







Visual and Pupillary Behavior in Neonatal Pain Assessment using Eye-Tracking

Roberto Gonçalves de Magalhães Júnior , Rafael Nobre Orsi , Tatiany Marcondes Heiderich , Marina Carvalho de Moraes Barros , Ruth Guinsburg , and Carlos Eduardo Thomaz 

Abstract—This paper introduces the application of novel eye-tracking metrics to assess visual attention and cognitive load in neonatal pain assessment. Our goal is to evaluate pediatrician experts, non-experts, and parents using the relative Explore-Exploit Ratio and the Task-Evoked Pupillary Response while analyzing the frontal faces of distinct newborns before and after painful procedures. All experiments were based on a benchmark image dataset, considering clinically relevant areas of interest. The Tobii TX300 system was used to record eye-tracking data in a closed room with controlled lighting. Our results show that visual attention described by traditional metrics does not correspond directly to the respective fixation patterns and pupillary changes quantified for all the groups of participants investigated, highlighting statistically significant differences in visual behavior between participants with or without clinical experience only when using the novel metrics proposed instead.

Link to graphical and video abstracts, and to code:
<https://latam.ieceer9.org/index.php/transactions/article/view/9742>

Index Terms—Eye-tracking, Neonatal Pain Assessment, Visual Behavior.

I. INTRODUCTION

EYE-TRACKING technology has allowed researchers to analyze visual attention and pupillary response. Since eye-tracking measures the natural visual activity of an individual, it has provided a feasible but non-invasive opportunity for the study of human visual cognition [1].

Traditional eye-tracking metrics used to study visual cognitive load have primarily focused on analyzing fixation count and duration [2]. However, there is no concrete and comprehensive guide to using eye-tracking movements to investigate a cognitive process that can be modulated by human visual behavior [3], especially when considering different areas of interest in computer interface technology.

The associate editor coordinating the review of this manuscript and approving it for publication was Guadalupe Dorantes-Mendez (*Corresponding author: Roberto Goncalves De Magalhaes*).

This study was financed in part by FEI, the Brazilian Federal Agency CAPES - Finance Code 001, and the Sao Paulo Research Foundation FAPESP (2018/13076-9).

Roberto Goncalves De Magalhaes, T. M. Heiderich, and C. E. Thomaz are with Federal University of São Paulo, São Bernardo do Campo, Brazil (e-mails: feimagjr@fei.edu.br, tatianymh@fei.edu.br, and cet@fei.edu.br).

R. N. Orsi is with Paula Souza Center, São Paulo, Brazil (e-mail: rafael.orsi@cpspos.sp.gov.br).

M. C. de M. Barros, and R. Guinsburg are with Federal University of São Paulo, São Paulo, Brazil (e-mails: marina.barros@unifesp.br, and ruth.guinsburg@unifesp.br).

In this work, we introduce a data processing methodology designed to address the exploration-exploitation paradigm and the pupillary response as relative metrics to evaluate different areas of interest in the assessment of neonatal pain. Newborns are frequently subjected to painful clinical procedures that, if prolonged, may cause short- and long-term health complications, including increased mortality and morbidity. The ability to quickly identify the presence or absence of pain is crucial to preventing the consequences of inadequate treatment or the unnecessary use of opioids¹ when pain has already subsided [4], [5].

Although newborns cannot express themselves verbally, bodily signals allow for an effective non-invasive pain assessment [6], motivating several clinical scales to be developed. Some methods are based primarily on the reactions of the subject, using measures such as crying, motor activity, and facial expressions [7]. These scales are clinically validated and applied in real-world settings, but they are inherently subjective, leading to results that may vary depending on the characteristics of the patient, the evaluator's experience and the surrounding environment [8]. Consequently, the challenge of accurately identifying and quantifying neonatal pain, coupled with the subjectivity of these assessments, remains a significant concern for the scientific community [9].

By analyzing the potential benefits of incorporating the metrics described in this work into existing eye-tracking protocols, our goal is to promote insights into the underlying mechanisms of visual cognitive processing and contribute to the development of complementary measures to assess mental load and performance in vital visual cognitive tasks like the assessment of neonatal pain. In other words, we believe that this study can contribute significantly to the field of health information technology, especially neonatal care, because our aim is to quantify the differences between experts and non-experts and explore whether specific cognitive strategies could be inferred from visual and pupillary metrics. This comparison allows for the identification and recognition of distinctive patterns in the expert behavior that can inform the design of training tools for non-experts and create diagnostic support systems that might help make the corresponding clinical decisions more accurate and individualized.

II. RELATED WORKS

Visual attention is an important process through which humans engage with the environment. We selectively focus

¹Opioids are potent analgesic and sedative medications.

on specific areas for detailed visual processing by directing our gaze. Tracking these eye movements allows us to collect valuable information that gives clues about our thoughts and intentions [10]. Therefore, investigating human visual behavior has aroused interest in research related to cognitive science in particular [11] and computer interface technology in general [12].

During a typical visual task, the observer has to constantly decide whether to remain in the currently fixated region or change their gaze to another. The observer must choose between further exploiting the information or continuing to explore other patches within a given scene [13]. The number and spatial distribution of eye movements' activities represented by saccades and short fixations characterize the visual exploration of a scene. Short fixations are often associated with search processing because their duration is insufficient for the observer to extract and interpret the associated information [14]. In contrast, long fixations indicate visual in-depth processing and characterize exploitation. However, given time constraints in observing and interpreting visual stimuli like images in decision-making tasks, there are conflicting demands between analyzing the entire scene and focusing on specific local regions. Consequently, visual behavior involves a continuous interaction between exploration and exploitation in natural, decision-making, or free-viewing visuo-cognitive tasks [3], [11], [15].

Another cognitive behavior phenomenon might occur when the observer visually scans a scene: changes in pupil size. These changes are associated with the Task-Evoked Pupillary Response (TEPR), a concept generalized by Beatty [16]. TEPR refers to the mean of a stimulus-locked pupil response observed relatively to a short pre-stimulus baseline and is commonly used to estimate arousal by assuming that pupil size is a linear transformation of task events. A substantial body of psychophysiological works indicates that TEPR might serve as a viable index of cognitive effort exertion [16]. In various areas of study, the greater cognitive efforts needed to achieve a correct response lead to larger TEPRs [17].

III. MATERIAL AND METHODS

To evaluate the exploration-exploitation paradigm and the pupillary response, we followed an eye-tracking study and used the same image dataset of newborn faces [17]. The objective of this previous study [17] was to experimentally monitor the eye movements of experts, non-experts, and parents while evaluating the presence or absence of neonatal pain.

A. Subjects and Apparatus

Volunteers (N = 102) were divided into 44 experts (4 pediatricians and 40 neonatologists, 33.48 ± 7.01 years old), 29 non-experts (39.82 ± 10.39 years old) with no professional or personal experience in the evaluation of neonatal pain, and 29 parents (30.48 ± 6.95 years old) of newborns interned in a Neonatal Intensive Care Unit. Only experts had training or clinical experience in pain assessment.

Eye-tracking data was recorded using the Tobii TX300 system (300Hz). Calibration was performed to ensure accurate

tracking of each participant's eye position before the start of the experiment. The study was carried out in a closed room with controlled lighting outside the participant's direct visual field to minimize fluctuations in pupil size caused by ambient luminosity changes. Data acquisition was performed using an auxiliary computer with Tobii Studio software and eye-tracking equipment installed and connected, respectively [17].

B. Stimuli and Experimental Procedure

Fig. 1 shows the procedure created and developed in which volunteers evaluated pain in different images of the neonatal face by assigning scores to each stimulus presented on a numerical analog scale ranging from 0 (representing no pain) to 10 (representing extreme pain) [17]. This rating scale was used primarily to promote task engagement and help maintaining participants' attentional focus throughout the whole experiment, and it was not used in the analysis. During each exploratory session, the eye-tracking equipment recorded data on how each volunteer visually assessed the stimuli, including their gaze patterns, fixation points, saccades, and overall scanning behavior. This eye movement evaluation strategy refers to the specific ways in which participants directed their attention, processed visual information, and prioritized different areas of interest while observing the stimuli.

We used a database of facial images authorized with the consent of family members or guardians not included as volunteers in the current experimental assessment. The project to construct and analyze these images was approved by the Ethics Committee for Research of the Federal University of Sao Paulo (UNIFESP), under the numbers 1299/09 and 3.116.146. All data were collected at the Hospital of Sao Paulo, a university affiliated hospital of UNIFESP.

A total of 20 images (referenced here as tasks) was chosen, representing 10 newborns. Each pair of images consists of one photograph captured while the neonate is at rest and another taken after a painful procedure, such as a clinically necessary puncture procedure [17].

The experiment began with an eye-tracking calibration, followed by standardized instructions displayed on the screen. The participants first completed two preliminary trials for familiarization, which were not included in the final analysis. During the experimental procedure, as illustrated in Fig. 1, facial images were displayed in a randomized order for 7 seconds each, followed by a 3-second response window during which participants provided their pain ratings orally. Each task began with a central fixation cross (2 seconds) to standardize the initial gaze positioning, reducing the variability in the early saccades [17].

C. Eye-Tracking Data Processing

We considered the Tobii I-VT filter in the pupil diameter signal to fill in gaps where valid data are missing [18]. After that, the remaining gaps were filled using linear interpolation [17].

Orsi *et al.* [17] have shown that 2 seconds of exposure to a facial expression is sufficient for a neonatal pain assessment. Based on this finding, we analyzed eye movement activities

by dividing the experimental timeline into two distinct phases. This time splitting allows us to distinguish between initial and later visual behavior processing, providing insights into the temporal dynamics of neonatal pain assessment.

Thus, we computed the number of occurrences of eye movement activities for each participant, considering the following experimental analysis: (1) time course, considering the signal eye-tracking data from 0 to 2 seconds, and from 2 to 7 seconds; (2) participant groups, with experts, non-experts, and parents. For all of them, we calculated the metrics compared to the corresponding fixation durations and counts as representative traditional metrics of visual attention, and the respective pupillary response diameter as a cognitive load indicator. We have also analyzed these metrics according to the following areas of interest clinically considered for facial pain assessment [19]: eyes, region between eyebrows, forehead, mouth, nasolabial grooves, chin, eyebrows, nose, and cheeks, as illustrated in Fig. 2. The gaze positions were analyzed based on these semantic clustering areas, independently of their exact spatial position in the scene, to ensure robust comparisons between time courses and participants.

The reasoning about the exploration-exploitation paradigm is to capture quantitatively the meaning that when participants explore the whole visual stimulus of the computer interface technology, the numbers of saccades and short fixations tend to increase, thereby increasing the mathematical ratio between exploration and exploitation eye movements. In contrast, when participants exploit the same parts of the stimulus, long fixations become predominant, causing a decrease in this ratio. However, long fixations are not always present in the eye-tracking signal, resulting in a division by zero, especially in the context investigated here with areas of interest. For this reason, we normalized this standard ratio, as mathematically defined by [12], proposing the following relative explore-exploit ratio

(EER):

$$EER = \frac{(saccades + short\ fixations)}{(long\ fixations) + (saccades + short\ fixations)},$$

where the parameter *saccades* represents the number of saccades; and *long fixations* and *short fixations* are the number of fixations longer than 200 milliseconds and shorter than 100 milliseconds, respectively. We set these values in all experiments as suggested by [15].

To measure changes in pupil size, we first applied a low-pass Butterworth filter to the data, following the findings in [20]. The filter was designed with a cutoff frequency (1.6 Hz), allowing only low-frequency components to pass while reducing luminance-driven effects. The first value of the pupil size was used as a baseline for each subject's task. In other words, the trials were segregated according to the neonatal image case, resulting in a TEPR per face image per participant. For each gaze point, the pupil diameter at that moment was subtracted from the baseline. The absolute deviations were then averaged over the time course to produce a single value, representing the average absolute change in the relative pupil response (RPR). The reasoning here is that higher values of RPR describe quantitatively larger deviations from the baseline and, consequently, greater changes in visual cognitive state.

IV. RESULTS

Fig. 3 shows the average of EER and Fig. 4 the average of RPR calculated from each time course and sample group of participants.

We first tested the normality of the data using the Kolmogorov-Smirnov test [21]. As the normality assumption was not met, we applied the Wilcoxon signed-rank tests [22] to determine the performance of the metrics and calculate

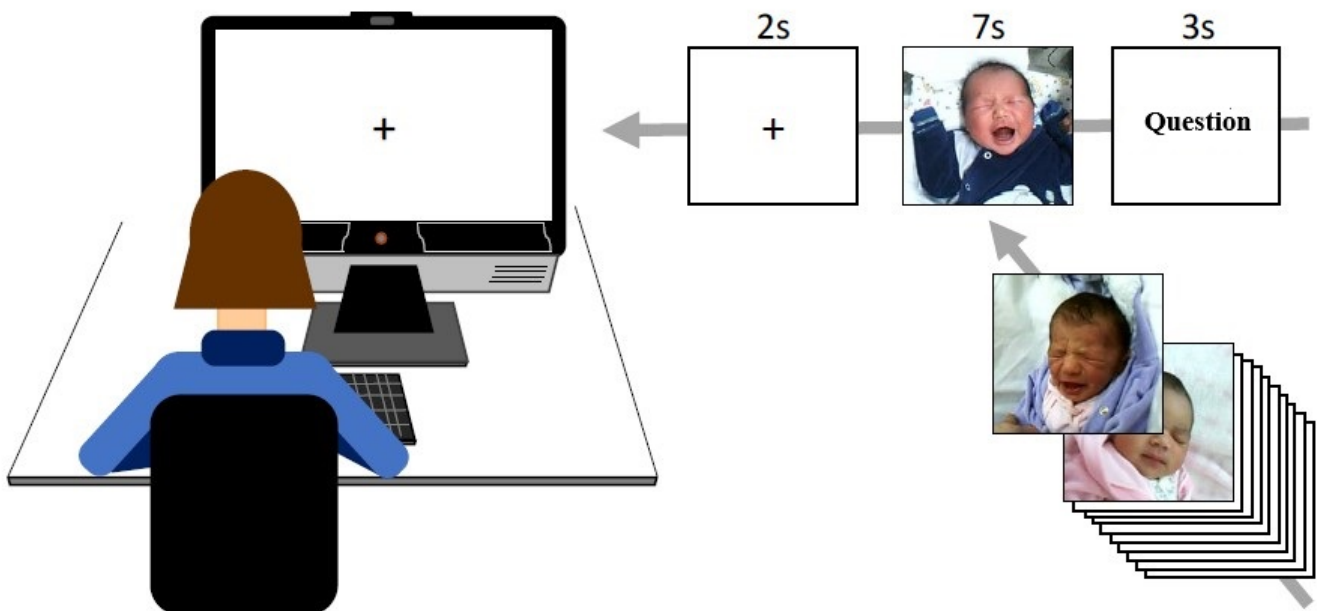


Fig. 1. Experimental procedure involving newborn faces [17].

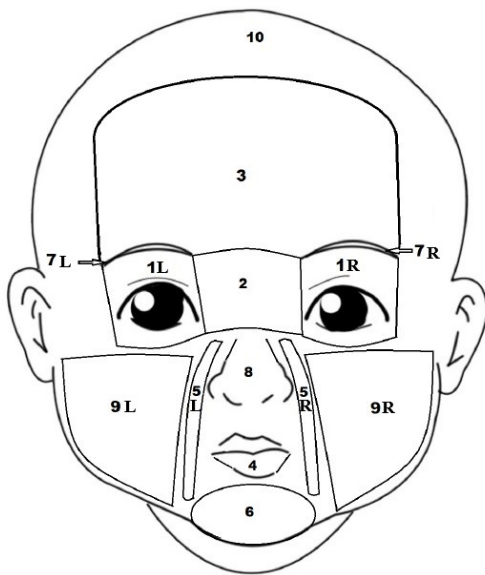


Fig. 2. Areas of interest clinically associated with the assessment of facial pain. 1) left eye (1L) and right eye (1R); 2) between the eyebrows; 3) forehead; 4) mouth; 5) left (5L) and right (5R) nasolabial grooves; 6) chin; 7) left (7L) and right (7R) eyebrows; 8) nose; 9) left (9L) and right (9R) cheeks; and 10) other regions of the face [19].

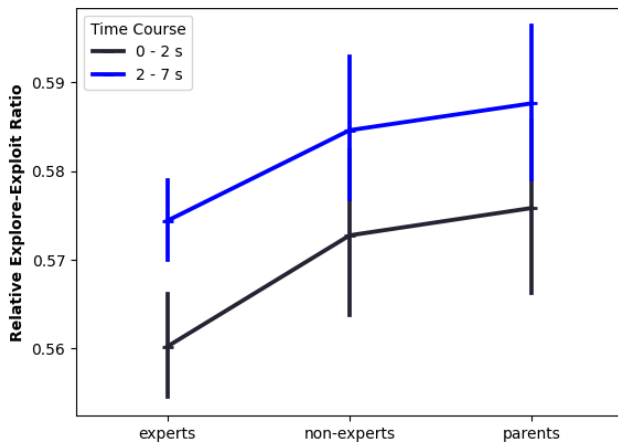


Fig. 3. Relative explore-exploit ratio calculated as the average of EER considering all tasks per time course and participant group.

the corresponding statistical differences. Specifically, for the expert group, a significant difference ($p < 0.05$) was observed between the initial (0 - 2 s) and later (2 - 7 s) intervals of their EER results (Fig. 3), with $p = 0.0056$. Non-experts and parents did not reach such statistical significance, with $p = 0.0664$ and $p = 0.0862$, respectively. Moreover, it is important to note in Fig. 3 that the experts had the lowest EER, implying that they spent more time focusing on specific areas of interest than scanning or exploring new regions.

Analogously, as illustrated in Fig. 4, the group of expert samples had a greater relative pupil response, implying a greater change in cognitive state. This result suggests that this behavior occurred because experts were in an important professional decision-making situation that could influence their

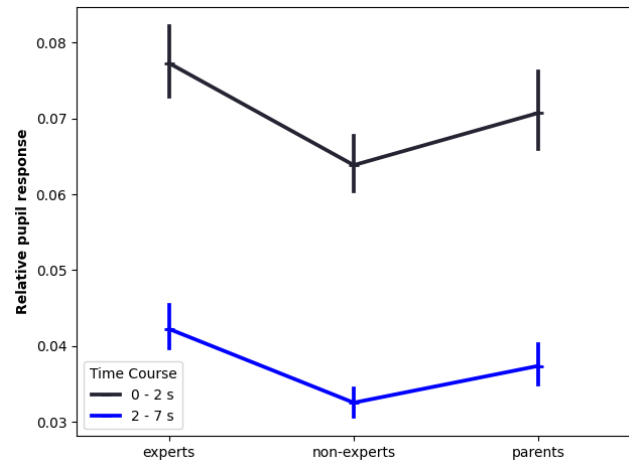


Fig. 4. Relative pupil response (RPR) calculated as the average of the absolute diameter variation from a baseline value, considering all tasks per time course and participant group.

choice to perform correctly. In contrast, for the group of non-expert samples, composed of participants with no professional or personal experience on this matter, it is reasonable that they achieved the lowest result regarding the change in cognitive state, as shown in Fig. 4. Statistical analyses confirmed these observations. For experts, the difference between the early (0 - 2 s) and later (2 - 7 s) intervals was highly significant ($p < 0.001$). Similarly, significant differences were observed for non-experts ($p < 0.001$) and parents ($p < 0.001$).

It is worth noting the inverse relationship between the EER and RPR metrics, particularly within the 0-2 s time course. During this initial period, experts exhibited lower EER values, indicating more focused visual attention, suggesting less exploratory behavior and more targeted information processing. Simultaneously, their higher RPR values reflect increased cognitive engagement. This inverse pattern suggests that, early in the task, experts not only directed their gaze more efficiently but also experienced greater cognitive effort during the decision-making process.

Fig. 5 provides a visual mapping of the proposed (EER and RPR) and traditional eye-tracking metrics (fixation count and fixation duration) in all the groups of participants (experts, non-experts, and parents) and time intervals (0-2 seconds and 2-7 seconds) considered. Each facial colored map represents the normalized metric values distributed among the specific areas of interest, highlighting variations in visual patterns. For instance, in the first 2 seconds, the chin, cheeks, and nasolabial grooves were the most dominant areas of interest in cognitive load, according to our metrics. However, the traditional eye-tracking measures - fixation count and fixation duration - indicate that the nose was the dominant area of interest in all groups. This result is consistent with previous research [23], suggesting that the nose serves as a central spatial position within facial images, facilitating a holistic analysis of neonatal pain expressions.

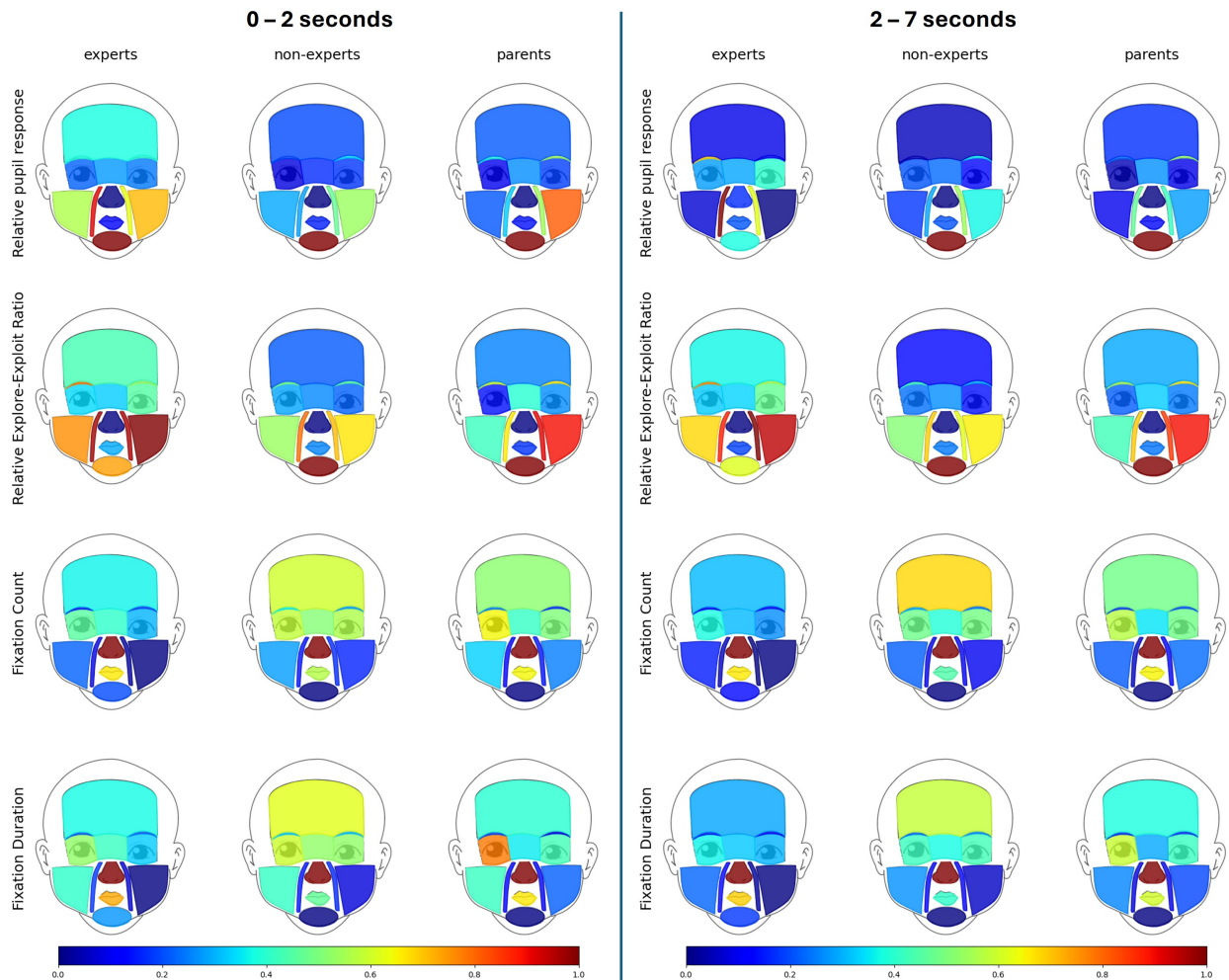


Fig. 5. Mapping of all areas of interest. Each color represents the average metric value for each area of interest, calculated for all tasks, time courses, and participant groups.

V. DISCUSSION

This work focused on studying the visual behavior and pupillary response of experts, non-experts, and parents while analyzing frontal face images of distinct newborns before and after painful procedures.

Comparing the time courses across all sample groups of participants, we showed (Fig. 4) that the RPR within the first two seconds was significantly larger than those observed in the remaining signal, indicating an immediate cognitive response at the beginning of the visual stimuli for all participants. Furthermore, we found (Fig. 3) that the EER was lower in the same initial time interval considered. However, this effect reached a statistically significant difference only within the expert sample group. These results, when considered alongside the findings of Orsi et al. [17], highlight that pupillary responses and associated visual behaviors play a critical role in this decision-making process, particularly within the first two seconds of exposure to the visual stimuli under investigation. Consequently, it reinforces and emphasizes the rapid and decisive nature of the visual cognitive process in neonatal pain assessment, particularly to the experts.

Tamanaka et al. [19] quantified that the nose area is consistently a focal visual point during the pain assessment of newborns for the same groups, regardless of the number of fixations and the time spent looking at the same areas of interest considered here. We extended this finding incorporating additional measures. As a result, we identified that the areas of interest to which participants direct their gaze do not necessarily reveal the cognitive meaning of what they look at. For example, while the nose may capture most of the visual attention (Fig. 5, third and fourth rows), the corresponding relative pupil response and explore-exploit ratio (Fig. 5, first and second rows, respectively) indicate that this area may not contribute to the decision-making process as the other regions of the face. This outcome reveals an important difference between the face regions participants look at and the way in which they process the visual information they gather during the cognitive process of assessing neonatal pain.

A common issue related to the explore-exploit paradigm is how to precisely determine the numerical boundaries between short and long fixations in a specific visual-cognitive experimental context and whether long fixations always reflect cognitive efforts related to the corresponding fixation areas.

For instance, the participant may be in a visually stationary state just because of the experimental setup, in which the participant is normally asked not to look away from the screen. For this reason, we also used the RPR metric in conjunction with the EER to better explain visual behavior in cognitive decision-making tasks. We combined as well fixation count and duration as a reference for comparison. These results are presented in Fig. 6, which shows the main areas highlighted by each combination of metrics, confirming the distinctive importance of the decision making of the nasolabial grooves for experts in both time intervals considered.

VI. CONCLUSION

This work approached the cognitive task of assessing neonatal pain, investigating visual and pupillary behavior in such a vital decision-making process. Intuitively, it is natural to think that the regions of visual stimuli with the greatest number and duration of fixations are the ones of greatest interest to the participants and consequently where they promote the greatest cognitive load. However, we show here that these traditional metrics may not be the main gaze results to properly discriminate the most dominant cognitive areas of interest, nor the expected differences between the groups of participants with or without clinical experience.

By introducing and applying the EER and RPR metrics, we provide new insights into how professionals and non-professionals visually engage with newborn facial expressions to assess pain. These findings underscore the potential for improving clinical pain assessment in neonatology, supporting the development of more objective, nuanced tools and training protocols that go beyond standard fixation-based analysis.

The central contribution of this study lies in the advancement of methods for understanding how visual and cognitive mechanisms are specifically related to the detection and evaluation of pain in infants - a task that has significant implications for early intervention, patient outcomes, and clinical decision-making in neonatal care. Nevertheless, we believe that our methodology can be used in a broader context involving other cognitive tasks.

Although the sample size of the tasks is limited, the exploratory data set was selected due to the availability of high-quality experimental data with ethical approval. In particular, other studies using the same dataset have reported significant results, as in [17], [23]. In future work, we suggest incorporating a larger and more diverse dataset to generalize and extend current findings. Additionally, we will consider including short and long saccade distances in the EER metric. Analogously to how short and long fixation durations are understood, we hypothesize that a longer saccade amplitude corresponds to greater visual exploration. By analyzing these saccade patterns as well, we aim to gain further insight into the strategies individuals use to engage visually with their environment for decision making.

REFERENCES

- [1] K. A. Phillips, K. L. Bales, J. P. Capitanio, A. Conley, P. W. Czoty, B. A. Hart, *et al.*, "Why primate models matter," *Am. J. Primatol.*, vol. 76, pp. 801–827, 2014. [Online]. Available: <https://doi.org/10.1002/ajp.22281>
- [2] P. M. Fitts, R. E. Jones, and J. L. Milton, "Eye movements of aircraft pilots during instrument-landing approaches," *Ergonomics*, Major Writings, 2004, p. 56. [Online]. Available: <https://www.semanticscholar.org/paper/227324185>
- [3] V. Skaramagkas, G. Giannakakis, E. Ktistakis, D. Manousos, I. Karatzanis, N. S. Tachos, *et al.*, "Review of eye-tracking metrics involved in emotional and cognitive processes," *IEEE Rev. Biomed. Eng.*, vol. 16, pp. 260–277, 2021. [Online]. Available: <https://doi.org/10.1109/RBME.2021.3066072>
- [4] R. C. X. Balda and R. Guinsburg, "A linguagem da dor no recém-nascido," *Doc. Científico do Dep. de Neonatol., Soc. Bras. de Pediatria*, 2018.
- [5] G. Zamzmi, R. Kasturi, D. Goldgof, R. Zhi, T. Ashmeade, and Y. Sun, "A review of automated pain assessment in infants: features, classification tasks, and databases," *IEEE Rev. Biomed. Eng.*, vol. 11, pp. 77–96, 2017. [Online]. Available: <https://doi.org/10.1109/RBME.2017.2777907>
- [6] R. V. E. Grunau and K. D. Craig, "Pain expression in neonates: facial action and cry," *Pain*, vol. 28, no. 3, pp. 395–410, 1987. [Online]. Available: [https://doi.org/10.1016/0304-3959\(87\)90073-X](https://doi.org/10.1016/0304-3959(87)90073-X)
- [7] R. Guinsburg, "Avaliação e tratamento da dor no recém-nascido," *J. Pediatr. (Rio J)*, vol. 75, no. 3, pp. 149–160, 1999. [Online]. Available: <https://doi.org/10.25060/residpediatr>
- [8] R. C. X. Balda and R. Guinsburg, "Perceptions of neonatal pain," *Neoreviews*, vol. 8, no. 12, pp. e533–e542, 2007. [Online]. Available: <https://doi.org/10.1542/neo.8-12-e533>
- [9] J. C. A. Soares, M. C. M. Barros, G. V. T. Silva, L. P. Carlini, T. M. Heiderich, R. N. Orsi, *et al.*, "Looking at neonatal facial features of pain: do health and non-health professionals differ?," *J. Pediatr.*, vol. 98, no. 4, pp. 406–412, 2022. [Online]. Available: <https://doi.org/10.1016/j.jpmed.2021.10.006>
- [10] P. Majoranta and A. Bulling, "Eye-tracking and eye-based human-computer interaction," in *Adv. Physiol. Comput.*, pp. 39–65, 2014. [Online]. Available: https://doi.org/10.1007/978-1-4471-6392-3_3
- [11] R. R. Gameiro, K. Kaspar, S. U. König, S. Nordholt, and P. König, "Exploration and exploitation in natural viewing behavior," *Sci. Rep.*, vol. 7, no. 1, p. 2311, 2017. [Online]. Available: <https://doi.org/10.1038/s41598-017-02526-1>
- [12] J. H. Goldberg and X. P. Kotval, "Computer interface evaluation using eye movements: methods and constructs," *Int. J. Ind. Ergon.*, vol. 24, no. 6, pp. 631–645, 1999. [Online]. Available: [https://doi.org/10.1016/S0169-8141\(98\)00068-7](https://doi.org/10.1016/S0169-8141(98)00068-7)
- [13] N. Malem-Shinitzki, M. Opper, S. Reich, L. Schwetlick, S. A. Seelig, and R. Engbert, "A mathematical model of exploration and exploitation in natural scene viewing," *bioRxiv*, 2020. [Online]. Available: <https://doi.org/10.1101/2020.04.16.044677>
- [14] V. Peysakhovich, "Study of pupil diameter and eye movements to enhance flight safety," Ph.D. dissertation, Univ. Toulouse, France, 2016. [Online]. Available: <https://doi.org/10.13140/RG.2.2.34241.35680>
- [15] B. R. Manor and E. Gordon, "Defining the temporal threshold for ocular fixation in free-viewing visuo-cognitive tasks," *J. Neurosci. Methods*, vol. 128, no. 1, pp. 85–93, 2003. [Online]. Available: [https://doi.org/10.1016/S0165-0270\(03\)00151-1](https://doi.org/10.1016/S0165-0270(03)00151-1)
- [16] J. Beatty, "Task-evoked pupillary responses, processing load, and the structure of processing resources," *Psychol. Bull.*, vol. 91, pp. 276, 1982. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/7071262>
- [17] R. N. Orsi, L. P. Carlini, T. M. Heiderich, G. V. T. Silva, J. D. C. A. Soares, R. C. X. Balda, *et al.*, "Visual attention during neonatal pain assessment: A 2-s exposure to a facial expression is sufficient," *Electron. Lett.*, vol. 59, no. 6, p. e12756, 2023. [Online]. Available: <https://doi.org/10.1049/ell2.12756>
- [18] A. Olsen, "The Tobii I-VT fixation filter." [Online]. Available: http://www.vinix.co.kr/ivt_filter.pdf
- [19] F. G. Tamanaka, L. P. Carlini, T. M. Heiderich, R. C. X. Balda, M. C. M. Barros, R. Guinsburg, and C. E. Thomaz, "Neonatal pain assessment: A Kendall analysis between clinical and visually perceived facial features," *Comput. Methods Biomech. Biomed. Eng. Imag. Vis.*, vol. 11, no. 3, pp. 331–340, 2022. [Online]. Available: <https://doi.org/10.1080/21681163.2022.2044909>
- [20] V. Peysakhovich, M. Causse, S. Scannella, and F. Dehais, "Frequency analysis of a task-evoked pupillary response: Luminance-independent measure of mental effort," *Int. J. Psychophysiol.*, vol. 97, no. 1, pp. 30–37, 2015. [Online]. Available: <https://doi.org/10.1016/j.ijpsycho.2015.04.019>
- [21] R. H. C. Lopes, "Kolmogorov-Smirnov test," in *Int. Encycl. Stat. Sci.*, Springer, pp. 718–720, 2011. [Online]. Available: https://doi.org/10.1007/978-3-642-04898-2_326

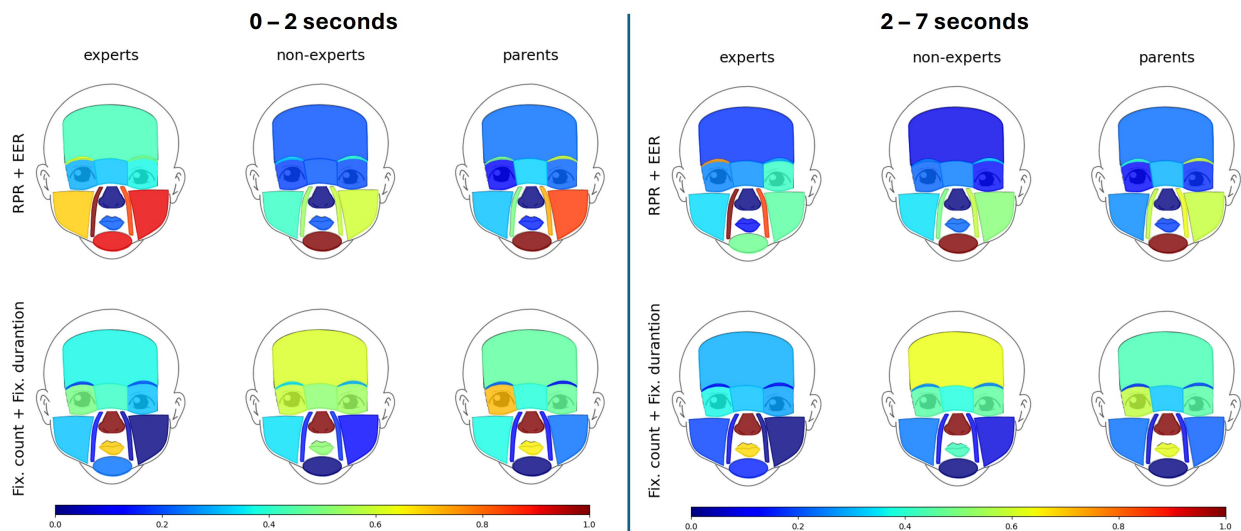


Fig. 6. Mapping of RPR in conjunction with EER; and fixation count in conjunction with fixation duration for each area of interest, calculated for all tasks, time courses, and participant groups.

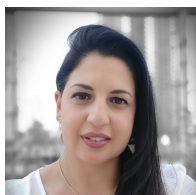
- [22] J. W. Pratt, "Remarks on zeros and ties in the Wilcoxon signed rank procedures," *J. Am. Stat. Assoc.*, vol. 54, pp. 655–667, 1959. [Online]. Available: <https://doi.org/10.1080/01621459.1959.10501526>
- [23] L. P. Carlini, G. A. S. Coutrin, L. A. Ferreira, J. C. A. Soares, G. V. T. Silva, T. M. Heiderich, *et al.*, "Human vs machine towards neonatal pain assessment: A comprehensive analysis of the facial features extracted by health professionals, parents, and convolutional neural networks," *Artif. Intell. Med.*, vol. 147, p. 102724, 2024. [Online]. Available: <https://doi.org/10.1016/j.artmed.2023.102724>



Roberto Gonçalves de Magalhães Júnior received his B.Sc. degree in Automation and Control Engineering and his M.Sc. degree in Electrical Engineering from the University College FEI, Sao Paulo, Brazil, in 2016 and in 2019, respectively. He is currently working toward the Ph.D. degree in the research area of Image and Signal Processing at FEI.



Rafael Nobre Orsi is an Associated Professor and Researcher at the Paula Souza Center, Government of the State of São Paulo. He holds the M.Sc. and Ph.D. degrees in Electrical Engineering from University College FEI, Sao Paulo, Brazil, and a B.Sc. degree in Mathematics. His research interests include artificial intelligence, machine learning, and statistical methods for pattern recognition.



Tatiany Marcondes Heiderich is Physiotherapist (2000) from Universidade de Mogi das Cruzes (UMC), Sao Paulo, Brazil, Master's in Biomedical Engineering (UMC, 2006), PhD in Pediatric Sciences (2015) from Escola Paulista de Medicina of the Federal University of São Paulo (EPM-UNIFESP), and PhD in Electrical Engineering from University College FEI (2024).



Marina Carvalho de Moraes Barros graduated in Medicine from the Escola Paulista de Medicina - Federal University of São Paulo, with specializations in Pediatrics (1990) and Neonatology (1992), and Master's (1996) and Doctorate (2005) from the same institution, including Master's in Business Administration from Ibmec - São Paulo (2009). She is Professor of the Postgraduate Program in Pediatrics and Sciences Applied to Pediatrics (since 2011) and Executive Editor of the Editorial Board of *Revista Paulista de Pediatria* since July 2015.



Ruth Guinsburg is Full Professor of Neonatal Medicine at Escola Paulista de Medicina of the Federal University of São Paulo (EPM-Unifesp), Head of the Neonatal Intensive Care Unit at São Paulo Hospital/EPM-Unifesp, Co-chair of the Neonatal Resuscitation Program of the Brazilian Society of Pediatrics, Scientific Coordinator of the Brazilian Network on Neonatal Research and Editor-in-Chief of *Revista Paulista de Pediatria*.



Carlos Eduardo Thomaz is Full Professor of Electrical Engineering at University College FEI, Sao Paulo, Brazil. He is head of the Image Processing Lab at FEI. Carlos Thomaz received, in 1993, his B.Sc. degree in Electronic Engineering from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil. After working for six years in industry, he obtained the M.Sc. degree in Electrical Engineering from PUC-Rio in 1999. In October 2000, he joined the Department of Computing at Imperial College London, UK, where he obtained the Ph.D. degree in Computing in 2004. His main research interests are in Pattern Recognition, Cognitive Perception and Machine Learning.