Applying Neuro-Metrics to the Development of Learning Solutions

Adam L. Hall
Chairman, Nervanix LLC
Fort Myers, Florida
Adam.Hall@Nervanix.com

Stephen J. Kenton
VP Development, SS&C Technologies
New York, New York
SKenton@sscinc.com

ABSTRACT

One significant challenge for organizations developing learning courses is the ability to effectively predict and evaluate their impact on learners. One way such evaluation can be accomplished is through using neuro-based technology to gauge participant responses during training. As the utility of neuro-studies in learning environments becomes clear, more out-of-lab approaches are being developed. However, to advance these studies beyond data mining exercises, the essential components of the scientific process must be assiduously followed. Towards this goal, the authors have performed several studies of digital course material using neuro-metric indicators (EEG) to analyze and provide recommendations for improving effectiveness and efficiency of instructional delivery. These studies were carried out using traditional techniques for data collection and processing. While these efforts produced a number of interesting insights into the efficacy of learning objects, the experience more importantly illuminated the need for a set of requirements to make the experimental process itself more effective. Based on these experiences, the goal of this paper is to supply a roadmap for building a framework to be used by an Instructional System Designer that can support repeatable experimental workflow and leverage the current neural research with evaluations informed by expert analytic language. Insights and methods presented in the paper offer new and powerful insights into a learning object's effectiveness and efficiency in delivering its intended results to an audience of learners. The subsequent paragraphs discuss the variance of expectations to actual results (as they pertain to attentional and emotional states, when/whether memory is being encoded and/or retrieved, and the synchrony of brainwave data across multiple subjects given common stimuli) in terms of content delivery analyzed through neuro-metric indicators.

ABOUT THE AUTHORS

Adam L. Hall is the Founder and Chairman of Nervanix, LLC. Adam is a well-respected ed-tech entrepreneur and business leader. After SS&C Technologies invested in Nervanix in 2016, Adam served as Senior Vice President at SS&C and ran their Learning Institute, focusing primarily on the financial services sector. In 2019, Adam left SS&C to focus Nervanix on other verticals, specifically those where high stakes learning is a priority. Adam is a serial entrepreneur. In 2000, Adam co-founded an education technology company, Impact Education, Inc., focused on delivering basic skills to the K-12 and Higher Education markets. After a successful ten years, Adam sold the company to Houghton Mifflin Harcourt in 2010. He then went on to found Nervanix, a highly innovative business focused on optimizing efficiencies and efficacy in learning by creatively applying neuroscience technologies to the process of instructional design. Adam received his B.A. in Economics from Columbia University. He was appointed and twice re-elected to the Education Board of directors for the Software Information Industry Association (SIIA) and is an active member of the Consortium for School Networking (CoSN). Adam volunteer coaches freshman football in his hometown of Fort Myers, FL and serves on the school board of his local Catholic High School. He and his wife Angie are proud parents of four.

Stephen J. Kenton is a Vice President in Development at SS&C Technologies where he has created a patented distributed system messaging and workflow architecture, developed financial valuation and transaction systems, and has worked extensively with data visualization. While attending MIT for physics, he worked with Margaret Hamilton at HOS, modeling highly reliable real-time software systems. He continued that work in critical systems in high speed parallel processing and DSP. In the late nineties, Steve began working in financial services, joining a startup hedge fund to develop a proprietary options trading operation system, and later formed a development consultancy with major banks as clients.

Applying Neuro-Metrics to the Development of Learning Solutions

Adam L. Hall
Chairman, Nervanix LLC
Fort Myers, Florida
Adam.Hall@Nervanix.com

Steve Kenton
VP Development, SS&C Technologies
New York, New York
SKenton@sscinc.com

INSTRUCTIONAL SYSTEM DESIGN

Instructional System Designers (ISDs) apply their skills to create instructional materials, bringing greater efficiency and effectiveness to the process of transferring knowledge to a learner. This discipline has been developing formally since at least the 1950s, when the idea of systems engineering was in its nascent state. Excitement about the work of Weiner, Turing and Von Neumann in computational approaches extended systems analysis to many other fields, particularly the educational area.

"The systems approach was born in the field of systems engineering and was first applied rigorously to the design of electronic, mechanical, military and space systems. Here it got involved with man-machine systems and from there it was a short step to organizational and management systems. It began to be used in training and then in education from the late 1950s and early 1960s. What was new was the systematic methodology; the general approach is as old as scientific method itself and has been practised [sic] by a few enlightened souls for generations".
A.J. Romiszowski – Designing Instructional Systems (1981)

Instructional Designers apply various models (ADDIE, Agile) to their craft, with the intention of optimizing the learning process given the conditions under which instruction occurs and the anticipated results of the instruction itself. ISDs adhere to various proven design theories such as Merrill's First Principals of Instruction, Gagne's Nine Steps of Instruction and Keller's ARCS model, and they concern themselves with the psychology of the learner (such as what stimulus might trigger an emotional or engaging response), the subject matter (how it's organized and at what levels), and methods, techniques and modalities of instruction (combinations of visual, auditory and kinesthetic elements to optimize the learning results). (Dianne Rees 2019)

How ISDs measure the success of their efforts has historically been through assessment, whether formative or summative; the learner experiences instruction, and then he/she is tested to see if learning actually occurred. Such reporting through traditional assessment techniques provides valuable feedback, but does not provide indicators in real time; the ISD does not actually know how much attention, what level of emotion, to what extent memory is being encoded or retrieved, or to what extent synchrony of brainwaves across multiple learners may manifest itself given a particular stimulus or instructional "event". Such neuro-metric feedback, gleaned through applications of EEG technologies, and its incorporation into the complex and iterative process of optimizing the efficiency and effectiveness of a learning object is the main focus of this paper.

Assuming an ISD were to have access to the neuro-metric data of a sampling of learners, and assuming the ISD could apply experimental design to the development process of a given learning object by identifying specific events within a course and setting expectations for each event (i.e. levels of attention, emotion, memory, etc.), then the ISD could compare actual neuro-metric results to hypothesized results and benefit from an iterative feedback process. If the ISD hypothesizes that a certain event within a course will elicit a significant emotional response, but the neuro-data of the sampling of learners under EEG demonstrates otherwise, then the ISD has important feedback to help modify and refine that instructional event within the course. Such knowledge, and the ability to formatively respond to it, is quite valuable to the ISD in managing expectations around how effectively and efficiently a course performs.

The benefit of having such insights into a course's potential for eliciting neuro-metric results could prove quite valuable to the ISD, whose craft involves a process replete with complex, interactive disciplines – subject matter experts (SMEs), graphics and textual support, LMS delivery systems (SCORM, xAPI), etc. The ISD must pull it all together, being familiar with the requirements of the SME and the limitations of the LMS, while also being keenly aware of the external environment of the learner, constantly demanding and manipulating every sensory opportunity – all 5 senses are fair game. (Anderson 2019) So, to craft a memorable learning experience, the ISD needs to compete with the standards established by high power players in entertainment and advertising. Movie trailers are regularly evaluated using EEG; education now has a natural ally to leverage: Brain science. Just as it's important to the marketing of a movie or a consumer product through various advertising techniques, psychological sophistication is key to effective teaching – ask any teacher. To compete for attention in a world of impressive presentations in traditional and social media, learning systems can't appear irrelevant or out of date. And the personalized interfaces of computer applications have raised expectations for an experience that caters to and targets the individual's preferences. (Barnett and Cerf 2015)

In an ideal world, ISDs would develop numerous versions of a course on the same topic, with each version catering to specific learner profiles. Student profiles including demographics, expertise and personality (e.g. the Big Five: Extroversion, Agreeableness, Openness, Conscientiousness, and Neuroticism) have been found to be relevant in classifying course effectiveness. It benefits the ISD to cater a course to an intended sub-audience, particularly in certain fields such as military training, where the characteristics of the broader general audience are quite diverse (soldiers vary in size, shape, gender, IQ, cultural and socio-economic backgrounds, etc.) as compared to the characteristics of the sub-audience who must learn a specific military tactic or competency with precision. (Reynolds 2019)

The ISD, working with the tools to monitor and enhance the learning experience, also takes into consideration how effectiveness can be measured. Course development becomes collaboration between the ISD and the supporting system platform. In this paper, the promise for the Instructional System Designer to use neuro-metrics to more effectively tailor courses for students with specific learning profiles will be explored. The background of neuro-methods used to analyze cues that are relevant to the learning process are discussed, as well as the technical issues encountered in linking the neuro-metrics to the course experience. Finally, aspects of a pilot study that examines the EEG results of a group of students taking a digital HR course are presented.

NEUROSCIENCE AND EXPERIMENTAL DESIGN

The enormous popularity of neuroscience is a tribute to its potential value but also can cloud understanding about what the field has actually accomplished to date. Similarly, as artificial intelligence (AI) becomes more of an everyday reality, the misimpression that machines can imitate the actions and performance of the human brain can have negative consequences. For this reason, when brain science is enlisted to assist in computer-based learning systems, realistic expectations must be clearly established and understood. To that end, it is important to understand how the brain works, how neuro-metrics are measured, how they can contribute to inference-building techniques and the validation of expectations, and most importantly how they are limited in predicting learning behavior.

Some Neuro Basics

To start, it helps to understand some basic features of brain science. Electrical activity in the synapses of the brain corresponding to neural events, Event Related Potentials (ERPs), are observed to produce electromagnetic waves in certain bands of relatively low frequency; these are the familiar alpha, beta and other Greek letter waves. The waves are generally measured using sensors arranged in a headset in contact with a subject's scalp. The wave amplitudes for each sensor are sent to a recording system to produce a stream of values for each sensor in real time. The graph of those values over time is the electroencephalograph, or EEG. The raw EEG contains all the different wave types combined, along with noise from muscle movement from the eyes or neck and environmental interference such as from the 60hz AC electrical lines. Preprocessing techniques, which include some sophisticated machine learning, are used to clean the data, and Fourier analysis is used to extract the frequency bands. From lowest to highest frequency, they are commonly (and in this paper) identified as follows in Table 1.

Table 1. Types of Frequency Bands

Frequency Bands	Frequency Ranges	Physiological Definitions
Delta	0.5 hz to 4 hz	Deep Sleep
Theta	4 hz to 8 hz	Drowsiness

Alpha	8 hz to 13 hz	Relaxed but Alert
Beta	13 hz to 30 hz	Highly Alert and Focused
Gamma	30 hz to 50 hz	Unity of Consciousness

The psychological cues of attention, emotion (attraction or avoidance), and memory activity are derived from the power signals of these frequency bands (10 ^Amplitude) in certain areas of the brain observed over 20 second epochs. The Attention metric is derived from the average power signal of the alpha waves observed in the frontal lobe (above the fore head). The Emotion metric is also derived from the frontal lobe alpha power signals, but depends on the asymmetry of the average signal on the right side of the brain from that on the left side. Memory activity measured is associated with the parietal and temporal lobes, on the top and side of the brain, and derives its metric from the combination of the average signals of the alpha, theta and gamma waves from the left side of the brain added to the same combination from the right side.

In addition to these standard indicators, one can measure the brain synchrony across subjects looking at the same event. This metric is known as Cross Brain Correlation (Barnett-Cerf 2017), also referred to as "Engagement". Alpha power signals can be measured across one subject's brain over an event epoch, and correlated in pairs of the same electrode with the other subjects' brains. A more recent "Engagement" metric has been proposed that actually uses some of the signal noise from muscle activity and eye movement as significant indicators. If similar correlated noise is observed among a number of subjects for the same event, while not strictly neural behavior, a group reacting viscerally (literally) to the same thing probably indicates an event worthy of analysis.

These neuro-metrics can provide valuable signposts around how instructional presentation methods affect an audience or a classroom. (Barnett and Cerf 2016) The characteristic differences that appear among learners when viewed through the lens of these metrics can signal methods that are more appropriate to one type of learner, or group of learners, as opposed to another. For military training, where a broad cross-section of students needs to be classified and motivated to particular disciplines with high degrees of expertise, determining the optimized delivery for instruction among these classifications (and sub classifications) is crucial. So, refining the focus of a course for particular learner profiles justifies serious commitment.

While the fast-evolving analytics using neuro-metrics are exciting and quite promising, there are significant technical and financial considerations with regard to the physical apparatus used in studies that must also be taken into consideration. Two leading technologies are EEG and fMRI. EEG is limited by signal to noise ratio issues and how deeply into the subcortical regions of the brain it can examine, although EEG gives result granularity of fractions of a second. Data collected from fMRI monitors blood flow deep in the brain and can show exciting indications of neural processing, but only with a granularity of several seconds. And the sophisticated equipment required for fMRI is complicated and expensive, costing on the order of $5 million, where a scientific quality EEG system is in the $20,000 range. (Harrell 2019)

Given the significant difference in cost between the less expensive EEG versus the costly fMRI, the practicality of their applications (EEG offers much more flexibility and a wide variety of use cases), and this paper's focus on drawing inferences around the qualities of a learning *object* versus the qualities of a learning *brain* (the former requiring less invasive pinpointing), it is the opinion of the authors that EEG is the better choice for accomplishing the types of neuro-metric learning applications discussed in this paper.

With that established, it is important to note what EEG *cannot* accomplish in the evaluation of learning. It has not yet been determined through EEG when or where learning happens. *"The field has yet to identify a clear, finite set of brain regions whose activity modulates when an individual processes content."* (Barnett and Cerf, 2017). However, what can be accomplished with EEG as it pertains to measuring learning is still quite significant. EEG offers valid, meaningful metrics such as attention, emotion, memory and synchrony (or engagement). These metrics alone, while they don't provide definitive proof of whether, when, where or how learning occurs, do in fact provide valuable insights into how the learning brain is functioning given specific instructional stimuli, and ISDs can make important inferences as they take an iterative approach to master the complexity of their learning designs. Definitive measurements of learning can't be observed in absolute real time. Testing is still the only way to determine how well content has been learned. But, inherent within the instructional design process are neural indicators that can provide significant clues around how instructionally impactful a course will be to a learner. Such indicators, and their appropriate applications within the instructional design process, provide a critical tactical advantage for the modern ISD.

The Experimental Design

While it's difficult to determine *how* the brain learns, it's far simpler to capture and apply specific brain behavior cues that assist an ISD in refining and honing instruction to elicit the optimal activities of the highly engaged and learning brain. In short, appropriate and unique application of EEG data can facilitate the designer's process.

In order to facilitate this process through applications of neuro-metrics, it is important to ask the proper questions and monitor the proper learning cues. The ISD's question should not be "is the brain learning?" but rather "are the neuro-metrics of a learning object optimized in a way where the brain has the best chance to learn?" For example, what exactly does a teacher see when looking at a class, determining whether or not the instruction is translating to knowledge? Many teachers would say facial cues and body language. Perhaps if all the students react the same way, together – like to the punch line of a joke – that would indicate the successful translation of instruction to knowledge? In neuro-metric terms, such group reaction is a synchrony event – or Engagement.

Such inferences, when based on actual neuro-metric data derived from raw EEG signals, are significant, and as this paper argues, can have a powerful positive impact on an ISD's ability to create efficient and effective learning material. However, these inferences and their utility can be exponentially enhanced and validated when hypotheses are established a priori. And the value of these hypotheses increases dramatically when they can be formulated using repeatable protocols, with a certain level of granularity (in the case of EEG neuro-metrics, 20 second epochs to get meaningful analytics) and with a certain level of pre-defined language (offered by both ISDs and neuroscientists as they study data) that correlate with different combinations of expected neuro-metric patterns. By establishing a priori hypotheses, the subsequent inferences become testable and repeatable, and not just episodic. While it is presumed that every ISD intends to build an excellent learning object, the ISD is empowered significantly more when his/her specific intentions within a course can be defined, categorized, qualified, quantified and measured.

However, before meaningful neuro-analytics can be applied, these course events must be defined, and neuro-metric expectations must be set. For example, when the learner sees the happy puppy, one might predict high positive emotional valence. Or, when the learner moves from an exciting moment to a reflective moment in the course, one might expect attention to be low and meditation to be high. By defining, categorizing, qualifying and quantifying events, and establishing a priori hypotheses, the ISD can much more easily and effectively map EEG neuro-metrics to content material, and glean meaningful and usable analytics. In fact, evidence points to EEG indicators corresponding to other measures of personality types (the Big 5). This relationship suggests that targeted messaging for different subject profiles can be made more effective for specific audiences, the holy grail for the commercial tailoring of marketers (Harrell 2019). These findings are an encouraging result for the ISD attempting to improve how effectively certain students interact with a course (Cerf, pending publication).

Effectively, the challenge to consider here is the of use rapidly evolving neuroscience, applied against a comprehensive solid framework to inform and improve the ever changing field of learning systems; actually using brain science to evolve education. However, due to current limitations it is wise to be humble when considering what might be inferred about the learning brain from neuro-metrics. Although even with the technological limitations, when applied carefully, neuro-metrics could be a promising tool to help Instructional Designers characterize learning outcomes of students.

ENFORCING THE SCIENTIFIC METHOD

In-lab studies have traditionally relied on developing ad-hoc protocols and processes that are particular to the experiment in hand. This is due to the groundbreaking nature of the studies, requiring researchers to constantly explore new paradigms. As a result, limited work has been conducted on how to develop easy to use and standardized neuro-study tools. This lack of easy to use software puts areas like education at a disadvantage when it comes to integrating brain science. Particularly, modern learning systems, where the presentation of digital courses is much more manageable, lend themselves well to the monitoring and coordination of analysis, and is a natural candidate for methodical software support. In a framework where digital courses can be judged with established neural-metrics as inputs, the basic process of hypothesis and testing, Scientific Method, can be repeated on new course material, providing a path to systematic experimental design in learning systems, distinct form the requirements of a brain study. This section describes the logic and processes the authors went through to begin determining what such a neuro-based software tool for education should encompass. The section starts by discussing the current limitations of study design and then moves into a discussion of specific interface elements the authors are including in a system prototype.

The field of lab support software addressing the issues mentioned is a rather sparse landscape. However, a notable contribution has been made by the University of Quebec with the LabPal system, described by Sylvain Halle in International Symposium on Software Testing in 2017. This software is designed to support study development in computational performance laboratories integrating a testing platform, report writing and data and graphical reporting, with an explicit goal of delivering repeatability and reproducibility in experimental design. In the cited paper, Halle describes doing a comparison of sorting algorithms where experimental operation and results gathering are supported by the LabPal system.

In psychological studies, reproducibility takes on particular challenges (e.g. every subject is unique). In fact, the issue has been raised to a crisis in a number of papers with alarming titles, for example “Is there a reproducibility crisis?” (Baker 2016), “Is science facing a reproducibility crisis, and do we need to?” (Fanelli 2016), and “A crisis of Confidence?” (Pashler and Wagenmakers 2012). But, at least the ability to reliably repeat an experiment within a study or between studies should be attainable for software based systems. From supporting the neuro-study on the human resources course described later in this paper, the authors hoped to deliver a set of requirements, a roadmap, for a system that could provide repeatable, efficient laboratory workflow. The approach to the problem became specifying a system that enforces and facilitates scientific method: hypothesis evaluation in a structured process for the ISD. The authors envisioned a report generating system that essentially produces a rough draft of the neuro-analytics report by starting with the hypothesis language established on the front-end, and filling in the language and graphical material that is produced by the neuro computational processing, and with screenshot documentation of the events from the course. The process for producing the ISD evaluation report became a set of requirements to support similar ongoing research by the authors for clients. For the purposes of this paper, the discussion is from the standpoint of those requirements rather than any particular implementation.

A prototype framework for such a system began to emerge from the authors’ consideration of these requirements. In Figure 1, an interface with a browser screen for displaying and stepping through the course by the ISD is shown integrated with a panel for marking the relevant events, in this case with html actions like clicks and URL changes. When an event is identified, a screen (Figure 2) appears for the ISD to enter the event description including the hypothesis language that is to be included in the draft report, and expectations to be tested. Possible terms to describe the metric behavior are provided for the ISD from a library, shown in Figure 2 as selectable using a slider control. These features allow the ISD to annotate a course before the study begins in order to indicate significant events. For the events to be appraised in terms of the brain science, specific psychological cues that will be evident in the EEG metrics need to be identified and the expectations need to be described for comparison to actual results. As a result, the ISD should begin by bringing up the course to go through and identify events of psychological significance.

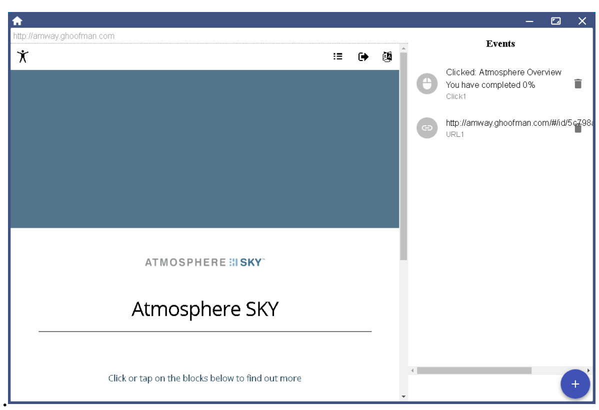

Figure 1. Example of Course Loading and Event Identification

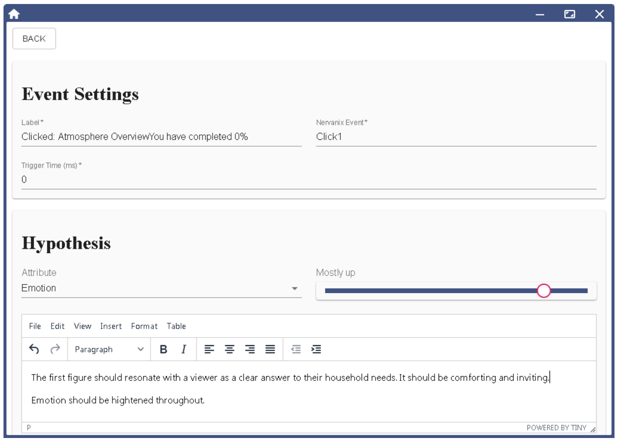

Figure 2. Example of Annotation: Event Descriptions and Hypotheses

The ISD analyzes the events in terms of the cues provided by the neuro-metrics previously discussed in this paper: Attention, Emotion, Memory and Engagement. Since these metrics have technical meaning that may not be in the ISD's skillset, verbal descriptions from the neuroscience literature and tutorial support should be available to assist an ISD in the process. In the studies used to guide the results presented in this paper, neuroscience experts were available to provide insights. But going forward, the intention is to use the results of these prior studies to provide a framework for an ISD to recognize what events and neural cues are most useful for improving a course.

When the events are initially identified, they are characterized in language as a hypothesis, describing what is significant in terms of the course presentation and how that presentation should be expected to affect the neural cues. The hypotheses define the context and expectations for the neuro-metric behavior consistent with the ISD's intentions regarding how the learning object should perform – neuro-metrically.

This is the chief function the ISD needs in a system; a way to develop a narrative of behavior around metrics that are well understood. While the actual learning process *cannot* be directly observed with current neural science, many of the cues that the ISD wants to manipulate can be. So the narrative that develops the hypothesis needs to be disciplined in hard, established meaning. However, the language associated with the neural cues is hardly standardized. While well-established concepts like attention and emotion may be understood intuitively, newer concepts, like the correlation of neural activity among subjects, are still evolving. Thus, Cross-Brain Correlation, Inter Subject Correlation, and Engagement are all terms in the literature that represent different flavors of this nascent metric.

Another consideration when attempting to build a standardized tool for integrating neuro-metrics into education is the ISD's familiarity with neuro-terminology. It is assumed that an ISD has a deep understanding of the issues involved in course development, but not necessarily a similar familiarity with brain science concepts. This, coupled with the fact that a neuroscientist coach cannot always be on hand, supports the argument that a system providing a selection of standardized language corresponding to the neural cues can provide a significant advantage. To accomplish this standardization, the authors drew from current descriptive language in the neuroscience field often used to categorize metric graphs. In addition, excerpts from papers that describe neural behavior from previous studies can be used as prompts to keep the ISD on track as they characterize events for current and future studies. Additionally, access to tutorials about the metrics would be highly useful in providing the ISD with an understanding of the measures being used. Having this language built into the platform the ISD uses to develop the hypotheses and to express predictions would be instrumental in keeping the ISD in synch with the neuroscience. Also, using standard language allows meaningful comparisons between the hypothesized result and the description of the actual metric observed.

Another benefit of having standardized language is the ability to use more advanced analytics methods. When the language and observations of renowned neuroscientists can be organized in a library for formulating conjectures and reporting behavior, machine learning classification algorithms can be used to attach the metric graphs to the corresponding language. When a new graph is presented for analysis, the appropriate language can be retrieved to compare to the hypothesis. However, the challenge is to incorporate cutting edge research and make it available as an established standard.

Cognitive Instructional Designer Tool Development

The key to experimental design is a methodical process to develop hypotheses about psychological events to be analyzed in the learning object. Initially, the idea is to check off what seems interesting in the course material, but with the intention of developing a narrative of the behavior that is intended by the ISD. It becomes in the ISD's interest to develop courses that can be checked for effectiveness, and to use whatever tools to make that manageable. An essential tool would be a platform that combines the display of the course where it can be tagged for events and coordinated with event descriptions that include the hypothesis definition, while facilitating and enforcing an organized, methodical, scientific structure for results analysis.

Tagging Markers

As stated above, the process begins with marking events in a course that correspond to significant psychological activity that the ISD needs to appraise. Attempts using video cameras and timestamp markers generally have proved unwieldy, although this approach is typically used in one-off neuro-studies that are performed in the constantly paradigm shifting advancement of brain science in academia.

Several methods have been identified to facilitate event tagging. Besides simplifying the manual time stamp to event marking, other automated approaches have been proposed by the authors. For instance, capturing screen images to characterize an event and then matching the stored event image to the passing images in the course display as the subject traverses the content can trigger the EEG recording software to mark the EEG data with annotations corresponding to the events.

Other methods of tagging events include monitoring the HTML/JavaScript activity and trapping on certain actions. New innovations in LMS tracking with both SCORM and xAPI embracing the Learning Record Store (LRS) model will provide standards for more granular tracking of a student's course progress. As LMS culture recognizes (and implements) the need for tracking course navigation, more handles for marking learning events for neuro-tracking and synchronization will become exploitable. The original DoD SCORM initiative recognized the need to monitor student progress in detail, but industry has been slow to implement these recommendations over the past 20 years.

Indicating events on a 20 second epoch scale (Barnett and Cerf 2017), provides a granularity that allows for specific significant neuro-observation and limits the range of stimuli in the learning object to manageable ISD issues. Events in an interactive course, as opposed to a video, will generally occur at different times for different students. So identifiers such as screen shots provide a way to "normalize" the event flow. This can limit the degree of granularity in the event observation of neural data to just screens images or pages. Innocuous markers such as QRcodes that appear regularly embedded in the course images can provide similar markers. Since the QRcode can be readily tracked and interpreted by software, such as the Python Imaging Library (PIL), but are not usually distracting to the learner, it's a fairly unobtrusive method and can be used to track progress through text that would be difficult without distinctive lesson images.

Course Analysis Framework

A number of requirements also came from the vicissitudes of operating quirky technologies to produce AI consumable data. A requirement for the course display and screen capture method came from the difficulty in image pattern matching in the study where screen matching for event mapping was problematic due to web casting idiosyncrasies with image quality and framing. To effectively use images for pattern matching, a consistent video interface is crucial.

The processing of EEG signals produces metrics of certain attributes over time, but these are not simple numbers. They are graphs of behavior that are best described with careful technical language. Attention over a 20 second epoch is not described by a single value but by terms like: attention was seen "to rapidly increase to a midlevel plateau and then fall gradually". To produce useful descriptions of these measures that help an ISD form a hypothesis, where the ISD may not generally have data interpretation skills, the authors have defined a machine learning approach where the graphs for actual epoch metrics are classified with the appropriate academic expression. The pattern classification methodology known as Random Forest is powerfully implemented in Python libraries, e.g., SciKitlearn (scikit-learn.org), to take a series, like the graph of an epoch of a neural metric, and associate it with a label – in this case the technical verbiage. The system is trained by passing examples provided from neural studies that have already been associated with language, such as in academic papers, or in the lab.

Some data normalization is required to cooperate with the algorithm. Epochs must be of the same time length, and kinks need to be removed, otherwise details that are not of interest can dominate the calculations. After the training period is complete, a new graph is passed to the system and the language associated with that graph classification is returned to the hypothesis as text for comparison to expectations.

For the following study, the level of granularity was sufficient for what could be observed from the activity on the screen. Narrative characters appeared in the lesson as part of the course that clearly demarked activity changes, and interview subjects were visible. But where currently we can only follow page changes with images and the html capture, the issue of greater granularity going forward needs to be addressed. Eye tracking camera software can map eyeballs to positions on a screen and could be used to mark where certain content is encountered, bringing greater detail to the event tagging process.

NEURO-STUDY: A HUMAN RESOURCES COURSE

Study Development

In late 2018, the authors conducted a study for a global investment bank using a course provided by the bank's Human Resources department focusing on "Interviewing Skills". The investment bank also provided 29 of their employees to serve as subjects in the study. Each subject completed the course, which consisted of a single lesson and simulation exercise, while undergoing neural observation via EEG. The course included multiple delivery platforms, including voiceover instruction, video lessons, text pop-ups and more extensive PDF files used for both instruction and as exemplars. The simulation involved a mock-interview with two potential candidates for a hiring position, during which various questions were available to interviewers, as well as documentation on the candidates' history. At the end of the interview, subjects were asked to select the candidate they would recommend for hiring.

Before the subjects experienced the course under EEG, the course itself underwent an evaluation phase with an instructional designer. Since no experimental design process (or supporting software) existed at the time to help facilitate and automate this evaluation process, the process itself was quite manual and time consuming. First, the ID identified 44 events within the course that she felt were important. She documented these events, but did not formally hypothesize what level of neuro-metrics (attention, memory, emotion, engagement) should be attributed to each event epoch. When the subjects were run in the lab, under EEG, experiencing the course, their interactions with the course were video recorded (a video camera was mounted behind the subjects, capturing their interactions with the course by recording the screen display). After all subjects were run through the lab, the lab administrators manually aligned the EEG data from each predefined event epoch with the event epoch itself. This required sifting through hours of recorded video content, finding the specific event epochs for each subject (these were asynchronous, as different subjects experienced different events at different times throughout their learning experience) and aligning the EEG data accordingly. Once the alignment process was complete, Dr. Moran Cerf, Nervanix's Chief Scientist, analyzed the data.

It should be noted that, while this study provided very valuable analyses to the instructional designers at the investment bank, with insights on how to improve their course design for efficiency and effectiveness, the process itself could be made dramatically more efficient and effective if a) hypotheses had been applied at the beginning of the process and b) if the identification of the event epochs were automated through software allowing the ISD to easily load the course and tag events with hypothesized neuro-metric outcomes. The courses (SCORM, video, xAPI files) would then go to the lab for neuro-testing, but would already be in a format where EEG data could auto-align to the event epochs. This would remove the tedious and time consuming requirement of video recording each session and manually aligning EEG results with event epochs.

Study Results and Discussion

Notable results of the study included data around Engagement and Emotion demonstrated by the subjects while experiencing the course. The Engagement metric is a more general indication of group behavior, but as the excerpt below referring to Figure 3 indicates, still applies best to specific events like the "start and end of a lesson"

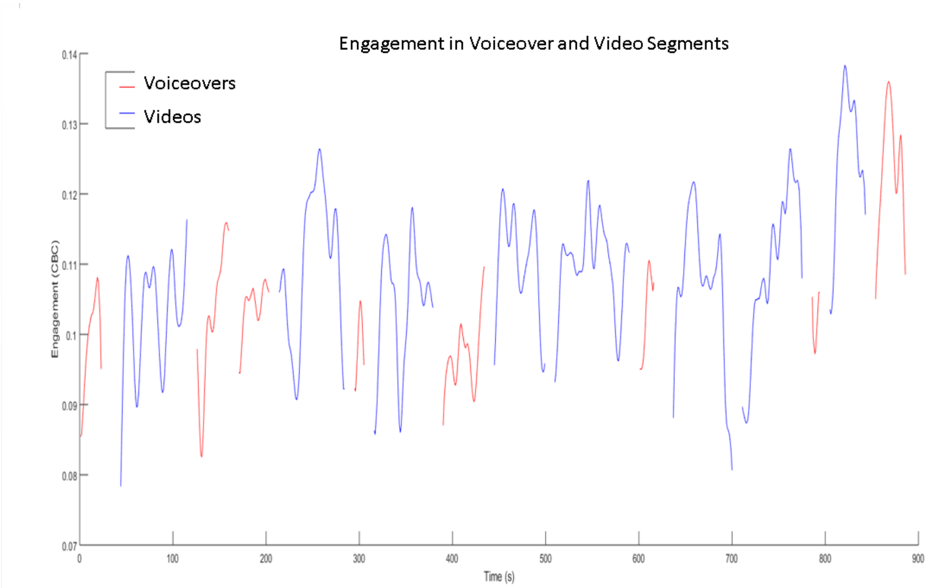

Figure 3. Example of Engagement Data Visualization

As Figure 3 shows, Engagement in the content was relatively constant over the course of the lesson, with slight increases in group-level engagement at both the start and end of the lesson. Subjects were significantly more engaged with the videos, which included novel content, than they were the periods that included voiceovers, which largely provided a segue between different pieces of information. Other typical examples of "Events" in this study would be the beginning of the interview simulation, and the points where the interviewees make certain statements describing their qualifications.

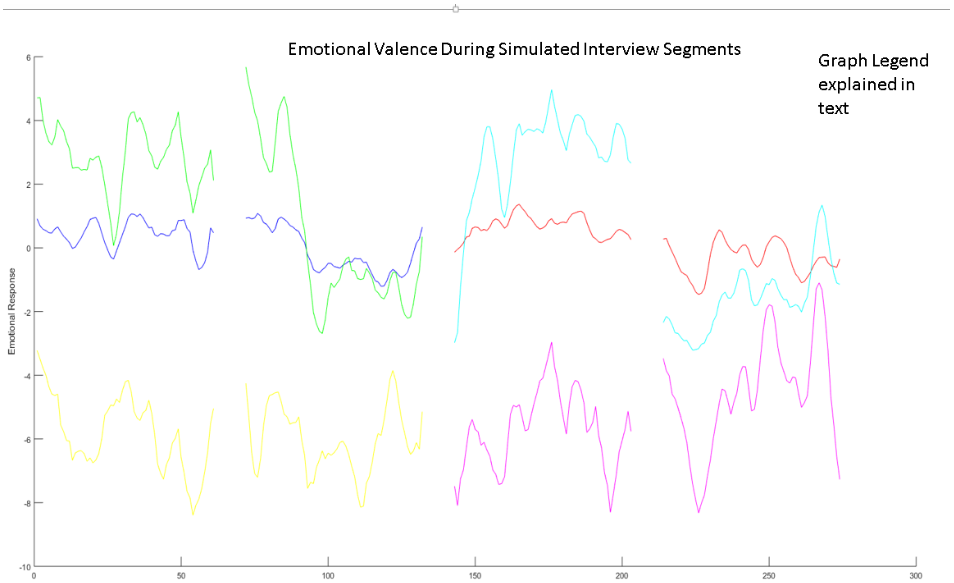

Figure 4. Example of Emotion Data Visualization

As seen in Figure 4, Notable neural differences between the various subject groups were found in the emotional response data. The human resources subjects (green/cyan), those who chose the first interviewee in the interview (yellow/magenta), and the subjects overall (blue/red) were compared. The human resources subjects demonstrate much greater emotional approach relative to the general cohort in the pre-resume period of the interview for both candidates, and while they came to be similar to the cohort overall post-resume, they were initially less approaching than average for the first interviewee and more approaching for the second interviewee. In general, those who made the wrong candidate choice during the simulation were markedly more avoidant than other subjects throughout the entire interview period. Further, when the event behavior analysis took into account the differentiating factor of professional expertise between the subjects, interesting results were obtained that could be relevant to tailoring the course for certain audiences based upon such professional expertise.

The differences between neural responses for human resources subjects and the general cohort were intriguing. The lesson was particularly relevant to subjects who worked in human resources, as they often are responsible for interview processes of the type outlined in the lesson. Though the study cohort was small, they nevertheless demonstrated a unique response style. Specifically, they seemed to be more emotionally attentive to the candidates during the interview period. At the onset of interviews, they were more approaching in both cases. However, subsequent to learning more about each candidate, their emotional response attenuated based on how appropriate the candidate would be for the position. While this is a particular result, it highlights an important factor that affective response can be highly relevant to one's judgments and that this is most effective within an individual's field of expertise, where they also tend to have the soundest emotional judgment. Given that the lesson is intended for those involved in Human Resources, it may be that one could isolate more particular and effective results using subjects exclusively drawn from a pool of such subjects. In high value training situations, being able to anticipate and accommodate different subsets of learners is a decided advantage.

Conclusion

Instructional System Designers need to be aware of a variety of factors when constructing learning material. The content itself is a challenge, of course, requiring careful language to convey accuracy and salient meaning. As well, the presentation medium, graphical material, interactive involvement with course navigation, and lesson sequencing are all up for consideration in a panorama of possible modes and combinations. But this embarrassment of riches requires subtle knowledge of the targeted learner for the ISD to deploy them effectively. It is important to understand what stimulus elicits which response from which kind of learner, in order to effectively tailor instruction and convey information that will translate to knowledge by the learner. Neuro-metrics and other methods to characterize students are essential tools in helping to track and improve learning material, especially where a broad cross-section of student backgrounds needs to be accommodated. One size does not fit all in education, especially when the instruction is narrowly targeted to a specific audience and critical in nature. The requirements to deliver well-prepared learners, replete with critical knowledge, delivered in an efficient and highly effective manner, is especially relevant to the soldiers and operators of the defense mechanisms that protect our citizenry, and is the responsibility of the educational institutions that serve them. Knowledge is power. *Optimizing the delivery of knowledge* builds capacity for power and can help ensure that during training, the warfighter gains the most benefits possible from the learning material.

ACKNOWLEDGEMENTS

The authors wish to acknowledge Christopher Rose, whose contributions helped make this paper possible.

REFERENCES

Anderson, James R. (2019). How to become an Instructional Designer. *eLearningIndustry.com*, *On-line Forum*.

Baker, Monya (2016). Is there a reproducibility crisis?. *Nature, vol 533, no 7604*, pages 452+.

Barnett, Samuel B. & Cerf, Moran (2015). Connecting on Movie Night? Neural Measures of Engagement Differ by Gender. *Advances in Consumer Research, vol 43*, pages 314–18.

Barnett, Samuel B., Cerf, Moran, & White, Hope M. (2016). Keep It Simple Stimuli: Brain-Vetted Elements of Movie Trailers Predict Opening Weekend Ticket Sales. *Advances in Consumer Research, vol 44*

Barnett, Samuel B. & Cerf, Moran (2017). A Ticket for Your Thoughts: Method for Predicting Content Recall and Sales Using Neural Similarity of Moviegoers. *Advance Access publication, January 23, 2017, Published by Oxford University Press on behalf of Journal of Consumer Research, Inc.*

Cerf, Moran (2019 – in review).

Fanelli, Daniele (2018). Is Science Facing a Reproducibility Crisis, and Do We Need To?. *Proceedings of the National Academy of Sciences of the USA, March 12, 2018*

Halle, Sylvain (2017). LabPal: Repeatable Computer Experiments Made Easy. *International Symposium on Software Testing, July 2017*

Harrell, Eben (2019). Neuromarketing: What You Need to Know. *Harvard Business Review, January 23, 2019*

Nervanix Client Study June 2018

Pashler, Harold & Wagenmakers, Eric-Jan (2012). Editor's Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence?. *Perspectives on Psychological Science, November 7, 2012.*

Rees, Dianne (2019). How to Become an Instructional Designer. *eLearningIndustry.com, On-line Forum (https://elearningindustry.com/free-ebook/become-instructional-designer/dianne-rees)*

Reynolds, Kenney (2019). How to Become an Instructional Designer. *eLearningIndustry.com, On-line Forum (https://elearningindustry.com/free-ebook/become-instructional-designer/kenney-reynolds)*

Romiszowski, A. J. (1981). Designing Instructional Systems. *RoutledgeFalmer, NY*