Information Need Awareness: an EEG study

Dominika Michalkova, Mario Parra Rodriguez, Yashar Moshfeghi
University of Strathclyde, Glasgow UK

ABSTRACT

A fundamental goal of Information Retrieval (IR) is to satisfy search/-ers' information need (IN). Advances in neuroimaging technologies have allowed for interdisciplinary research to investigate the brain activity associated with the realisation of IN. While these studies have been informative, they were not able to capture the cognitive processes underlying the realisation of IN and the interplay between them with a high temporal resolution. This paper aims to investigate this research question by inferring the variability of brain activity based on the contrast of a state of IN with the two other (no-IN) scenarios. To do so, we employed Electroencephalography (EEG) and constructed an Event-Related Potentials (ERP) analysis of the brain signals captured while participants experiencing a realisation of IN. In particular, the brain signals of 24 healthy participants were captured while performing a Question-Answering (Q/A) Task. Our results show a link between the early stages of processing, corresponding to awareness and the late activity, meaning memory control mechanisms. Our findings also show that participants exhibited early N1-P2 complex indexing awareness processes and indicate, thus, that the realisation of IN is manifested in the brain before it reaches the user's consciousness. This research contributes novel insights into a better understanding of IN and informs the design of IR systems to better satisfy it.

KEYWORDS

Information Need, Information Retrieval, Anomalous State of Knowledge, Memory Error, EEG, ERP

1 INTRODUCTION

The goal of Information Retrieval (IR) is to satisfy searchers' information needs (IN) [1, 2]. Many researchers in the past have tried to better understand this complex phenomenon [3–7] to satisfy it. Notable examples of these studies include Taylor's Question-Negotiation Process [3], Anomalous State of Knowledge Model by Belkin et al. [4], Wilson's Information Seeking Behaviour [5], Kuhlthau's Information Search Process [6], Ingwersen's psychological and cognitive aspects of IR [8] and Cognitive IR Theory [7], Cole's framework of information processing and knowledge frames [9] and many other conceptual analyses [10–15]. Despite their invaluable contributions, they all have investigated the IN phenomenon indirectly, via traditional methods of user behaviour, i.e. some sort of mediator, such as questionnaires/interviews [16], reflective diaries [6, 17], user-system interactions and search logs [18]. Past studies have shown that IN phenomenon is often difficult for the searchers to understand and describe [19], leaving current user-base studies limited in capturing and describing how exactly such a phenomenon appears at its very early stage.

, , 2021
2022. ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

To tackle this problem, new research has emerged employing neuroimaging techniques to directly monitor brain activity of subjects experiencing a realisation of IN [20–22]. This research has led to a discovery of cognitive processes and brain regions associated with the realisation of IN [22, 23]. However, these studies have used fMRI technique which has a high spatial, but low temporal resolution. Thus the question of "how a realisation of IN occur from a temporal aspect?" has still remained unanswered. This study aims to investigate the possibility of capturing the realisation of IN in real time from the brain signals.

We aim to go one step further and investigate the possibility of capturing the brain activity of searchers experiencing a memory error (or gap) [24]. This is the situation where searchers are under the assumption that they know (or have recalled) the information for a task or a question, but the recalled information is, in fact, incorrect. In such a scenario, an individual would not experience a realisation of IN and hence would not engage in an information retrieval and seeking process, leaving the memory error unresolved. Depending on the context, this could have negative consequences for the user (e.g., in an education context) and, in some cases, catastrophic ones (e.g., in a medical context).

It can be argued that in a situation like this, where an individual is not aware of an anomaly in their state of knowledge, IR systems could be helpful. One way could be to make users aware of their memory error. This could lead to a realisation of IN, an initiation of a search process and possibly prevention or correction of a mistake. Another way would be to alleviate the information overload situation by not flooding users with information when they have a correct memory retrieval/recall and they do not need information. This is particularly important in research directions such as proactive IR [25] and/or personal assistant systems [26–28]. However, this would be possible, if users' memory errors could be captured. This work aims to investigates the possibility of detecting situations where searchers experiencing a memory error in real time from their brain activity. The question then would be how different the brain signals of memory error would be from the ones when a realisation of IN occurs? And, how different they would be from the ones when a correct information retrieval/recall occurs? This lead us to the following research questions (RQs):

(1) **RQ1:** *"What are the temporal dynamics of the neural manifestation of 'IN realisation' scenario?";*
(2) **RQ2:** *"Is there a clear, detectable, neural manifestation of 'memory error' and 'correct recall' scenarios? and if so, what are the temporal dynamics of these manifestations?"; and*
(3) **RQ3:** *"Do neural manifestations of these three scenarios differ?, i.e. when searchers realise an IN (Know that Don't know and want to know scenario), have a memory error (Don't know that Don't know scenario), and have a correct recall (Know that Know scenario)?".*

Our study utilises an experimental design involving an interactive Question-Answering (Q/A) retrieval task. The information processing at participants was induced by questions of general knowledge while their brain activity was recorded using Electroencephalography (EEG) in real-time. Depending on the answer the participants provided with each question, we encoded the neural processing into three pre-defined Scenarios: 1) Correct, meaning participant selected a correct answer (CorrectRecall), 2) Incorrect, meaning participant falsely endorsed the incorrect answer choice as the correct (MemoryError) and at last, 3) Not Know, representing recognised and acknowledged ASK with an interest to learn the answer (NeedToSearch, i.e. representing the expression of IN). We focused on the inspection of the modulation of ERP components. These represent stimulus-evoked voltage deflections as these are seen as an indicative of unfolding of cognitive processes. We developed a data-driven Event Related Potentials (ERP) analytical framework contrasting the time segments of significant spatial neural activity associated with each of the three states. Our central aims are to: (i) evidence on the distinct neural correlates to project a state of INs (in particular in an early state) with the focus on its temporal dynamics and (ii) evaluate if the brain activity is significantly modulated depending on the distinct states of knowing (in the context of our three scenarios). Specifically, answering (i) can be the basis to detect IN in a Q/A task in real-time. This study accounts for neurophysiological objective inputs to IR systems which might to lead to improved (even proactive) IN detection.

2 RELEVANT WORK

Complex Nature of IN. IN is a core concept of the IR process of delivering relevant documents and satisfying user's IN [29]. Its subjective and user-dependent nature makes it an complex input to IR [11]. Multiple studies have attributed IN to an anomaly in knowledge [4, 30], uncertainty [6], feelings of unease [9], feeling of dissatisfaction [19] or doubt [3]. Taylor described the complexity of IN on a hierarchical model of IN in respect to user's awareness with their IN. Whilst Taylor pointed out that low levels of IN are visceral and within-brain attributes, Belkin et al. [4] specified that the ASK variants are dependent on the levels of user's knowledge. Both of these anticipate the inherent complexity in how the user might perceive their INs. This paradox influences the quality of main elements in the IR process, most notably the formulated query as the system-representation of IN. The situation becomes even more pronounced when IN is ill-defined [14]. The user receives only vague, uncertain and fuzzy memory retrievals [9]. It makes it all the more challenging for the user to map their actual (inexpressible) IN to a (expressive) query [3].

The IN were empirically investigated via self-reported mechanisms or utilisation of search logs [6]. Recently appealing is the area of contextualisation of the INs, to derive information such as search intent [31]. For instance, utilising additional search metadata, e.g. real-time contextual activity of the mobile queries [32–34] shows positive results in this area. These scenarios however evaluated IN only indirectly which makes it difficult to create a link between the actual INs [3] and their system formulation. Interdisciplinary*NeuraSearch* research is an empirical response to the overcome this issue by evaluation of the user brain responses as an alternative source of user behavioural patterns [35].

Furthermore, a rapid development of intelligent IR engines, such as conversational agents [36, 37] proceeds to explore the proactive design-thinking of IR [38]. The potentials of proactive zero-query systems [25] might benefit from the NeuraSearch approach and address the "IN-query gap" [20]. In a sense, the anticipation of INs and proactivity of IR is an intervention with the user's awareness. Utilising the link between awareness and knowledge anomaly, could proactive IR system alter the user's awareness and even rectify user's unawareness? The investigation of memory error that has not been recognised by the user is, thus, a novel attempt to bring a new perspective the concept of ASK and expand the discussion in the area of proactive IR systems.

Neuroscience & IR Research. Despite traditional behavioural studies concerning INs offer beneficial insight into the qualitative characteristics of IN [39–41], their capabilities are limited in the area of data acquisition of the internal (within-brain) formation of INs as well. To operationalise these efforts, a branch of interdisciplinary research within IS&R has been established under the term *NeuraSearch* [42]. NeuraSearch is an evidence-based research employing a carefully designed user-studies with a real-time acquisition of neural manifestations (i.e., activity) of user experiences in IS&R [43, 44]. The employed neuroimaging techniques advance to evidence on the location of the source of the activity and to create a fine-grained temporal profile of the activity strengthening the link with the cognitive theoretical approaches in IS&R [7, 8, 45].

A rise in adoption of neuroimaging techniques across the key areas of IS&R resulted in numerous *NeuraSearch* applications with the most notable being relevance judgment [46, 47], IN realisation [20], search transitions [23] or query construction [48]. Commonly used neuroimaging techniques include electroencephalography (EEG) [46, 48–52], functional magnetic resonance imaging (fMRI) [20, 22, 49], magnetoencephalography (MEG) [53], functional Near Infrared Spectroscopy (fNIRS) [54, 55]. The EEG studies manifested the effectivity of EEG to capture the brain activities of complex cognitive processes utilising its fine-grained temporal resolution. EEG technique allows to capture even the smallest variability between the graded levels of the same phenomenon, e.g. three ERP components, P300, N400 and P600 were modulated depending on the grades of relevance judgmens [47]. The current development and availability of portable EEG devices strengthens its applicability as the signal input into Brain-Computer Interfaces (BCI) systems [56].

Conceptually relevant and informative to our research is a series of fMRI studies investigating IN realisation developed by Moshfeghi et al. [20, 21]. fMRI technique has the advantage of highly precise spatial resolution to locate the brain source of significant activity within mm. The first study [20], thus, identified topological brain maps of the significant activity found critical hubs orchestrating activity between IN and noIN states. Second study [21], positively confirmed the predictive capabilities of fMRI data. The series culminated in the introduction of Neuropsychological model of IN realisation with three components: Memory Retrieval, Information Flow, A high level Perception component [22]. Specifically the first two are supported by discoveries of functional networks involving factual searches in memory and self-awareness context, as the sources of cognitive mechanisms involved in the user's adaptability on a new stimuli and support the user's readiness, e.g. whether there is a need to search or not [9]. In these works, IN was reduced to a binary concept, represented a state of knowledge in which users felt a higher probability of not knowing the answer than knowing the correct answer. No deeper investigation concerning and potentially diversifying the user's state of the knowledge, was done. However, as the task involved user's access to knowledge, the authors anticipated that arise of IN was connected to a feeling of anomaly, and thus, connected IN to ASK [4]. There is not yet known to us a study focusing on investigation of temporal signatures of IN within IR context using EEG. We hypothesise that the mental formation of the realisation of not-knowing based on measured brain signals provide a reliable information about the variability of neurocognitive processes and might lead to a potential detection of IN on the visceral basis.

3 PRELIMINARIES

Electroencephalography (EEG). Brain generates small electrical activity, on the order of millionth of a volt. The EEG represent a procedure how to capture this activity in real-time using EEG's main advantage of high temporal resolution [57]. Time frequency (or sampling rate) is the measure of data sampling (e.g., 500 Hz/s). A sensor, i.e., electrode, is attached to a specific location of the skull surface to capture the cortical electrical activity. A common approach is to use an "EEG cap", made of an elastic light-weight fabric, as the recording interface on top of which multiple electrodes attached to in order to get a signal originating different areas on the brain. The electrode configuration on each cap follows the standardised "10:20 System of EEG Placement" [58] as depicted in Figure 2 which allows comparable results of spatial activity between subjects within a study as well as between the studies. The acquired data must go through a data pre-processing pipeline (see Section 4.1 to increase the Signal-to-Noise Ratio with the noise imposed by "the artefacts" originating from inside the body, such as heart activity, eye blinks, eye movements, facial and other muscle movements amplifying the signal.

Event Related Potentials (ERP) is a common approach of analysing EEG data based on averaging EEG response waveforms, time-locked to stimuli onset (start) and offset (end), across people and trials. ERP Component then represent a deflection from the baseline of EEG activity denoting neural activation or deactivation linked to some form of neural processing. The standard convention for labelling ERP components is to use: 'Letter' which refers to the polarity of the component: Positive (P) or negative (N) and 'Number', which refers to the time point on the series where this deflection reaches its local maximum, e.g. component N1 meaning negative deflection peaking at 100 ms post-stimulus. **Data Epoch** is the time window within which the relevant brain responses, i.e. ERP Components, are expected to emerge. In our experiment, a typical epoch expanded from 200 ms pre-stimulus onset to 800 ms post-stimulus onset to cover the stimuli presentation time and baseline activity. **Time Window** is a specific selections of the time series, i.e. epoch, that captures particular ERP components. This allows investigation of the temporal dynamics of the relevant brain responses. **Region of Interest (ROI)** represents a set of neighbouring electrodes that jointly contribute to a particular ERP component linked to the investigated function. The topological distribution of ROI (Figure 2) allows for investigation of the spatial dynamics of relevant brain responses. **ERP Analysis** focuses on extraction and analyses of the amplitude of an ERP component to contrast between the given experimental conditions in a study. In general, the amplitude of an ERP component reflects the voltage elicited by the neural activity giving rise to the ERP component. The higher the voltage (higher negativity or positivity) the higher will be the ERP amplitude indicating a greater amount of neural resources (activity) recruited to support the specific neural operation [59].

4 METHODOLOGY

Design. A "within-subjects" design was used in this study, in which the participants performed a Q/A task (see Section 4.1.1 to explain the task). The main aim of the experiment was to evaluate the modulations in the brain activity posed by the three pre-defined states of knowledge (Scenarios). These were linked to the participants responses which were controlled by responding to questions viewed on the screen. These represented the independent variable (Scenarios) in the study: "Correct" (CorrectRecall), "Incorrect" (MemoryError) and "I do not know" meaning acknowledged ASK. The level "I do not know" was further subsetted according to the participant follow up decision indicating either i) care to know the answer (NeedToSearch) or ii) no care to know the answer (NoNeedToSearch). By limiting the response space to three levels we had a control over the categorisation of participants and the information processing associated with these levels. For each level, we extracted the relevant ERP activity (refer to EEG Glossary in Section 3). The dependent variable was the mean amplitude of relevant ERP components drawn from the EEG signals synchronised with the Q/A task.

Participants. Twenty-four healthy university students volunteered. Initial insight into data showed that data were unbalanced across our Scenarios (see Table 1). As the methods of the analytical framework we devised (see Section 4.2) rely on individual averages, we need a sufficient samples in each Scenarios. We, thus, filtered the data where only participants satisfying a threshold of minimum number of responses in each level will be moved to final sample and undergo the statistical analysis. The criterium was satisfied by fourteen participants (see Section 4.2.2), out of which there were 13 females (93%) and 1 male (7%) within an age range between 18 and 39 years and a mean age of 23 years (sd 6.5). The participants were volunteers recruited by the **anonymised university** and they received no monetary payments, but were eligible for academic credits. Participants completed the task (without the breaks) on average in 44 min (sd=4.62, med=43.40).

Q/A Dataset. We constructed a dataset of general knowledge questions taken from the following sources (1) TREC-8 and TREC-2001[1] (widely applied in IR and *NeuraSearch*, such as [20]) and (2) B-KNorms Database[2] (used in cognition and learning studies [60]). Two independent assessors separately evaluated the questions difficulty (Cohen's Kappa: 0.61). We then selected a subset of 120 questions [3] where both annotators agreed upon their difficulties. Data were equally distributed between easy and difficult (60:60). Its purpose was to create balanced trials for the participants to experience scenarios of knowledge anomaly corresponding to

[1] https://trec.nist.gov/data/qamain.html
[2] http://www.mangelslab.org/bknorms
[3] The data will be available upon request.

our three Scenarios. We expected easy questions would trigger more of KNOW (either CorrectRecall or MemoryError) and difficult questions more of “I do not know", creating thus the premise of IN realisation. Here are two examples of a) a difficult question from the dataset: “What is the length of the coastline of the state of Alaska?" and b) an easy question" “What primary colours do you mix to make orange?". Questions covered a diverse range of topics: History, Science and Technology, Geography, Culture and Art, General. The questions were of the open domain and closed-ended answer. The question length (measured by the number of words the question consisted of) resulted from 3 words to 13 words. The question attributes (length, difficulty, category) were not used as an independent variable in our main analysis as we did not investigate their effects on cortical activity. Their impact was however tested using final data distributions, with no significant effects. Due to space limitations we do not further report on the outcomes.

4.1 Procedures

Ethical permission to carry out this study was obtained from the **anonymised** Ethics Committee. The experiment was performed in a research laboratory setting. All participants fulfilled the inclusion criteria to take part in the study, i.e. healthy people between 18 – 55 years, fluent in English, without any prior or current psychiatric or neurological Scenariosthat could influence EEG signal.

Participants were first distributed two questionnaires requesting their demographics information and inquiring about their habits with information searching and search engines. They were then informed about the task (described in Section 4.1.1). To ensure that all the participants had a good understanding of the task, they underwent a practice session. The practice session consisted of five questions not used in the main task. This section was not limited in time, and participants were allowed to repeat it, if required, until they felt comfortable to proceed to the main session. In the main session the question order, as well as the answer options on the screen, were randomised across participants. The randomisation allowed to eliminate any order bias and response tendencies. There was no time limit to provide responses, which were entered by the participants via a button press using three keys previously allocated to each option. Two breaks were provided (after completion of 1/3 and 2/3 of questions) to avoid fatigue. Participants were asked to remain still and minimise any body movements, particularly blinking, as these represent noise-introducing artefacts to the brain signal.

Recording Scenarios were kept constant across participants. After the main session, there was a debriefing session and participant were required to fill out a final post-task questionnaire related to their subjective perception with the task. A prior to data collection, a pilot study with two volunteers was conducted to ensure that the experimental procedure ran smoothly. Feedback from the participants in the pilot study was used to improve the procedure. Feedback from the participants in the pilot study was used towards the procedure improvements and data collected in the pilot study are not used in the present analysis.

4.1.1 Procedure of the Main Experiment Task. Figure 1 illustrates the Task sequence with an example of a 5-word question stimulus. Each participant was subjected to 120 trials. Every trial followed

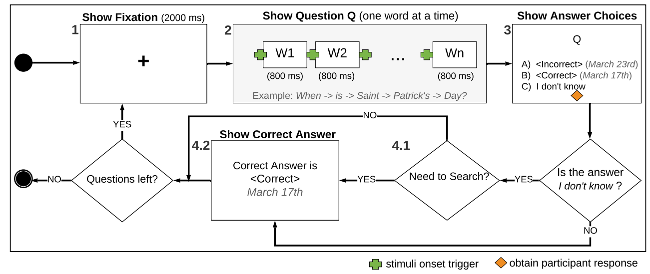


Figure 1: Diagram of the task flow.

the same order of steps: Step 1 Showing the Fixation Cross, Step 2 *Question Presentation*, Step 3 *Presentation of Answer Choices* and Step 4.1 *Search Decision*. The brain activity was captured during the entire course of the trial, but the core section of data represent the data captured in Step 2 of Figure 1). Here we captured the brain activity of participants during the sequential sentence reading in order to evaluate the brain correlates associated with the processes of information processing. After a random question ran, a participant recorded their response and the associated three Scenarios were added to each word of the question. Each word of the question then became a part of the same Scenarios. We did not differentiate between separate words, e.g. due to their qualitative criteria, or extracted brain activity with particular words in the question. Our work is similarly as in [47] built on the approach of knowledge accumulation. As the word sequentially appears on the screen, the question progresses corresponding to the amount of received stimuli (input) information.

Trial. The trial started with viewing a fixation cross in the middle of the screen for a duration of 2000 ms that indicated the location of the next stimuli on the screen and was a way to minimise eye movements on the screen. Next, the participants viewed a sequential presentation of a question randomly selected from the dataset. Each word was displayed for 800 ms. Within this step, the participant processed the information as it was coming (word-by-word). After the last word (Wn in Figure 1) of a generated question ran, the participants moved to Step 3. Here, they were presented with now fully-displayed question and three on-screen answer choices associated with the question. They were requested to select the correct answer or choose the option “I do not know". Next step depended on the response outcome in Step 3. In the case participants answered CorrectRecall or MemoryError, they were presented with the correct answer (Step 4.2) which they terminated by a button press after which they were moved on the beginning, Step 1. Alternatively, if the participants answered “I do not know", they were asked to make a Search decision (Step 4.1), whether they want to search (NeedToSearch) for the correct answer or not (NoNeedToSearch). The ‚search’ was simplified here, as no actual search was performed as part of the study task, with participants being aware of it. Search interaction would increase the EEG artefacts onto data, such as introduction of motor-related artefacts. Search represented here the decision of participants whether they want to learn the correct answer (move onto Step 4.2) and, thus, satisfy their “I do not know" Scenarios(NeedToSearch) or not (NoNeedToSearch) and move into the next trial. The process was repeated for all 120 questions.

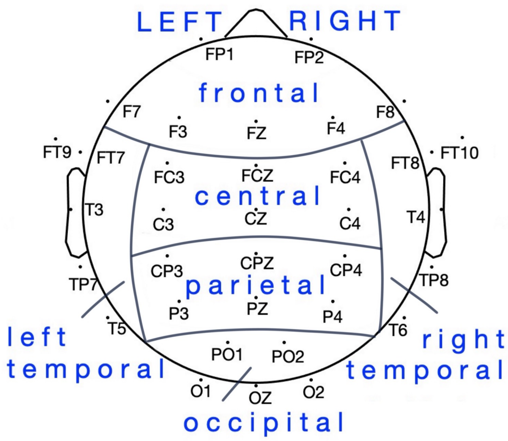


Figure 2: Placement of electrodes on EEG cap and their regional assignment (ROI).

Experimental decisions. **Sequential presentation** of a stimulus has been applied in studies examining neurological correlates of reading [61], as well as in IR-related studies of relevance [47] or query construction [48]. Its aim is to have control over free-viewing and to minimise presence of any confounding artefacts (i.e., saccades). As a result, the ERPs were time-locked to the word onset presentation, i.e. we captured the temporal processing of each stimuli ERPs during the preseWords on the screen were presented using a **Fixed Rapid Serial Visual Presentation** (fixed-RSVP) of 800 ms for each word. As previous studies of sentence processing [62, 63] pointed out, a ratio above 700 ms enables engagement of higher cognitive abilities. These are important for us as we want to capture the human information processing and to extract EEG signal associated with particular words, i.e. events. The final ratio of 800 ms was determined by the outcomes pilot study. It was found sufficient for fluent reading [63] and to avoid the overlapping effect of two consecutive words on the ERPs [64]. To determine **IN expression** we used the information related to participants decision we obtained during the study. We, thus, specified the "I do not know" level with an added decisions of participant, i.e. the want to search (i.e., NeedToSearch) and, thus, created the level NeedToSearch. We translated this level as an expressed form of participants' INs.

Apparatus. Interactive Q/A system which ran the task (Section 4.1.1) was developed in a behavioural research software e-Prime2 and was synchronised with an EEG system. A 40 electrodes NeuroScan Ltd. system with a 10/20 configuration cap was used for data acquisition. EEG was recorded with a sampling frequency of 500Hz. Impedances were kept below 10 kΩ and signals were filtered online within the band of 0.1 - 80Hz. EEG recordings were subsequently pre-processed offline using toolbox EEGLAB version 14.1.2 [65] executed with Matlab R2018a. Further stage of statistical analyses was done in RStudio with R 3.6.1.

EEG Data synchronisation and Reduction of artefacts. EEG signals were time-locked to the word presentation. In order to associate EEG recordings with behavioural data, we applied synchronous triggers, i.e. unique identifiers for each trial. Triggers were sent to a separate file at the onset of each word (as depicted by a green cross icon in Figure 1) and at the time of the button press indicating the response (depicted by an orange square in Figure 1). As was reported, the triggers corresponding to Scenarios encoded the sequence of EEG signal according to one of those levels. By design we eliminated any possibility of signal contamination by neural correlates corresponding to motor responses (i.e. hand movement to make a button click) as response-preparation component as it generates a separate signal (often called noise) affecting the underlying brain signal [59]. User read the textual stimuli on the screen following the order of word in a question. Only after the entire question was presented, the participant was asked to make an explicit judgment.

4.2 ERP Analytical Framework

Evaluation of conditional neurocognitive response is our main objective. Initial insight into ERP waveforms showed a clear manifestation of ERP components (see Section 3), which resulted in a decision to apply a combination of exploratory and component-driven approach, focusing on evaluation of the ERP deflection within specific time windows where such ERP deflection occurs. Spatio-temporal investigation of neural activity requires unbiased selection of i) relevant spatial regions where the activity is significant and ii) splitting of the overall timeline into smaller time windows. We framed a procedure that allows us to achieve unbiased results. Its details are provided in Section 4.2.3 and 4.2.4.

4.2.1 Data Pre-processing Pipeline. First, the individual participant data were pre-processed using the pipeline constructed according to guidelines for the standardisation of processing steps for large-scale EEG data [66]. Before passing the data through a high-pass filter at 0.5Hz and then through a low-pass filter at 30Hz, we removed the power line noise at 50Hz. Next, we down-sampled data to 250Hz. We then proceeded to reconstruct a low-quality EEG signal of selected electrode/-s whose recordings were found to be very noisy or their recording was interrupted. After that, we used average re-reference.In the next step, we performed Independent Component Analysis (ICA), a data cleaning technique used to separate the noise-introducing artefacts (see Section 3) from the genuine brain signal, i.e. the sole brain effect to stimuli. Completion of ICA resulted in an artefact-free signal which was then epoched (see 3), from 200ms pre-stimulus presentation to 800 ms post-stimulus. All epochs were baseline-corrected using the -200 to 0 ms window using the baseline activity from prior to the onset of the first word of each question (represented by Screen 1 within S1 in Figure 1). It is used to remove DC-offset or in other words to compensate for signal drifts in electrophysiological recordings [67]. After that, for every participant we averaged all epochs which belonged to one of the three Scenarios (CorrectRecall, MemoryError, NeedToSearch). These average participant data entered the further stages of the statistical analysis.

4.2.2 Sample Size Pre-processing. As a second step, we explored the distribution of the responses (see Initial column in Table 1 to uncover potential issues with the existing data that might prevent us from the application from the further stages of data analysis. As our methods rely on within-subject methods and ERP averages, we need sufficient samples in each Scenario per participant.

Table 1: Distribution of responses per participant before and after filtering.

Scenarios	*% of All Trials (Responses) per participant*	
	Initial (24 participants)	*After filtering (14 participants)*
CorrectRecall	56	50.50
MemoryError	25	21
NeedToSearch	15	23.50
NoNeedToSearch*	4	5

* Excluded from analysis due to low sample size

NoNeedToSearch was the least frequent response across among the participants. It was excluded from the further analysis due to very low sample size across and within participants (only 4% of responses per participant). Next, on average each participant's NeedToSearch responses accounted only up to 15 % of overall responses.

We checked the individual records for each participant, sorted them in ascending order by the number of records per each Scenarios. The participants with the lowest proportion responses were selected and subsequently visualised using ERP analysis. The stimulus triggered activity, i.e. ERP, emerged when averaging 12 trials (representing 10% of all trials participants were subjected to), which we then set as our threshold. Altogether 14 participants satisfied this condition for each Scenarios. Table 1 shows the original distribution of full dataset responses (24 participants) across Scenarios and updated values after the filtering was applied.

The comparison of distributions between NeedToSearch and NoNeedToSearch responses points out to a strong preference of participants wanting to resolve their gap in knowledge, manifesting as a high amount of NeedToSearch responses in contrast to NoNeedToSearch responses. We retrospectively learnt from the exit questionnaires that the questions mostly evoked curiosity and interest in learning the correct answer. We select few of such responses that support this conclusion:

"I responded positively in every one because I found it as an opportunity to learn."

"I like finding out new information. Having that option was really good it kept you interested even when you didn't knew the answer."

"Genuine interest. The information may be useful in future."

4.2.3 Identifying Regions of Interest (ROI). Literature [59] suggested to deal with both the selection of time windows and ROIs by separating statistical tests for each time point at each electrode, combined with correction for multiple comparisons. Following this approach, we searched for significant differences across Scenarios with a combination of the 2-sample paired Monte Carlo permutation test and non-parametric bootstrapping running 10,000 permutations[4]. Outcome of each pairwise comparison was a set of significant (p<0.001) electrodes and their assigned time point where the activity significantly differed. We assigned these electrodes into clusters (ROI) based on their spatio-temporal properties, i.e. local proximity (according to Figure 2) and significance within the same time window.

[4]A solution for multiple comparison problems and does not depend on multiple comparisons correction or Gaussian assumptions about the probability distribution of the data.

4.2.4 Time Windows. To ensure the decision of setting the boundaries of time frames capturing ERP components is not arbitrary, we used unbiased data-driven procedure we now describe. We averaged all the epochs corresponding to the three Scenarios across all ROIs and participants. This reflects the overall brain responses regardless of task Scenarios and topological distributions with baseline set y=0. From these grand waveforms, we selected the time points where the waveforms, i.e. ERP components, abandoned and returned to baseline. We then calculated the mean signal within the intervals for each relevant ROI.

4.2.5 Statistical Methods. Our main comparative measure is the mean signal, calculated as the mean of the ERP activity, precisely the amplitude that describes the ERP activity, which occurred within particular time window (see 4.2.4) and significant ROI (see 4.2.3). We created a mixed linear model for each time window with the parameters: "Scenarios" as the independent variable, "Participant" as the random variable, "Mean signal within ROI" as the dependent variable. In order to test if and how much each participant's neurocognitive response (amount of potential elicited) varied across the Scenarios, we applied ANOVA repeated measures test with 3-levels factorial design for each time window and for each significant ROI. Data met the assumptions required by ANOVA.

5 RESULTS

Task Perception. Participants perceived the task mostly as Interesting (58%). In terms of the difficulty, the task was perceived as relatively Easy (54%), followed by perceptions of some degree of difficulty (Not so Easy 21%, Slightly Difficult 17%). One participant (4%) perceived it as Difficult. Those who perceived some degree of difficulty also found the task to some degree Challenging (21%). In general, the task was not perceived as Stressful (4%) nor Familiar (4%). We pick a few of additional participants' comments which expand on their perceptions with the Q/A dataset. In general, participants agreed that Q/A dataset was an appropriate mix of easier and more difficult questions of general knowledge:

"Topics varied widely which was very interesting as there was quite a mix of things I knew and things I did not."

"Some answers I thought I knew and I was wrong and vice versa."

Main Findings. The results show at first, a separation of the activity into ERP components: N1, P2, N400 and the late positivity of P6 across all Scenarios. The main findings of significant pairwise ERP modulations are presented in the the Table 2. First column *Time Window* shows the temporal intervals where ERP components occurred (*ERP*). The column *ROI* specifies the location where the mean of the corresponding ERP shows statistical significance. The column *F value* quantifies the value of the statistical test with (2,26) degrees of freedom. The column M_{diff} compares the mean values of two significantly different Scenarios identified by the pairwise post-hoc tests. At last, *p-value* shows the level of statistical significance. The Figures 3 and 4 complement this table by illustrating the corresponding ERP waveforms at the significant ROIs. ERP

Table 2: Significant differences in ERP amplitudes and the pairwise contrasts (p-value adjusted using Bonferroni corrections <0.05 *, <0.01 **, <0.001 *)**

Time Window	*ERP*	*ROI**	*F value*	M_{diff}	*p-value*
90 - 150 ms	N1	RF/RFC	F[2,26]=3.50	$\bar{x}$(MemoryError)= -0.02 μV, NeedToSearch)= -0.16μV	*
150 - 270 ms	P2	LTP	F[2,26]=3.97	$\bar{x}$(MemoryError)= -1.93 μV, $\bar{x}$(NeedToSearch)= -1.70 μV	*
270 - 430 ms	Onset N400	RFT	F[2,26]=5.93	$\bar{x}$(CorrectRecall)= -0.39 μV, $\bar{x}$(MemoryError)= -0.73 μV	**
		C/CP	F[2,26]=4.83	$\bar{x}$(CorrectRecall)= -0.10 μV, $\bar{x}$(MemoryError)= 0.04 μV	*
430 - 570 ms	Offset N400	RFT	F[2,26]=4.69	$\bar{x}$(MemoryError)= -0.23 μV, $\bar{x}$(NeedToSearch)= -0.08 μV	*
		PO/O	F[2,26]=5.19	$\bar{x}$(MemoryError)= 0.29 μV, $\bar{x}$(NeedToSearch)= 0.08 μV	**
570 - 800 ms	P6	-	-	-	-

* (For spatial reference of ROIs see Figure 2): L - Left, R - Right, F - Frontal, C - Central, T - Temporal, P - Parietal, O - Occipital.)

waveforms of each scenario are encoded by a specific colour and a marker. The scales are uniform for all corresponding plots. Following the chronological latency of the ERP components, we now describe the main findings.

N1-P2 component. ERP modulation was found to emerge already before 100 ms poststimulus exhibiting N1 component. We found right frontal/front-central distributed activity (RF/RFC) discriminating between the mean levels of Condition. Post-hoc test specified that the difference is driven by a significantly greater negativity of N1 exhibited for NeedToSearch in contrast to MemoryError. Further, the amplitude of P2, with the latency between 150 - 270 ms, over left temporo-parietal (LTP) ROI was found to be significantly affected by the different processing for Scenarios. Here, a significantly greater negativity of P2 emerged for MemoryError level in contrast to the mean amplitude of NeedToSearch.

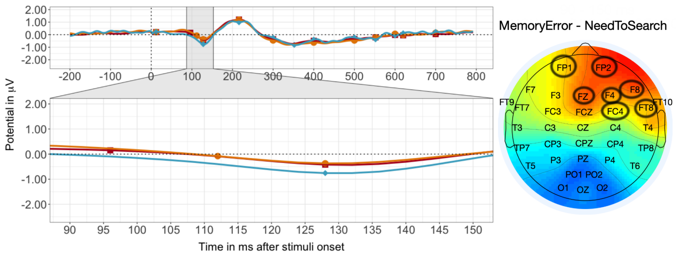
N400 component. The long amplitude of N400, with a latency between 270 - 570 ms, describes: 1) the onset of the window describing the accumulation of the resources to support a cognitive operations until their peak 2) followed by a return of the amplitude to the baseline of values after the cognitive decision was made. We, therefore, split the time window into two frames: (1) 270 - 430 ms comprising the onset of the amplitude and (2) the window 430 - 570 ms with the activity offset and, thus, to explore if any differences are involved in these two time frames. The findings support this decision. In the (1) window 270 - 430 ms we found two ROIs where the activity was significantly modulated by the Scenarios: a) right front-temporal (RFT) ROI and b) central/centro-parietal (C/CP) area. Post-hoc tests specified that in both ROIs the signal significantly varied between the pair CorrectRecall and MemoryError. Whereas, in (a) the negative amplitude of N400 onset was significantly larger for MemoryError, in (b) a larger negativity was measured at CorrectRecall level. In the (2) window we found statistically significant contrast over different ROIs. First, the RFT ROI, similarly to findings from earlier, demonstrates the significant differences across the Scenarios. Post-hoc test, however, alters the previous pairwise findings. Here, the significant difference is driven by a pair of MemoryError and NeedToSearch levels with the mean amplitude for MemoryError having a significantly greater negativity. The same direction was confirmed for the second ROI, parieto-occipital/occipital (PO/O). Pairwise post-hoc tests revealed a significantly greater amplitude of MemoryError over NeedToSearch. The distribution of ROIs at the offset of N400, thus, suggests anterior-posterior pattern with the amplitude for MemoryError being significantly higher than that of NeedToSearch.

P6 component. At last, we lacked to find any significance in the window spanning the P6 component that would evidence on the significant effects driven by the Scenarios.

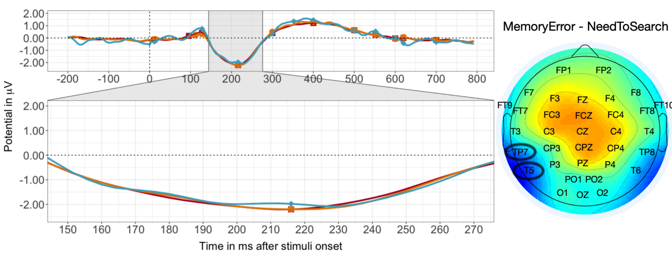
Summary. In summary, the data evidence on the identifiable spatio-temporal patterns for NeedToSearch, as the expression of IN, addressing **RQ1**. Specifically it found to elicit 1) significantly amplified negativity of N1 component in RF/RFC region relative to MemoryError and 2) relative to MemoryError, significantly lower amplitude of P2 over LTP and of the offset of N400 over both RFT and PO/O ROI. Furthermore, 3) the onset of N400 is not modulated by NeedToSearch and 4) the latest activity of P6 is not altered by NeedToSearch. Furthermore, addressing **RQ2** and **RQ3**, the evidence indicates a quantitative separation of the MemoryError level as manifested by the increased 1) amplitude of P2 in LTP ROI, 2) amplitude of the onset of N400 over RFT ROI against CorrectRecall and 3) amplitude of the offset of N400 over both RFT and PO/O ROI relative to amplitude of NeedToSearch.

6 DISCUSSION

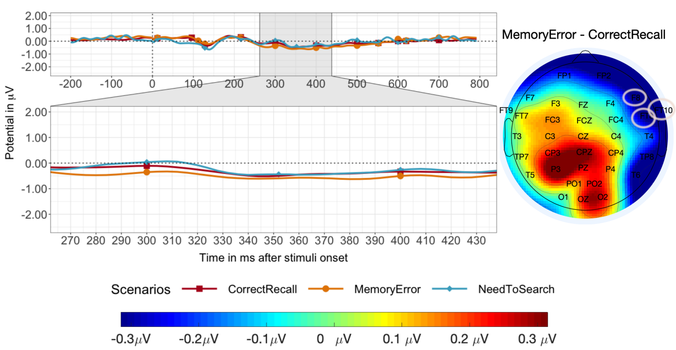
Before we proceed to address the specific RQs, we present a model driven by the exhibited ERP components we found and their association with the cognitive processes. Our data suggest support for two processes: 1) *Awareness*, produced in early processing demonstrated by early components of N1, P2 and 2) *Memory*, pronounced in evoked component N400 and P6. This process can be characterised by early employment of subconscious processes transformed to latter conscious processes. Awareness is updated by output from the initialised memory checks. Here we talk about orchestrated activity supporting this adaptive behaviour, where the synchronicity between regions, whose resources support different Scenarios, are needed to inform other parts of the process in order to make a



(a) N1 (90-150 ms) over RF/RFC ROI

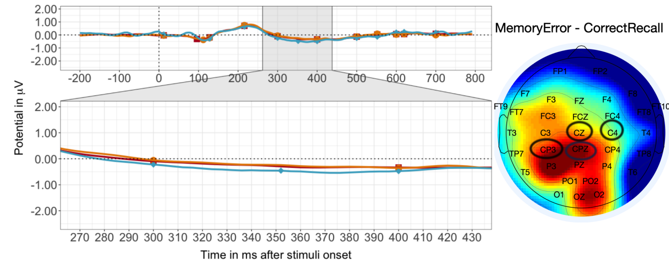


(b) P2 (150-270 ms) over LTP ROI

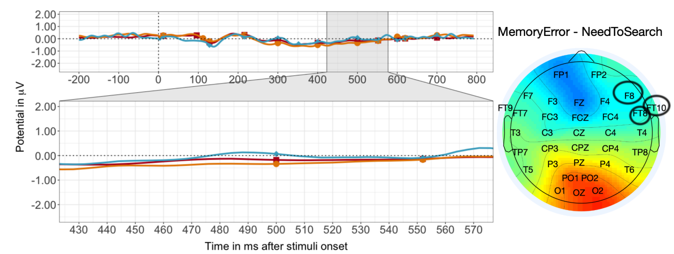


(c) N400 onset (270 - 430 ms) over RFT ROI

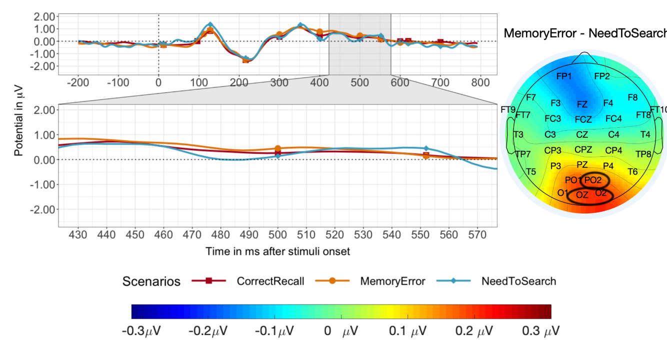
Figure 3: Grand ERP waveforms of three Scenarios with a zoom on time windows 90-150ms, 150-270ms and 270-430ms over significant ROIs (see Table 2) and the topological maps of the location of the corresponding ROI. Averaged over 14 participants.



(a) N400 onset (270-430 ms) over C/CP ROI



(b) N400 offset (430-570 ms) over RFT ROI



(c) N400 offset (430-570 ms) over PO/O ROI

Figure 4: Grand ERP waveforms of three Scenarios with a zoom on time windows 270-430ms and 430-570ms over significant ROIs (see Table 2) and the topological maps of the location of the corresponding ROI. Averaged over 14 participants.

decision in each level. Our results suggest a transition network between attention and awareness process as an act upon stimuli input and knowledge retrieval. This proposition is supported by different latencies of relevant components supporting the notion of overlapped cognitive activity. We use this model as a reference and to link the findings of separate RQs with the cognitive underpinnings of the Scenarios.

Temporal signatures of IN realisation. To address **RQ1** we aimed to bring evidence that IN phenomenon is happening prior to the user's explicit decision to search and can be detected in real-time based. Topologically, the earliest activity associated with NeedToSearch depicting the N1 component was concentrated in F/FC area with the right lateralisation. The evidence of RF/RFC ROI might be seen as an earliest sign associated with NeedToSearch level (i.e., IN). Next, the shift to left lateralisation of TP ROI was found to be a significant driver of P2 difference between NeedToSearch and MemoryError level, with the lowest averages for NeedToSearch. In summary, these findings of early rapid neural correlates suggest early engaged awareness processes responsible for realisation of IN. Following the findings of Paynter et al. [68], N1-P2 complex signalises that the early activity emerging before the realisation reaches user's consciousness. The second phase of the timeline complements the information process from the engagement of memory. Here, the processing differed at the offset of N400 component. The reduced amplitudes of NeedToSearch caused significant differences with the level MemoryError with the discriminative activity distributed in two ROI with anterior-posterior locality. In conclusion, it seems that the signature features of NeedToSearch are concentrated in the early phase of Awareness process which indicate that IN is indexed by very early ERP components (N1-P2 component). The result of reduced activity for NeedToSearch in anterior-posterior areas at the offset of N400 might suggest an overlapping activity between N400 and P6 which makes it difficult to separate the effect.

Variability in the cognitive processing. We found a consistency in the neural manifestations, namely N1-P2, N400 and P6, underlying each of the three Scenarios (**RQ2**). This finding supports the data-driven model presented at the beginning of this Section and describing the temporal dynamics of the cognitive operations. The modulations of ERP suggest that during information processing, a variety of cognitive processes are relied upon to different degrees (**RQ3**). The early distinction between NeedToSearch and MemoryError levels might be indexing process of awareness about one's own knowledge, demonstrated by the occurrence of N1 component. Moreover, the highest neural activity elicited for NeedToSearch level early in the time indicates not only early availability of knowledge cues but also, that knowledge cues might predicts the absence of knowledge (and potentially lead to early IN realisation). The lowered amplitude of N1 for MemoryError indicates that when people think they know the answer, less neural resources are recruited [57] than when they think they don't (NeedToSearch). The early emergence of N1 is believed to be triggered regardless of the task demands [69] and, as such, is used to measure early perceptual processing [70]. The modulations of N1 are also contributed to attention [69]. High early activity, demonstrated by P2 component with differences pronounced in TP region suggests further support for activated processes of attention awareness and deploying necessary resources to link attention and memory. P2 component was found to be associated with early low-level sensory processing, triggering early input processing, such as registration and input classification [71]. Significantly higher P2 amplitudes for MemoryError in contrast to NeedToSearch might suggest differences in the cognitive effort associated with memory recall [72, 73], and precisely higher memory strength [74]

The N1-P2 effect could represent some kind of early recognition of the stimulus with an increased attention [75] to an item which has a certain degree of familiarity, possibly employing the retrieval from contextual memory [76] and recall of past experiences [74], which might interpret the amplified amplitude for MemoryError level. Contextual (source) memory stores the background context of the person's past experiences as the core of episodic memory [77].

The immediate sequence of the offset of P2 and the onset of N400 seems to be an indication of a link between early and late processes [74, 78]. P2 emergence confirms the availability of memory and, thus, supports the decision by memory search and verification, which is amplified as N400. Modulation of N400 component is considered to correlate with familiarity, however, there is not a consensus what this component indexes. As Diana et al. [74] suggest, the later ERP effects might index co-occurring memory phenomena or an initiation of memory search and, thus, not necessarily index just one process. The emergence of N400 in connection to memory marks the attempts of deeper memory search [74].

The late activity depicting the offset of N400 with the significant anterior-posterior distribution, signalising memory-driven processes. A sustained activity of P6 might be triggered by final verification checks maintaining the previously triggered processes of awareness and memory search and the flow of information between significant parts of the brain resulting in an conscious response to stimuli [79].

MemoryError level and Proactive IR support. According to design of the study, the participants were asked to choose the correct answer to the question or to acknowledge the state of not knowing. What is the explanation for a choice of MemoryError, then? As our data showed, MemoryError was part of several pairwise contrasts. Specifically, the amplified frontal signal N400 for MemoryError level, significantly different in contrast to CorrectRecall, might signalise a potential issue with the controlled (later) processes of memory search. This could imply that the incorrect knowledge was already encoded in their memory, causing such difference to exhibit, and the phenomenons such as memory error [24] of false memory appeared [80].

Could IR intervene and rectify the awareness about this misleading and unknown-to-user knowledge anomaly? More research is need to answer this question, but the present indicates that the neural correlates are sensitive to this phenomenon. As we noted at the beginning, the anticipation of INs is one of the desired functionalities of proactive systems. The functionality extends over the potential support for MemoryError level in order to prevent users from failing and reduce the future INs linked with MemoryError. More is however needed to investigate to understand the associated user behaviours. At last, more insight might provide the investigation of the answer choices themselves acting as the alternatives of the user judgments. Specifically, what is the trade-off between the choices, e.g. MemoryError and NeedToSearch, and how it can impact the user's behaviours in situations which require to assess their knowledge.

7 CONCLUSIONS

Our study contributes with a novel interdisciplinary *NeuraSearch*-based outtake to evaluate the spectrum of states of knowledge. Our study meets the criteria for objective evaluation of EEG response based on data-driven analysis. We inferred the variability of brain activity based on the contrast of IN state with the two other (no-IN) scenarios. We formally assessed the overall pattern of electrical activity split into time windows to project a state of INs. Our findings differentiated between ERP components elicited for the present Scenarios and provided us with the evidence of orchestrated activity between the cognitive functions. Further insight can advance the interpretation of the realisation of IN in the brain, e.g. to directly decode IN from the brain data and validate the robustness of these patterns outcomes in a study with larger sample of participants and in different contextual IN scenarios. Furthermore, the question of how the present knowledge could be exploited in the search applications, is desirable. The study serves to account for a more accurate model of an a-priori state of the user determining INs. The discovered neural mechanisms are the basis for modelling of user interactions and behaviours and, in turn, open the discussion about the system (proactive) response.

REFERENCES

[1] W.S. Cooper. A definition of relevance for information retrieval. *Information Storage and Retrieval*, 7(1):19–37, 1971.

[2] Peter Ingwersen. *Information Retrieval Interaction.* Taylor Graham Publishing, GBR, 1992.

[3] Robert S. Taylor. Question-Negotiation and Information Seeking in Libraries. *College & Research Libraries*, 29(3):178–194, 1968.

[4] Nicholas Belkin, R.N. Oddy, and H.M. Brooks. ASK for information retrieval: Part I. Background and theory. *Journal of Documentation*, 38:61–71, 12 1982.

[5] T.D. Wilson. On user studies and information needs. *Journal of Documentation*, 37(1):3–15, 1981.

[6] Carol C. Kuhlthau. Inside the Search Process: Information Seeking from the User's Perspective. *Journal of the American Society for Information Science*, 42(5):361–371, 1991.

[7] Peter Ingwersen. Cognitive perspectives of information retrieval interaction: Elements of a cognitive ir theory. *Journal of Documentation*, 52:3–50, 01 1996.

[8] P. Ingwersen. Psychological aspects of information retrieval. *Social Science Information Studies*, 4(2):83–95, 1984. Special Issue Seminar on the Psychological Aspects of Information Searching.

[9] Charles Cole. *A theory of information need for information retrieval that connects information to knowledge.* Information Today Inc., 2012.

[10] Reijo Savolainen. Conceptualizing information need in context. Technical Report 4, 2012.

[11] Reijo Savolainen. Information need as trigger and driver of information seeking: a conceptual analysis. *Aslib Journal of Information Management*, 69(1):2–21, 2017.

[12] Richard L. Derr. A conceptual analysis of information need. *Information Processing & Management*, 19(5):273–278, 1983.

[13] Brenda Dervin and Michael Nilan. Information Needs and Uses. *Journal of the Medical Library Association*, 9, 11 2002.

[14] Charles Cole. Taylor's Q1 "Visceral" level of information need: What is it?". *Information Processing & Management*, 57(2), 2020.

[15] Waseem Afzal. A proposed methodology for the conceptualization, operationalization, and empirical validation of the concept of information need. *Information Research*, 22(3):1–16, September 2017. Includes bibliographical references.

[16] Karen Markey. Levels of question formulation in negotiation of information need during the online presearch interview: A proposed model. *Information Processing Management*, 17(5):215–225, 1981.

[17] Carol C. Kuhlthau. Developing a model of the library search process: Cognitive and affective aspects. *RQ*, 28(2):232–242, 1988.

[18] Michael Bendersky and W. Bruce Croft. Analysis of long queries in a large scale search log. In *Proceedings of the 2009 Workshop on Web Search Click Data*, WSCD '09, page 8–14, New York, NY, USA, 2009. Association for Computing Machinery.

[19] Robert S. Taylor. The process of asking questions. *American Documentation*, 13(4):391–396, 1962.

[20] Yashar Moshfeghi, Peter Triantafillou, and Frank E. Pollick. Understanding Information Need. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval - SIGIR '16*, 2016.

[21] Yashar Moshfeghi, Peter Triantafillou, and Frank Pollick. Towards predicting a realisation of an information need based on brain signals. In *The World Wide Web Conference*, WWW '19, page 1300–1309, New York, NY, USA, 2019. Association for Computing Machinery.

[22] Yashar Moshfeghi and Frank Pollick. Neuropsychological model of the realization of information need. *Journal of the Association for Information Science and Technology*, 05 2019.

[23] Yashar Moshfeghi and Frank E. Pollick. Search process as transitions between neural states. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, page 1683–1692, Republic and Canton of Geneva, CHE, 2018. International World Wide Web Conferences Steering Committee.

[24] Victor van der Veen, Nitish dutt Sharma, Lorenzo Cavallaro, and Herbert Bos. Memory errors: The past, the present, and the future. In Davide Balzarotti, Salvatore J. Stolfo, and Marco Cova, editors, *Research in Attacks, Intrusions, and Defenses*, pages 86–106, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[25] Sumit Bhatia, Debapriyo Majumdar, and Nitish Aggarwal. Proactive information retrieval: Anticipating users' information need. In Nicola Ferro, Fabio Crestani, Marie-Francine Moens, Josiane Mothe, Fabrizio Silvestri, Giorgio Maria Di Nunzio, Claudia Hauff, and Gianmaria Silvello, editors, *Advances in Information Retrieval*, pages 874–877, Cham, 2016. Springer International Publishing.

[26] Johanne R. Trippas. Spoken conversational search: Audio-only interactive information retrieval. *SIGIR Forum*, 53(2):106–107, mar 2021.

[27] Jianfeng Gao, Chenyan Xiong, and Paul Bennett. *Recent Advances in Conversational Information Retrieval*, page 2421–2424. Association for Computing Machinery, New York, NY, USA, 2020.

[28] Manoj Kumar Chinnakotla and Puneet Agrawal. Lessons from building a large-scale commercial ir-based chatbot for an emerging market. In *The 41st International ACM SIGIR Conference on Research amp; Development in Information Retrieval*, SIGIR '18, page 1361–1362, New York, NY, USA, 2018. Association for Computing Machinery.

[29] Hamid Keshavarz. Human information behaviour and design, development and evaluation of information retrieval systems. *Program: electronic library and information systems*, 42, 09 2008.

[30] N. Belkin, R. Oddy, and H. Brooks. ASK for Information Retrieval: Part II. Results of a Design Study. *The Journal of Documentation*, 38:145–164, 1982.

[31] Markus Koskela, Petri Luukkonen, Tuukka Ruotsalo, Mats SjÖberg, and Patrik Floréen. Proactive information retrieval by capturing search intent from primary task context. *ACM Trans. Interact. Intell. Syst.*, 8(3), jul 2018.

[32] Jan R. Benetka, Krisztian Balog, and Kjetil Nørvåg. Anticipating information needs based on check-in activity. In *Proceedings of the 10th ACM International Conference on Web Search and Data Mining*, WSDM '17, page 41–50, New York, NY, USA, 2017. Association for Computing Machinery.

[33] Jan Benetka, John Krumm, and Paul Bennett. Understanding context for tasks and activities. In *The Fourth ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR 2019)*. ACM, March 2019.

[34] Annika M. Hinze, Carole Chang, and David M. Nichols. Contextual queries express mobile information needs. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '10, page 327–336, New York, NY, USA, 2010. Association for Computing Machinery.

[35] Yashar Moshfeghi and Joemon M. Jose. An effective implicit relevance feedback technique using affective, physiological and behavioural features. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval - SIGIR '13*, 2013.

[36] Mateusz Dubiel, Martin Halvey, Leif Azzopardi, and Sylvain Daronnat. Interactive evaluation of conversational agents: Reflections on the impact of search task design. In *Proceedings of the 2020 ACM SIGIR on International Conference on Theory of Information Retrieval*, ICTIR '20, page 85–88, New York, NY, USA, 2020. Association for Computing Machinery.

[37] Mateusz Dubiel, Martin Halvey, Leif Azzopardi, and Sylvain Daronnat. Investigating how conversational search agents affect user's behaviour, performance and search experience. 07 2018.

[38] Salvatore Andolina, Valeria Orso, Hendrik Schneider, Khalil Klouche, Tuukka Ruotsalo, Luciano Gamberini, and Giulio Jacucci. Investigating proactive search support in conversations. In *Proceedings of the 2018 Designing Interactive Systems Conference*, DIS '18, page 1295–1307, New York, NY, USA, 2018. Association for Computing Machinery.

[39] Ian Ruthven. The language of information need: Differentiating conscious and formalized information needs. *Information Processing and Management*, 56(1):77–90, 2019.

[40] Jaime Arguello, Adam Ferguson, Emery Fine, Bhaskar Mitra, Hamed Zamani, and Fernando Diaz. Tip of the tongue known-item retrieval: A case study in movie identification. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*, CHIIR '21, page 5–14, New York, NY, USA, 2021. Association for Computing Machinery.

[41] Sosuke Shiga, Hideo Joho, Roi Blanco, Johanne R. Trippas, and Mark Sanderson. Modelling Information Needs in Collaborative Search Conversations. pages 715–724, 2017.

[42] Yashar Moshfeghi. Neurasearch: Neuroscience and information retrieval. In Omar Alonso, Stefano Marchesin, Marc Najork, and Gianmaria Silvello, editors, *Proceedings of the Second International Conference on Design of Experimental Search & Information REtrieval Systems, Padova, Italy, September 15-18, 2021*, volume 2950 of *CEUR Workshop Proceedings*, pages 193–194. CEUR-WS.org, 2021.

[43] Gernot Müller-Putz, René Riedl, and Selina Wriessnegger. Electroencephalography (eeg) as a research tool in the information systems discipline: Foundations, measurement, and applications. *Communications of the Association for Information Systems*, 37:911–948, 01 2015.

[44] Jacek Gwizdka, Yashar Moshfeghi, and Max L. Wilson. Introduction to the special issue on neuro-information science. *Journal of the Association for Information Science and Technology*, 70(9):911–916, 2019.

[45] Tefko Saracevic. The stratified model of information retrieval interaction: Extension and application. *Proceedings of the ASIS Annual Meeting*, 34:313–327, 01 1997.

[46] Marco Allegretti, Yashar Moshfeghi, Maria Hadjigeorgieva, Frank E. Pollick, Joemon M. Jose, and Gabriella Pasi. When relevance judgement is happening? an eeg-based study. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '15, page 719–722, New York, NY, USA, 2015. Association for Computing Machinery.

[47] Zuzana Pinkosova, William J. McGeown, and Yashar Moshfeghi. The cortical activity of graded relevance. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '20, page 299–308, New York, NY, USA, 2020. Association for Computing Machinery.

[48] Lauri Kangassalo, Michiel Spapé, Giulio Jacucci, and Tuukka Ruotsalo. Why do users issue good queries?: Neural correlates of term specificity. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR'19, pages 375–384, United States, 2019. ACM, Association for Computing Machinery.

[49] Yashar Moshfeghi, Luisa R. Pinto, Frank E. Pollick, and Joemon M. Jose. Understanding relevance: An fMRI study. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2013.
[50] Giulio Jacucci, Oswald Barral, Pedram Daee, Markus Wenzel, Baris Serim, Tuukka Ruotsalo, Patrik Pluchino, Jonathan Freeman, Luciano Gamberini, Samuel Kaski, and Benjamin Blankertz. Integrating neurophysiologic relevance feedback in intent modeling for information retrieval. *Journal of the Association for Information Science and Technology*, 70(9):917–930, 2019.
[51] Angel Jimenez-Molina, Cristian Retamal, and Hernan Lira. Using psychophysiological sensors to assess mental workload in web browsing. *Sensors*, 18, 01 2018.
[52] Jacek Gwizdka, Rahilsadat Hosseini, Michael Cole, and Shouyi Wang. Temporal dynamics of eye-tracking and eeg during reading and relevance decisions. *Journal of the Association for Information Science and Technology*, 68, 08 2017.
[53] Jukka-Pekka Kauppi, Melih Kandemir, Veli-Matti Saarinen, Lotta Hirvenkari, Lauri Parkkonen, Arto Klami, Riitta Hari, and Samuel Kaski. Towards brain-activity-controlled information retrieval: Decoding image relevance from meg signals. *NeuroImage*, 112:288 – 298, 2015.
[54] Serena Midha, Horia A. Maior, Max L. Wilson, and Sarah Sharples. Measuring mental workload variations in office work tasks using fnirs. *International Journal of Human-Computer Studies*, 147:102580, 2021.
[55] Horia A. Maior, Richard Ramchurn, Sarah Martindale, Ming Cai, Max L. Wilson, and Steve Benford. Fnirs and neurocinematics. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, page 1–6, New York, NY, USA, 2019. Association for Computing Machinery.
[56] Carlos de la Torre-Ortiz, Michiel M. Spapé, Lauri Kangassalo, and Tuukka Ruotsalo. Brain relevance feedback for interactive image generation. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, UIST '20, page 1060–1070, New York, NY, USA, 2020. Association for Computing Machinery.
[57] Steven Luck. *An Introduction to The Event-Related Potential Technique*. 01 2005.
[58] The ten twenty electrode system: International federation of societies for electroencephalography and clinical neurophysiology. *American Journal of EEG Technology*, 1(1):13–19, 1961.
[59] Emily Kappenman and Steven Luck. Best practices for event-related potential research in clinical populations. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 1, 11 2015.
[60] Elizabeth F. Chua, Rifat Ahmed, and Sandry M. Garcia. Effects of HD-tDCS on memory and metamemory for general knowledge questions that vary by difficulty. *Brain Stimulation*, 10(2):231–241, 2017.
[61] Olaf Dimigen, Werner Sommer, Annette Hohlfeld, Arthur Jacobs, and Reinhold Kliegl. Coregistration of eye movements and eeg in natural reading: Analyses and review. *Journal of experimental psychology. General*, 140:552–72, 07 2011.
[62] Peter Hagoort. Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, 15(6):883–899, 2003.
[63] Tali Ditman, Phillip J. Holcomb, and Gina R. Kuperberg. An investigation of concurrent ERP and self-paced reading methodologies. *Psychophysiology*, 44(6):927–935, 2007.
[64] Manuel J.A. Eugster, Tuukka Ruotsalo, Michiel M. Spapé, Oswald Barral, Niklas Ravaja, Giulio Jacucci, and Samuel Kaski. Natural brain-information interfaces: Recommending information by relevance inferred from human brain signals. *Scientific Reports*, 2016.
[65] A. Delorme and Scott Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134:9–21, 2004.
[66] Nima Bigdely-Shamlo, Tim Mullen, Christian Kothe, Kyung-Min Su, and Kay A. Robbins. The prep pipeline: standardized preprocessing for large-scale eeg analysis. *Frontiers in Neuroinformatics*, 9:16, 2015.
[67] Phillip Alday. How much baseline correction do we need in erp research? extended glm model can replace baseline correction while lifting its limits. *Psychophysiology*, 56, 07 2017.
[68] Christopher A. Paynter, Lynne M. Reder, and Paul D. Kieffaber. Knowing we know before we know: Erp correlates of initial feeling-of-knowing. *Neuropsychologia*, 47(3):796 – 803, 2009.
[69] Steven Hillyard, Robert Hink, Vincent Schwent, and Terence Picton. Electrical signs of selective attention in the human brain. *Science (New York, N.Y.)*, 182:177–80, 11 1973.
[70] Elizabeth Bauer, Kayla Wilson, and Annmarie MacNamara. *Cognitive and Affective Psychophysiology*. 01 2020.
[71] Joel L. Voss and Ken A. Paller. pages 81–98. Number September 2016. Elsevier, 3rd edition, 2016.
[72] Nicole Franke, Ralph Radach, Arthur Jacobs, and Markus Hofmann. No one way ticket from orthography to semantics in recognition memory: N400 and p200 effects of associations. *Brain Research*, 1639, 02 2016.
[73] Patricia Bauer and Felicia Jackson. Semantic elaboration: Erps reveal rapid transition from novel to known. *Journal of experimental psychology. Learning, memory, and cognition*, 41, 08 2014.
[74] Rachel A. Diana, Kaia L. Vilberg, and Lynne M. Reder. Identifying the erp correlate of a recognition memory search attempt. *Cognitive Brain Research*, 24(3):674 – 684, 2005.
[75] Viktor Mueller, Yvonne Brehmer, Timo von Oertzen, Shu-Chen Li, and Ulman Lindenberger. Electrophysiological correlates of selective attention: A lifespan comparison. *BMC neuroscience*, 9:18, 02 2008.
[76] Chun yan Guo, Li Duan, Wen Li, and Ken A. Paller. Distinguishing source memory and item memory: Brain potentials at encoding and retrieval. *Brain Research*, 1118:142–154, 2006.
[77] Larry R. Squire and Stuart M. Zola. Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, 93(24):13515–13522, 1996.
[78] Joseph Dien, Charles Michelson, and Michael Franklin. Separating the visual sentence n400 effect from the p400 sequential expectancy effect: Cognitive and neuroanatomical implications. *Brain research*, 1355:126–40, 10 2010.
[79] Haopei Yang, Geoffrey Laforge, Bobby (Boge) Stojanoski, Emily Nichols, Ken McRae, and Stefan Köhler. Late positive complex in event-related potentials tracks memory signals when they are decision relevant. *Scientific Reports*, 9, 12 2019.
[80] C. Brainerd and V. Reyna. *The Science of False Memory*. Oxford University Press, 2005.