Towards Detecting Tonic Information Processing Activities with Physiological Data

Kaixin Ji
RMIT University
Melbourne, Australia
kaixin.ji@student.rmit.edu.au

Damiano Spina
RMIT University
Melbourne, Australia
damiano.spina@rmit.edu.au

Danula Hettiachchi
RMIT University
Melbourne, Australia
danula.hettiachchi@rmit.edu.au

Flora D. Salim
The University of New South Wales
Sydney, Australia
flora.salim@unsw.edu.au

Falk Scholer
RMIT University
Melbourne, Australia
falk.scholer@rmit.edu.au

ABSTRACT

Characterizing Information Processing Activities (IPAs) such as reading, listening, speaking, and writing, with physiological signals captured by wearable sensors can broaden the understanding of how people produce and consume information. However, sensors are highly sensitive to external conditions that are not trivial to control – not even in lab user studies. We conducted a pilot study ($N = 7$) to assess the robustness and sensitivity of physiological signals across four IPAs (READ, LISTEN, SPEAK, and WRITE) using multiple sensors. The collected signals include Electrodermal Activities, Blood Volume Pulse, gaze, and head motion. We observed consistent trends across participants, and ten features with statistically significant differences across the four IPAs. Our results provide preliminary quantitative evidence of differences in physiological responses when users encounter IPAs, revealing the necessity to inspect the signals separately according to the IPAs. The next step of this study moves into a specific context, information retrieval, and the IPAs are considered as the interaction modalities with the search system, for instance, submitting the search query by speaking or typing.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in ubiquitous and mobile computing**; • **Information systems → Users and interactive retrieval**.

KEYWORDS

information processing activities, physiological signals, information seeking

ACM Reference Format:
Kaixin Ji, Damiano Spina, Danula Hettiachchi, Flora D. Salim, and Falk Scholer. 2023. Towards Detecting Tonic Information Processing Activities with Physiological Data. In *Adjunct Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing & the 2023 ACM International Symposium on Wearable Computing (UbiComp/ISWC '23 Adjunct), October 08–12, 2023, Cancun, Quintana Roo, Mexico.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3594739.3610679

UbiComp/ISWC '23 Adjunct, October 08–12, 2023, Cancun, Quintana Roo, Mexico

ACM ISBN 979-8-4007-0200-6/23/10.
https://doi.org/10.1145/3594739.3610679

1 INTRODUCTION

To better understand user experience, researchers use physiological data (especially collected from wearable sensors) in more challenging contexts, such as information-seeking and retrieval tasks. In these tasks, the stimuli are usually texts, which have less effect than videos [1]; consequently, fewer physiological responses can be captured.

Information processing tasks comprise two fundamental activities, information input and output. Both activities have two forms, i.e., input via visuals (reading) or audio (listening), and output via motion (writing or typing) or speech (speaking). Critchley [7] noted cognitive, affective, and motor activities involve different brain regions and impact the Electrodermal Activities (EDA, i.e., skin conductance) and cardiovascular responses. The cognitive states and motor activities vary in four tonic (i.e., essential) information processing activities (IPAs) – READ, LISTEN, SPEAK, and WRITE. For example, all activities involve various cognitive states; SPEAK and WRITE also involve different motor activities. Besides, the EDA responses are different for visual stimuli (READ) and auditory stimuli (LISTEN) [28]. Hence, we anticipate they produce distinct physiological signals. To date, physiological data have been applied to detect cognitive loads to infer stress or interruptibility [11, 12, 32], or engagement [8, 10, 26]. However, these studies were built upon inclusive scenarios, i.e., including two or more tonic activities. The physiological responses varied across scenarios. Without a first look at the responses separately, there is a risk that the signals captured other variances instead of the desired responses.

What still remains unclear is the baseline of physiological measurements for all tonic activities when multiple wearables are used simultaneously (e.g., earphones, wristbands, and eye-trackers). In other words, how the physiological responses generally describe the IPAs remains unknown. In this work, we aim to discriminate the four tonic IPAs – READ, LISTEN, SPEAK, and WRITE– using the physiological signals in a rigorous-controlled user study. We use self-reported engagement as a measurement proxy and multiple human-generated signals. In considering mobility and scalability, we use the signals that can be captured by commercial wearable sensors, EDA, Blood Volume Pulse (BVP), gaze, and head motion.

Table 1: Extracted Features. Skin Conductance Level (SCL), Skin Conductance Response (SCR), Standard deviation (std), Inter-Beat Interval (IBI) measured in milliseconds, Pupil Diameter (PD). *The baseline pupil diameter is taken as the median from the RELAX task of each section. Features with statistically significant differences are in boldface.

Signals	Category	Features
EDA	SCL	mean, max, min, std & mean amplitude
	SCR	max, min, range & mean magnitude, mean amplitude,**peak count**; mean rise time, summation of rise time, mean recovery time
BVP	-	mean, max, **min**, std & mean amplitude
	IBI	count, **mean**, std; **max time**, ratio of max time to total time
Head Motion	gyroscope accelerometer	mean, std of magnitudes of *x, y, z* coordinates
Gaze	PD	**mean**, max, min, std of PD; **mean**, max, **min**, std of the difference to the baseline*
	movement	**count**, **total duration**, **mean duration of fixation** and saccade

Trend analysis and statistical tests indicated that the participants' physiological signals in IPAs differed from those with no IPAs. Although there are individual differences, some consistent trends were observed. In particular, a downtrend in EDA and an uptrend in pupil changes for relaxing activity (RELAX), a steady trend in EDA for READ, and a rapid EDA uptrend for SPEAK. The statistical analysis further indicated that 10 out of 42 features have significant differences across all activities. The post-hoc statistical test revealed significant effects of BVP features in SPEAK and significant effects of gaze features in RELAX.

2 METHODOLOGY

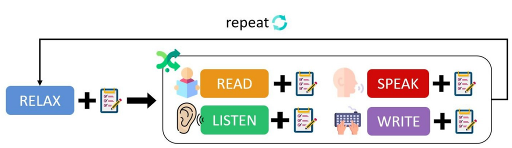


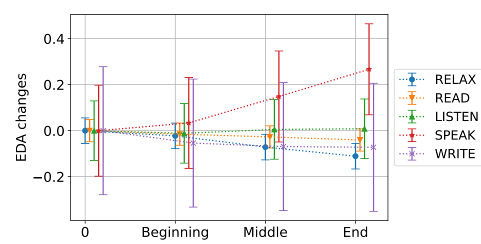
Figure 1: Experiment Procedure for Pilot Study.

As presented in Figure 1, the study consists of two sections, each corresponding to two pre-defined activity complexity levels (low and high). A 5-minute break is provided after one section. For READ and LISTEN IPAs, complexity is defined using low and high readability scores [30]. The materials are new articles with around 500 words, some synthesized into speech for LISTEN. For SPEAK and WRITE, complexity is estimated by the type of questions and the length of expected answers. For example, an easy question is, 'what was your routine this morning? (100 words minimum)' and a hard question is 'does social media make you in general happier or sadder? why? (300 words minimum)'. Each section starts with watching a relaxing video (RELAX), where minimal cognitive efforts are involved. Then, the participant completes four randomized IPAs: READ, LISTEN, SPEAK, and WRITE. Specifically, the participant needs to read one article, listen to one article, answer a question by speaking, and answer a question by typing. After each task (including RELAX), the participant completes an engagement scale [25].

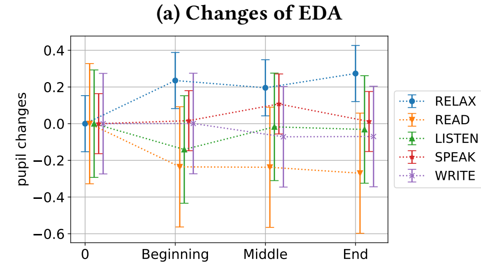
The participant is invited to the experiment room where has a desktop PC mounted with a Tobii eye-tracker and a webcam. The participant sits in front of the computer and wears a pair of Nokia eSense earphones and the E4 wristband on the non-dominant hand. All participants used the computer mouse with their right hand. The consent form is signed before the experiment starts. After setting up the sensors, the participant is instructed to complete calibration, followed by the debrief page and a pre-task survey. The survey asks for sleeping hours prior and caffeine intake on the experiment day, which might affect the cognition status [3]. The participant is informed of the right to pause or terminate the experiment whenever they feel uncomfortable. To avoid interruption, the instructor leaves the room after the calibration and instruction.

3 DATA AND RESULTS

The signals are segmented according to the timestamps recorded during the experiment. As listed in Table 1, overall, we extract 42 features from four types of signals, 14 features from EDA, 10 features from BVP, 4 features from head motion, and 14 features from the gaze.



(a) Changes of EDA



(b) Changes of Pupil Diameter

Figure 2: Aggregation of EDA/Pupil Diameter changes at three segments (Beginning, Middle, End) in each IPA for all participants. The dot points present the overall mean, and the error bars present the standard deviation.

Trend Analysis. To handle the issue of time spent variations, we use the same strategy provided by Wise et al. [31]. We divided the signals into three segments with equal time lengths – the beginning, middle, and end of the task – computed the changes by taking the arithmetic mean for each segment and subtracting from the second when starting the task (the second is not included in the segment).

As presented in Figures 1B and 1C, the EDA overall declines while the pupil overall dilates in RELAX. This indicates the participants were relaxing and engaging in the video. The EDA slightly decreases with the pupil constricts in READ throughout. The result aligns with prior work [6] that reports EDA decreases over time when reading. For LISTEN, the EDA remains unchanged, with only minor increases observed, while the pupil falls initially and rises in the middle. The EDA notably increases in SPEAK, and the pupil dilates more, while both EDA and pupil drop in WRITE. The following statistical analysis found more intensive BVP responses in SPEAK than WRITE. These responses also cause a rising in EDA. One possible explanation is speaking requires more physical effort than typing [9].

ANOVA (repeated-measures) Analysis. As presented in Table 1, we have extracted a total of 42 features. Our results indicate that ten features have significant differences (Table 2), 1 EDA, 3 BVP, and 6 gaze features. The head motion features did not have any significant effect. The post-hoc tests revealed fewer group pairs with significant differences (Table 3), possibly due to the limited number of participants. Despite this, RELAX is overall different from others in gaze-related features. RELAX has significant higher *min_pupil_change*, longer *total_fixation_duration* and *mean_fixation_duration* than READ. RELAX also has significant more *num_fixation* and longer *total_fixation_duration* than SPEAK. The BVP features, SPEAK has significant higher *BVP_min*, shorter *IBI_mean* and *time_max_BVP* than WRITE. *BVP_min* is also significantly larger in READ than in WRITE. Surprisingly, while the EDA signals revealed different trends among IPAs, none of the EDA features exhibited statistically significant differences. Interestingly, LISTEN has more fixations than READ and SPEAK, although the statistical testing does not reveal significant differences.

Table 2: One-way Repeated ANOVA (F-statistic) for the features that show significant differences, with generalized Eta squared effect size (η^2_G) and the Greenhouse-Geisser corrected $p^\wedge$ (for the violation of sphericity). Degree of Freedom is 24. η^2_G >= .02 means small effect size, >= .13 means medium effect size, and >= .26 means large effect size [4].

Feature	F	$p^\wedge$	η^2_G
SCR_peak_count	6.187	.013	.436
BVP_min	4.565	.036	.379
IBI_mean	4.962	.005	.336
time_max_BVP	6.618	.001	.394
min_pupil_diameter	10.618	.000	.294
mean_pupil_change	11.073	.000	.536
min_pupil_change	3.124	.033	.245
num_fixation	9.595	.011	.556
total_fixation_duration	7.927	.022	.516
mean_fixation_duration	10.314	.000	.564

Table 3: Post-hoc t-test (t-statistic) for the features and the pairs that show significant differences (p^a < .05, adjusted using Bonferroni correction). Degree of Freedom is 6.

Feature	t	p^a	η^2	A	B
BVP_min	4.412	**.045**	.549	READ	WRITE
BVP_min	5.417	**.016**	.645	SPEAK	WRITE
IBI_mean	−4.380	**.047**	.463	SPEAK	WRITE
time_max_BVP	−6.457	**.007**	.536	SPEAK	WRITE
mean_pupil_diameter	9.209	**.001**	.407	RELAX	READ
mean_pupil_diameter	4.372	**.047**	.291	RELAX	LISTEN
mean_pupil_diameter	5.176	**.021**	.326	RELAX	WRITE
mean_pupil_change	9.293	**.001**	.843	RELAX	READ
mean_pupil_change	5.007	**.024**	.639	RELAX	LISTEN
mean_pupil_change	6.217	**.008**	.737	RELAX	WRITE
min_pupil_change	6.614	**.006**	.423	RELAX	READ
total_fixation_duration	5.240	**.019**	.708	RELAX	READ
total_fixation_duration	4.804	**.030**	.688	RELAX	SPEAK
mean_fixation_duration	4.868	**.028**	.631	RELAX	READ
number_fixation	4.788	**.030**	.635	RELAX	SPEAK

4 DISCUSSION & LESSONS LEARNED

In this paper, we compare the differences in physiological responses between four IPAs, READ, LISTEN, SPEAK, WRITE, and RELAX. Based on the data collected from 7 participants, our trend analysis highlights differences among activities. A statistical test further confirms these findings, showing ten features have a significant effect. We acknowledge the limitation of our study and anticipate that the results may change with a larger sample size. However, we note that the current results are sufficient to demonstrate a difference among the activities. We suggest researchers consider different responses evoked by the different interaction forms when using physiological signals related to information processing activities. In addition, our analysis suggests the need to use multi-modal data to analyze these activities. Our results above have revealed differences across IPAs, so looking at each activity separately is necessary.

We have also identified problems in the current experiment setup using the methodology described in Ji et al. [13]. The first regards the materials provided for the reading and listening tasks. Although we controlled the readability levels and word counts, other variables, e.g., users' cognitive bias [5, 19, 21], or misinformation [16], can exist and possibly influence the physiological responses. The second regards a large variation in the task duration across participants which also influences the physiological responses.

5 NEXT STEPS

Based on the results and the limitations, we have moved toward a specific scenario, Information Retrieval (IR). Due to the diversity of search systems, the 4 IPAs are considered as different interaction modalities in the search process [14, 15, 29]. Specifically, enter the search query by speaking or typing, or receive the search result by reading or listening.

The interplay among affect, cognition, and physical behaviors in the search process is not new in IR. The Information Search Process model [17] and the Social-Biological Information Technology model [24] interpret their transition through the search process. So far, few studies investigated these internal activities (affect and cognition) with brain-generated signals, and found the differences between the search stages [22, 23, 27]. However, the other physiological signals, such as EDA and BVP, have not yet been applied in this context. Therefore, our next step is to investigate the potential of physiological signals used in IR. In addition to the sensors used in the previous study, we also add the EEG sensor. EEG and pupil data can provide a robust metric for measuring cognitive loads. With

EEG and pupil data as a reference to the internal activities during each stage, we aim to investigate whether the peripheral data, e.g., EDA and BVP, can achieve a similar result.

5.1 User Study

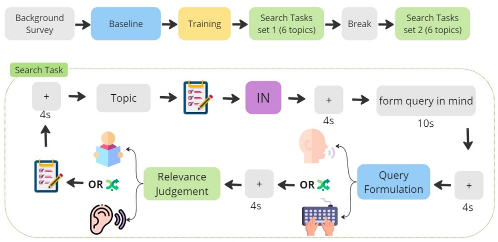


Figure 3: Experiment Procedure for Study 2.

A general search iteration includes a realization of information need, query formulation and execution, search result list viewing, relevance and satisfaction judgment. Depending on the result of the judgment, a search session might involve many search iterations transferring by reformulating the query. Focusing on one search iteration, an fMRI study by Moshfeghi and Pollick [22] investigated the neuro- and cognitive activities in the transition of search stages. To compare with their results, we design a similar setting with the same search stages as theirs: Information Need (IN), Query Formulation (QF), Query Submission (QS), and Relevance Judgment (RJ). Note that as we are interested in the different responses among search stages rather than particular search evaluation criteria, only highly-relevant search results are included for RJ.

There are 12 topics and corresponding backstories selected from the *TREC2002-InformationNeed* dataset [20]. To avoid risks of triggering irrelevant emotions or cognitive bias [2], the selected topics are from the Understanding category as it only requires the participants to find information and gain some understanding. Besides, the topics related to crises, wars, conspiracy, or any politically sensitive topics are removed. Then for each topic, we manually select a few relevant articles from the TREC English document collections and use ChatGPT to write a 150-word summary based on these articles and a binary factual judgment question.

The procedure is presented in Figure 3. After calibration and background survey, the participants are instructed to have an eyes-open section that collects the baseline data, followed by a training section. Each participant then needs to complete the search tasks for all 12 topics. Each search task is expected to take 2:30 minutes to complete. A break time is provided in the middle. For the search task, a topic is shown after a 4-second blank, then the participants need to rate their interests, familiarity and perceived difficulty about the topic. After that, a backstory to introduce the Information Need is presented. Next, the participants are asked to form a search query in mind, followed by an instruction about submitting the query by speaking (QF-S) or typing (QF-T). Then the participants either read (RJ-R) or listen (RJ-L) to the search result. A 4-second gap is provided after each search stage (i.e., IN, QF, QS, RJ) to clean the data. In the end, the participants need to answer a judgment question (to ensure they are focused) and rate perceived relevance and difficulty in understanding the search result. The sequences of topics and the combination of interaction modalities are randomized.

5.2 Hypotheses

Based on the findings we have so far and the literature, we have built the following hypotheses. H1. There are more sub-processes in Query Formulation (QF) than in the other search stages; thus the data can be an indicator of cognitive activities are effective in detecting QF. H2. There are more emotions be evoked in Relevance Judgment (RJ, high relevance) and Information Need (IN); thus the data can be an indicator of emotions are effective in detecting RJ and IN. H2a. The negative emotions are mostly evoked in IN, due to the feeling of uncertainty; thus the data has a decreasing trend for IN. H2b. If the search result is highly relevant, the positive emotions are mostly evoked in RJ, due to the feeling of uncertainty relieved; thus the data has an increasing trend for RJ. H3. Using features indicating motor activities are effective in detecting interaction modalities. For example, the BVP features are more differentiable between typing and speaking (higher in Speaking). Pupil dilates more in listening and EDA increases in reading.

5.3 Open Challenges

The study setting presents several open challenges that need to be addressed. The first challenge concerns how this study can receive accurate results and be representative of information processing activities. Secondly, this study tries to split the search process into multiple sub-processes; thus, the sub-processes have a short duration. For example, entering search queries only takes a few seconds. A challenge is ensuring follow the steps in short-duration tasks and do not combine them once they learn the procedure. Meanwhile, because the physiological responses sometimes delay or last longer, minimizing the impact caused by the sequence of the tasks is also a challenge. Furthermore, each physiological signals have different response and delay times. The last challenge concerns integrating multi-modal data and getting accurate information from them.

6 CONCLUSION

In conclusion, this work aims to detect internal activities (e.g., cognition and affect) during information processing. Concerning the high sensitivity of physiological signals, this work carefully looks at the signals respectively by different interaction modalities, i.e., the four tonic information-processing activities (IPAs) – READ, LISTEN, SPEAK, WRITE. It first conducts a preliminary study to investigate whether the signals are impacted by the IPAs. The results have revealed differences among IPAs. Although only the data from 7 participants are used for the analysis, it is sufficient to indicate that it is necessary to look at the signals separately. With the results and lessons learned from the preliminary study, the next step is moving forward to a specific scenario where the internal activity is more transitional within the process, which is Information Search [17, 18]. The results of this work will contribute to detecting any complex factors, e.g., cognitive bias [2, 15], in further experiments.

ACKNOWLEDGMENTS

This research is partially supported by the Australian Research Council (CE200100005, DE200100064).

REFERENCES

[1] Ioannis Arapakis, Ioannis Konstas, and Joemon M. Jose. 2009. Using Facial Expressions and Peripheral Physiological Signals as Implicit Indicators of Topical Relevance. In *Proceedings of the 17th ACM international conference on Multimedia (MM'09)* (Beijing, China) *(MM '09)*. Association for Computing Machinery, New York, NY, USA, 461–470. https://doi.org/10.1145/1631272.1631336

[2] Leif Azzopardi. 2021. Cognitive Biases in Search: A Review and Reflection of Cognitive Biases in Information Retrieval. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval* (Canberra ACT, Australia) *(CHIIR '21)*. Association for Computing Machinery, New York, NY, USA, 27–37. https://doi.org/10.1145/3406522.3446023

[3] Ebrahim Babaei, Benjamin Tag, Tilman Dingler, and Eduardo Velloso. 2021. A Critique of Electrodermal Activity Practices at CHI. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 177, 14 pages. https://doi.org/10.1145/3411764.3445370

[4] Roger Bakeman. 2005. Recommended effect size statistics for repeated measures designs. *Behavior research methods* 37 (2005), 379–384.

[5] Nattapat Boonprakong, Xiuge Chen, Catherine Davey, Benjamin Tag, and Tilman Dingler. 2023. Bias-Aware Systems: Exploring Indicators for the Occurrences of Cognitive Biases When Facing Different Opinions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 27, 19 pages. https://doi.org/10.1145/3544548.3580917

[6] Anne-Marie Brouwer, Maarten Hogervorst, Boris Reuderink, Ysbrand van der Werf, and Jan van Erp. 2015. Physiological signals distinguish between reading emotional and non-emotional sections in a novel. *Brain-Computer Interfaces* 2 (2015), 76–89. Issue 2-3. https://doi.org/10.1080/2326263x.2015.1100037

[7] Hugo D Critchley. 2002. Electrodermal responses: what happens in the brain. *The Neuroscientist* 8, 2 (2002), 132–142.

[8] Yi Ding, Yaqin Cao, Vincent G Duffy, Yi Wang, and Xuefeng Zhang. 2020. Measurement and Identification of Mental Workload During Simulated Computer Tasks with Multimodal Methods and Machine Learning. *Ergonomics* 63, 7 (2020), 896–908.

[9] Margaret Foley, Géry Casiez, and Daniel Vogel. 2020. Comparing Smartphone Speech Recognition and Touchscreen Typing for Composition and Transcription. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3313831.3376861

[10] Nan Gao, Wei Shao, Mohammad Saiedur Rahaman, and Flora D. Salim. 2020. N-Gage: Predicting in-Class Emotional, Behavioural and Cognitive Engagement in the Wild. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 3, Article 79 (Sept. 2020), 26 pages. https://doi.org/10.1145/3411813

[11] Martin Gjoreski, Mitja Luštrek, and Veljko Pejović. 2018. My Watch Says I'm Busy: Inferring Cognitive Load with Low-Cost Wearables. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (Singapore, Singapore) *(UbiComp '18)*. Association for Computing Machinery, New York, NY, USA, 1234–1240. https://doi.org/10.1145/3267305.3274113

[12] Eija Haapalainen, SeungJun Kim, Jodi F. Forlizzi, and Anind K. Dey. 2010. Psycho-Physiological Measures for Assessing Cognitive Load. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing* (Copenhagen, Denmark) *(UbiComp '10)*. Association for Computing Machinery, New York, NY, USA, 301–310. https://doi.org/10.1145/1864349.1864395

[13] Kaixin Ji, Damiano Spina, Danula Hettiachchi, Flora Dilys Salim, and Falk Scholer. 2023. Examining the Impact of Uncontrolled Variables on Physiological Signals in User Studies for Information Processing Activities. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Taipei, Taiwan) *(SIGIR '23)*. Association for Computing Machinery, New York, NY, USA, 1971–1975. https://doi.org/10.1145/3539618.3591981

[14] Diane Kelly. 2009. Methods for Evaluating Interactive Information Retrieval Systems with Users. *Found. Trends Inf. Retr.* 3, 1—2 (jan 2009), 1–224. https://doi.org/10.1561/1500000012

[15] Johannes Kiesel, Damiano Spina, Henning Wachsmuth, and Benno Stein. 2021. The Meant, the Said, and the Understood: Conversational Argument Search and Cognitive Biases. In *Proceedings of the 3rd Conference on Conversational User Interfaces* (Bilbao (online), Spain) *(CUI '21)*. Association for Computing Machinery, New York, NY, USA, Article 20, 5 pages. https://doi.org/10.1145/3469595.3469615

[16] Lauren Kirkwood and Randall Minas. 2020. Approaching Fake News at the Expense of Truth: A Psychophysiological Study of News on Social Media. In *Proceedings of the Annual Hawaii International Conference on System Sciences*. 10 pages. http://hdl.handle.net/10125/64486

[17] Carol Collier Kuhlthau. 2005. Information Search Process. *Hong Kong, China* 7 (2005), 226.

[18] Gary Marchionini. 1995. *Information Seeking in Electronic Environments*. Number 9. Cambridge University Press.

[19] Randall K Minas, Robert F Potter, Alan R Dennis, Valerie Bartelt, and Soyoung Bae. 2014. Putting on the Thinking Cap: Using NeuroIS to Understand Information Processing Biases in Virtual Teams. *Journal of Management Information Systems* 30, 4 (2014), 49–82.

[20] Alistair Moffat, Peter Bailey, Falk Scholer, and Paul Thomas. 2014. Assessing the Cognitive Complexity of Information Needs. In *Proceedings of the 2014 Australasian Document Computing Symposium* (Melbourne, VIC, Australia) *(ADCS '14)*. ACM, New York, NY, USA, Article 97, 4 pages. https://doi.org/10.1145/2682862.2682874

[21] Patricia Moravec, Randall Minas, and Alan R Dennis. 2018. Fake News on Social Media: People Believe What They Want to Believe When it Makes No Sense at All. *Kelley School of Business research paper* (2018), 36 pages. https://doi.org/10.2139/ssrn.3269541

[22] Yashar Moshfeghi and Frank E. Pollick. 2018. Search Process as Transitions Between Neural States. In *Proceedings of the 2018 World Wide Web Conference* (Lyon, France) *(WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1683–1692. https://doi.org/10.1145/3178876.3186080

[23] Yashar Moshfeghi, Peter Triantafillou, and Frank E. Pollick. 2016. Understanding Information Need: An FMRI Study. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Pisa, Italy) *(SIGIR '16)*. Association for Computing Machinery, New York, NY, USA, 335–344. https://doi.org/10.1145/2911451.2911534

[24] Diane Nahl. 2007. Social–biological information technology: An integrated conceptual framework. *Journal of the American Society for Information Science and Technology* 58, 13 (2007), 2021–2046.

[25] Heather L. O'Brien, Paul Cairns, and Mark Hall. 2018. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies* 112 (2018), 28–39. https://doi.org/10.1016/j.ijhcs.2018.01.004

[26] Akshay Palimar Pai, Jayasankar Santhosh, and Shoya Ishimaru. 2023. Real-Time Feedback on Reader's Engagement and Emotion Estimated by Eye-Tracking and Physiological Sensing. In *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers* (Cambridge, United Kingdom) *(UbiComp/ISWC '22 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 97–98. https://doi.org/10.1145/3544793.3560329

[27] Md. Hedayetul Islam Shovon, D (Nanda) Nandagopal, Jia Tina Du, Ramasamy Vijayalakshmi, and Bernadine Cocks. 2015. Cognitive Activity during Web Search. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Santiago, Chile) *(SIGIR '15)*. Association for Computing Machinery, New York, NY, USA, 967–970. https://doi.org/10.1145/2766462.2767784

[28] Daniel Tranel and Hanna Damasio. 1989. Intact electrodermal skin conductance responses after bilateral amygdala damage. *Neuropsychologia* 27, 4 (1989), 381–390. https://doi.org/10.1016/0028-3932(89)90046-8

[29] Johanne R. Trippas, Damiano Spina, Paul Thomas, Mark Sanderson, Hideo Joho, and Lawrence Cavedon. 2020. Towards a Model for Spoken Conversational Search. *Information Processing & Management* 57, 2 (2020), 102162. https://doi.org/10.1016/j.ipm.2019.102162

[30] Robert T. Williams. 1972. A Table for Rapid Determination of Revised Dale-Chall Readability Scores. *The Reading Teacher* 26, 2 (1972), 158–165.

[31] Kevin Wise, Hyo Jung Kim, and Jeesum Kim. 2009. The Effect of Searching Versus Surfing on Cognitive and Emotional Responses to Online News. *Journal of Media Psychology* 21 (2009), 49–59. Issue 2. https://doi.org/10.1027/1864-1105.21.2.49

[32] Manuela Züger and Thomas Fritz. 2015. Interruptibility of Software Developers and Its Prediction Using Psycho-Physiological Sensors. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 2981–2990. https://doi.org/10.1145/2702123.2702593