition effects (Freunberger, Klimiesch, Doppelmayr, & Holler, 2007; Evans & Federmeier, 2007). Most relevant to the interpretation of our research results is a proposed theory by Phillips and Takeda (2009) which posits that the P200 indexes a selective attention process that identifies meaningful stimuli through feature suppression. Thus, the findings from the Object Block may suggest that selective attention drives the perceptual recognition of object-like features, which in turn results in a more efficient comparison of sensory inputs with long-term memory. If so, this selective attention process is more effective in expert than novice submariners[2].

Generalizability of Intuitive Decision-Making across Tasks

The results from the Periscope Block offer critical insights about the generalizability of intuitive decision-making skills. The experimental procedures in the Periscope Block were designed to mimic the Object Block as much as possible, with the major exception that the task was specific to a particular domain of knowledge. Again, participants responded to briefly-presented visual images based on their initial hunches, in this case with regard to current course safety. Overall, performance accuracy on the Periscope Block task was not as good (~61%) as the Object Block (~75%), which suggests that the periscope task was, on average, more difficult than the general object detection task. In addition, the results varied systematically as a function of periscope operator skill.

Regarding reaction time, the submariners took on average approximately 1¼ seconds to log their correct response. This is slightly slower than response times from the Object Block. This is not unexpected, given that the stimuli were more complex (in full color, showing vessels at different distances and in varied weather conditions rather than simple white lines on a black background), and the decision-to-be-made was more complicated (heavily dependent on the precision perception of vessel direction and interpolation of intersecting courses vs. a perception of an intact object). Regardless, this task still required quick perceptual processing based on only 400 ms of visuals, and a reliance on long-term memory in order to make fast intuitive decisions. Statistical analyses revealed a significant performance advantage for experts. Specifically, experts were ~6% more accurate at intuiting course safety than were novices. Again, this performance enhancement was not due to a speed accuracy trade-off, because there were no significant response time differences between the two groups. Taken in conjunction with results from the Object Block, this result provides added evidence to support the hypothesis that expert submariners possess a special skill (e.g., selective attention, enhanced application of stored knowledge), which allows them to perform more accurately on this task.

The ERP results (as seen in Figure 3, right) in the Periscope Block are similar to those from the Object Block. Starting from the time when periscope view stimuli were visible on screen (at 0 ms), brain activity was similar between novice and expert submariners until about 200 ms. This again suggests that initial perceptual processing did not differ between groups. However, at the P200 (between 175-275 ms), brain activity was again much more positive for experts (~5 μV, black line) than for novices (~3 μV, red-dashed line). As mentioned above, we believe that selective attention was again useful in identifying meaningful stimuli through feature suppression during this time frame on this task as well, and this process was better in experts than novices (indexed by the P200). We believe that this was one factor that contributed to their enhanced intuitive decision-making performance. Following the P200 was a second between-group difference: the timing of the N300 peak (i.e., N300 latency). As depicted in Figure 3 (right), the N300 peaks earlier in experts (~330 ms) than in novices (at ~360 ms). Recent related research suggests that the N300 indexes knowledge-driven pattern matching (KDPM) and the start of decision processes (Schendan & Ganis, 2015). According to this theory, under ambiguous circumstances, top-down KDPM mechanisms quickly and iteratively pose and test

[2] As the expertise test served to group participants according to their periscope operator skill level, it may have had the unintended consequence of dividing them up according to their selective attention skills or abilities as well.

hypotheses drawn from prior knowledge starting from features perceived from a stimulus during the selective attention process (i.e., from the P200 activity). As such, N300 activity is proposed to index heavy and repeated access to domain-specific knowledge while perceived features are compared to models predicted from memory. When this pattern matching process concludes (e.g., match with model of safe course in memory), a decision is made. In the Periscope Block, as the N300 reached its peak sooner in experts compared to novices, we can posit that experts possessed an enhanced skill of applying of domain-specific knowledge to these KDPM processes during this task. This is consistent with other researchers' cognitive theories about the impact of expertise on intuitive decision-making as well: with increasing expertise, individuals become more adept at automatically identifying critical patterns in the environment and using those patterns to activate relevant information from long-term memory (Evans, 2010; Reber, 2013).

Study Limitations

Like every study, this one has its limitations. For example, in the Object Block, participants were not required to indicate the actual name of the object that they perceived. Instead, they only indicated whether they perceived the line drawing as an object. Of course, this experimental choice was intended to minimize explicit/declarative thinking and encourage quick implicit/intuitive thinking. While the research team did know which stimuli represented real objects (the fragmented line drawings) vs. non-objects (scrambled versions), there was no way to know if participants correctly perceived the identity of the real objects. At best, all we know is that they reported perceiving a gestalt; we did not inquire as to what the gestalt represented. However, as we were also attempting to replicate the Luu and colleagues (2010) study as closely as possible, we decided that this limitation was worth the risk.

Similarly, in the Periscope Block, while we did know the correct answer (safe vs. unsafe to continue), it was based on the judgment of a single subject matter expert (SME). If there had been additional time and funds, we would have recruited a panel of SMEs to ensure group consensus. While this is a study limitation, it is important to note that the SME in question is a retired Navy Commodore who was responsible for certifying whether submarine crews had completed all of their proficiency checks, and were in fact safe to deploy. As a result, he had intimate knowledge of periscope operations, among numerous other submarine tasks.

Recommendations and Best Practices for Neurocognitive Assessment

EEG recordings provide millisecond-level temporal resolution of electrical brain activity. Therefore, it is one of the most direct measures of real-time cognitive operations. The EEG techniques reported here are non-invasive. No gel was required to ensure a reliable signal between the electrode and the scalp, which reduced the amount of setup time and the time between data collection sessions (e.g., the time required to clean and prepare the equipment for the next participant). The equipment and techniques described in this study can be flexibly applied to any domain of interest where the goal is to assess the nature of intuitive, perceptually-driven decision-making. Generally speaking, participants should be required to make binary decisions (e.g., yes/no, safe/not safe) based on quickly-presented and relevant perceptual input. The actual decision should be made using an input device, such as a mouse or keyboard; they should not require a verbal response. With these perceptual and decision-making constraints in place, the intuitive task required of the participants transforms into an implicit category matching task wherein individuals utilize implicitly perceived cues from the stimuli, subconsciously referenced against their stored knowledge and experiences in long-term memory, to rapidly respond based on a hunch or gut feeling. The ERP components discussed here (the P200 and the N300), correspond to critical mental activities during intuitive decision-making, and should help to differentiate between novices and experts on similar tasks across a variety of domain areas.

This study was able to identify ERPs with a low-density (20 channel) EEG system, which we were able to rent from a commercial vendor (rather than purchase), which great reduced our overall research budget. Based on our experiences, we believe that researchers do not need to use high-density EEG arrays with up to 128 channels, as long as they place electrodes in the appropriate scalp locations: spanning an area from the top of the head (i.e., vertex) to the forehead (i.e., frontal, fronto-central, and frontopolar sites), along with the appropriate ground and reference channels. In addition, an essential component to conducting event-related EEG research involves the ability to identify, within the EEG data, the precise time when stimuli were visible to participants. We recommend using a stable and fast automated connection between the stimulus presentation software and the EEG recording software. PsychoPy is free, open-source application (Peirce, 2007; Peirce, 2009) that can be used for presenting stimuli (e.g., images, video clips, sounds) and recording responses with precise timing. PsychoPy is well-suited for neuroscience experiments because it provides a functionality that sends experimenter-designated event signals to the EEG recording system

(using a parallel port cable) to synchronize the stimulus onset with participant responses and EEG data. This type of data is called "event-labeled" data. Once event-labeled EEG data is collected, a free toolbox package called ERPLAB (Lopez-Calderon & Luck, 2014) can be used to process and analyze EEG data within Matlab. The ERPLAB documentation (see www.erpinfo.org) provides an excellent step-by-step tutorial describing the processes of averaging EEG data into ERPs, exporting amplitude and latency values for statistical analyses, and plotting the results. An alternative to Matlab is the open source R statistical software (see www.r-project.org) when supplemented with the freely-available "erpR" package (Arcara & Petrova, 2014).

Unlike expert observer ratings, neurocognitive assessments are unbiased. They are also more scalable and cost-effective as the number of learners increases. To utilize these methods for training assessment, one should first identify the pattern of ERP component activity that indicates the average expert vs. the average novice within that field of expertise. An offline test of skill will be useful in classifying trainees based on their level of expertise. The assessment task should be performed by individuals who are stationary (e.g. seated in front of a computer) and not speaking (e.g., responding with a button press) because movement artifacts can mask the neural signals. To assess the effectiveness of training, the learner's ERP can be compared to the expert and novice averages (that were computed through the same EEG protocols) both before and after training.

A particularly useful validation design will test both experts and novices on the new training program as well as a control training task. If the new training program is useful at developing expertise which the experts already have, then the novices will improve on the task while the experts should not. Similarly, if the control training task does not train as well as the new training program, novices will improve (but not as much as those who were in the new training program) and, again, experts will stay the same. In all cases, if the training is successful at developing expertise, the learners' ERPs should change from appearing more similar to the novice average (prior to training) to more similar to the expert average (after training).

Summary & Conclusion

Results from this study provide evidence for the existence of both generalized and domain-specific neurocognitive markers of intuitive decision-making. These markers represent the processes of selective attention and knowledge-driven pattern matching – both of which are necessary for quick and accurate intuitive decision-making. Occurring first (in time), the P200 indexes the selective attention process. Next, working from the features that were emphasized during selective attention, the N300 indexes a complex series of pattern matching processes which heavily involves internal models retrieved from memory (i.e., this is a very expertise-dependent process). Because P200 effects were observed in both the Object and Periscope Blocks, we concluded that the underlying processes were not dependent on submarine-domain specific knowledge and that, instead, they represented an enhanced generalizable skill – selective attention - that was inherent to the expert submariner group. In contrast, because N300 effects were detected only in the Periscope Block, we concluded that the processes underlying activity during this time (knowledge-driven pattern matching) were related to submarine-specific knowledge or current skill level. Critically, neural correlates of domain-specific intuitive decision-making differed according to domain skill level (novice vs. expert) during the N300, indexing a perceptual hypothesis testing process that draws heavily from knowledge/expertise. As such, the techniques reported here should be useful when examining expertise-based intuitive decision-making.

ACKNOWLEDGEMENTS

Funding for this study was provided by the Office of Naval Research to Aptima, Inc. The views expressed in this paper are solely those of the authors, and do not necessarily reflect the opinions of their sponsors at the Office of Naval Research, the Naval Air Warfare Center Training Systems Division, or any other Department of Defense agency, unless stated in official directives. The authors wish to thank the many reviewers of this paper for their thoughtful feedback.

REFERENCES

Arcara, G. & Petrova, A. (2014). *erpR: Event-related potentials (ERP) analysis, graphics and utility functions* (R package Version 0.2.0). Available from https://CRAN.R-project.org/package=erpR

Evans, J. St. B. T. & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives in Psychological Science*, 8, 223–241.

Evans, J. St. B. T. (2010). Intuition and reasoning: A dual-process perspective. *Psychological Inquiry*, 21 (4), 313-326.

Evans, K. M. & Federmeier, K. D. (2007). The memory that's right and the memory that's left: Event-related potentials reveal hemispheric asymmetries in the encoding and retention of verbal information. *Neuropsychologia*, 45, 1777-1790.

Freunberger, R., Klimiesch, W., Doppelmayr, M. & Holler, Y. (2007). Visual P2 component is related to theta phase-locking. *Neuroscience Letters*, 426, 181-186.

Kahneman, D. & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, 64 (6), 515-526.

Kofka, K. (1922). Perception: An introduction to the Gestalt-theorie. *Psychological Bulletin*, 19, 531-585.

Landsberg, C. R., Mercado, A. D., Van Buskirk, W. L., Lineberry, M., & Steinhauser, N. (2012). Evaluation of an adaptive training system for submarine periscope operations. In *Proceedings of the 56th Annual Meeting of the Human Factors and Ergonomics Society*, 56, 2422.

Lopez-Calderon, J. & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8, 213.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.

Luu, P., Geyer, A., Fidopiastis, C., Campbell, G., Wheeler, T., Cohn, J., & Tucker, D. M. (2010). Reentrant processing in intuitive perception. *PloS One*, 5, 9523.

Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods*, 162 (1-2), 8-13.

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2, 10.

Phillips, S. & Takeda, Y. (2009). An EEG/ERP study of efficient versus inefficient visual search. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, 31, 383-388.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R. Jr., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology*, 37, 127–152.

Reber, P. J. (2013). The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research. *Neuropsychologia*, 51 (10), 2026-42.

Schendan, H. E. & Ganis, G. (2015). Top-down modulation of visual processing and knowledge after 250 ms supports object constancy of category decisions. *Frontiers in Psychology*, 6 (1), 1289.

Snodgrass, J. G. & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6 (2), 174–215.

Techentin, C., Voyer, D, & Voyer, S. D. (2014). Spatial abilities and aging: a meta-analysis. *Experimental Aging Research*, 40 (4), 395-425.

Woods, D. L., Wyma, J. M., Yund, E. W., Herron, T. J., & Reed, B. (2015). Age-related slowing of response selection and production in a visual choice reaction time task. *Frontiers in Human Neuroscience*, 9: 193.