Visual attention during neonatal pain assessment: A 2-second exposure to a facial expression is sufficient

Rafael Orsi[1], Lucas Carlini[1], Tatiany Heiderich[2], Giselle Silva[2], Juliana Soares[2], Rita Balda[2], Marina Barros[2], Rute Guinsburg[2], and Carlos Thomaz[3]

[1]FEI
[2]Unifesp EPM
[3]Centro Universitario da FEI

November 28, 2022

Abstract

Facial expression has been widely used in clinical practice to assess pain in newborns. However, the inherent visual attention required to make such vital inference is poorly understood. It is also unknown whether this inference occurs differently when comparing health professionals with other adults. To investigate these issues, we have recorded and monitored the pupil size signal of 102 subjects (44 experts, 29 parents, and 29 non-experts) while visually analyzing 20 frontal face images of 10 distinct newborns after a painful procedure and at painless rest. Our experimental results have showed that neonatal pain assessment is more cognitively demanding when analyzing the presence of pain rather than its absence. Moreover, our results disclose that a 2-second exposure to a facial expression is sufficient to make this assessment, regardless whether done by health professionals or non-health ones, suggesting that this highly specific visual task is not driven by clinical experience.

Visual attention during neonatal pain assessment: A 2-second exposure to a facial expression is sufficient

R. N. Orsi, L. P. Carlini, T. M. Heideirich, G. V. T. Silva, J. C. A. Soares, R. C. X. Balda, M. C. M. Barros, R. Guinsburg and C. E. Thomaz

Facial expression has been widely used in clinical practice to assess pain in newborns. However, the inherent visual attention required to make such vital inference is poorly understood. It is also unknown whether this inference occurs differently when comparing health professionals with other adults. To investigate these issues, we have recorded and monitored the pupil size signal of 102 subjects (44 experts, 29 parents, and 29 non-experts) while visually analyzing 20 frontal face images of 10 distinct newborns after a painful procedure and at painless rest. Our experimental results have showed that neonatal pain assessment is more cognitively demanding when analyzing the presence of pain rather than its absence. Moreover, our results disclose that a 2-second exposure to a facial expression is sufficient to make this assessment, regardless whether done by health professionals or non-health ones, suggesting that this highly specific visual task is not driven by clinical experience.

Introduction: Newborn infants are exposed to painful experiences, routinely practised clinically, that might increase their short and long-term morbidity and mortality. Since the pioneering Neonatal Facial Coding System [1] proposed in 1987, facial expression has been widely used in clinical practice to non-invasively evaluate pain among newborns.

However, despite the fact that recognizing pain is a natural human skill [2] and faces are visual patterns that can be specifically recognized by humans with not only natural but also heritable ability [3, 4], the inherent visual attention required to properly make this evaluation is less established. In other words, when assessing the presence or absence of pain in newborns using faces, it is poorly understood whether such facial expression inference occurs more rapidly or effortlessly when comparing health professionals, with specific knowledge in decoding pain during the neonatal period, to non-health ones with personal or no experience on this matter.

To investigate these issues, we have carried out in this work a controlled visual attention procedure to record and monitor the involuntary pupil dilations and constrictions of experts, parents and non-experts' subjects while analyzing frontal face images of distinct newborns before and after a painful procedure. To the best of our knowledge, this is the first work that has showed experimentally that neonatal pain assessment is fast, but not intuitive and neither driven entirely by clinical experience.

Materials - Subjects, Apparatus and Stimuli: This work 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 (HSP), a university affiliated hospital of UNIFESP.

A total number of 102 adults participated in this study on a voluntary basis. They were divided into three groups: 44 experts (4 pediatricians and 40 neonatologists, 33.48 ± 7.01 years old) with clinical experience in pain assessment (N-PASS [5]), 29 parents (30.48 ± 6.95 years old) of newborns interned in the Neonatal Unit of the HSP without training in pain assessment, and 29 non-experts (39.82 ± 10.39 years old) without training or clinical experience in pain assessment. All subjects had normal or corrected to normal vision. Written informed consent was obtained from all subjects.

Pupil sizes, more specifically, pupil diameters, were recorded with an on-screen Tobii TX300 equipment that comprises a table-mounted eye tracker unit integrated to the lower part of a 23in TFT screen display. The eye tracker performs binocular tracking at a data sampling rate of 300Hz. Calibration, monitoring and data collection were performed as implemented in the Tobii Studio software running on an attached notebook (Core i7, 16Gb RAM and Windows 7).

Stimuli consisted of 20 face images of 10 different full-term newborns from the UNIFESP neonatal face database [6]. All face images are colorful and taken against a non-homogenous background in an upright frontal position, without artifacts that might overlap parts of the newborn's face. Each pair of face images is composed of one picture of the newborn at painless rest (negative stimulus) and another picture of the same newborn after a painful procedure (positive stimulus) routinely practised clinically, such as blood specimen collection, Hepatitis B vaccine or Inborn Errors of Metabolism. All visual stimuli were presented at a distance of 60cm using the 23in TFT monitor with a screen resolution of 1280 x 1024 pixels. Figure 1 illustrates some of the face images presented.

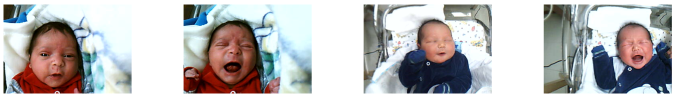

Fig. 1 *Examples of the face stimuli used in this work. From left to right, the first and third images are negative stimuli (newborns at painless rest) and the second and fourth positive ones (same newborns after a painful procedure).*

Methods - Visual Attention Procedure: Subjects were seated in front of the eye tracker, instructed about the apparatus and assessment, and asked to evaluate the visual attention stimuli as accurately as possible. All assessments were carried out in a closed testing room with artificially controlled lighting.

Every visual attention procedure began with an introductory screen with instructions to the subject, presentation of two prior trials to allowing a better understanding of the experiment by the subject, and then the proper experimental trials [7, 8]. On each trial, a central fixation cross was presented for 2 seconds followed by a neonatal face image randomly selected (Latin Square Design [9]) and shown non-centralized on a new screen. The face image was displayed for 7 seconds and followed by a scoring question on a new screen that required a response in relation to its corresponding pain assessment. The subject had 3 seconds to verbally answer the pain score of the face image displayed, using a numerical rating scale from 0 (no pain) to 10 (severe pain). Each response was subsequently followed by the central fixation cross, which preceded the next face stimulus until all the 20 face images were presented and evaluated.

The total time to perform the visual attention procedure was 5 minutes for each subject.

Methods - Pupil Signal Processing: Pupil diameters of the subjects' right eyes were acquired in millimeters directly from the eye tracker using the Tobii Studio software. We considered only pupil data from subjects for whom on average 70% or more of their gaze samples were collected by the eye tracker. Loss of pupil data owing to blinking or other small temporary occlusions was addressed based on the Tobii I-VT fixation filter approach [10], following its gap fill-in and moving average (with window size parameter equals to 5 samples) functions for linear interpolation and noise reduction, respectively. This signal processing was performed for all pupil sizes of all face stimuli and all subjects. Moreover, for each neonatal face trial onset of each subject, the first pupil size value was used as baseline and subtracted from its subsequent values, adjusting all raw pupil sizes originally acquired to their corresponding smoother pupil signal changes.

Methods - Classification Scoring: All numerical ratings verbally answered by the subjects on either positive or negative face stimuli were classified considering the following scoring intervals [11]: 0-2 (negative, absence of pain), 3-5 (positive, low/mild presence of pain), and 6-10 (positive, moderate/severe presence of pain).

Results: We first compared the classification results (Figure 2) of the experts, parents, and non-experts sample groups. Experts correctly classified the positive and negative face stimuli with mean accuracies of 96.59% ($\pm 5.27\%$) and 86.36% ($\pm 14.46\%$), respectively. Parents with 89.31% ($\pm 11.94\%$) and 85.17% ($\pm 15.65\%$), and non-experts with 92.67% ($\pm 10.27\%$) and 72.00% ($\pm 28.67\%$). Using analysis of variance [12], the overall (positive plus negative face stimuli) classification results between experts and non-experts were statistically different ($p < 0.05$) and there were differences statistically significant between positive and negative stimuli within the experts ($p < 0.05$) and non-experts ($p < 0.05$) sample groups.

These results (Figure 2) have showed that experts have performed as well as parents and better than non-experts when scoring the presence or absence of pain in frontal face images of newborns. Neither the former nor the latter are surprising results. Balda et. al [13] originally reported that there was no difference in the recognition accuracy of facial expression of pain in newborns between health professionals and parents when analyzing the face of full-term infants in an analogous experiment using printed photographs. Recently, researchers have found as well that such classification accuracy is not only similar, but also based holistically on the same facial parts [7, 11, 14, 8]. However, since the sample group of non-experts is composed of subjects with neither professional nor personal experience on attempting to score pain in newborns, non-experts have in general overestimated the presence of pain in the experimental setting carried out, showing the worst classification results on the painless rest situations of negative stimuli.

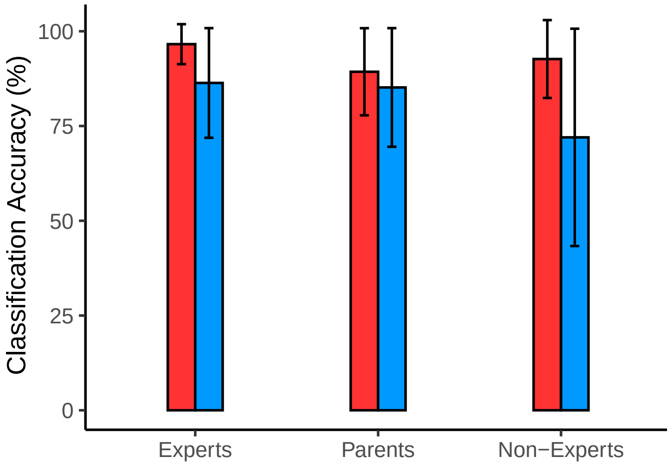

Fig. 2 *Classification accuracy on positive (in red) and negative (in blue) stimuli of experts, parents, and non-experts.*

To analyse the visual attention process, we compared the mean of the pupil diameter changes during the 7-second timeline of all positive and negative stimuli presented to the sample groups of experts (Figure 3), parents (Figure 4) and non-experts (Figure 5). All curves (Figures 3, 4 and 5) indicate that the pupil responses for both positive and negative assessments and all sample groups present a common temporal pattern with the following main transients or inflexion points: an initial transient of decrease in the pupil size from immediately after the neonatal face trial onset to about 0.063s (minimum value, from experts for negative stimuli) and 0.257s (maximum value, from parents for positive and negative stimuli); an intermediate transient of steep increase in the pupil size with peak reached at about 1.640s (minimum value, from non-experts for positive stimuli) and 1.770s (maximum value, from parents for positive stimuli); and a termination transient of smooth decrease in the pupil size.

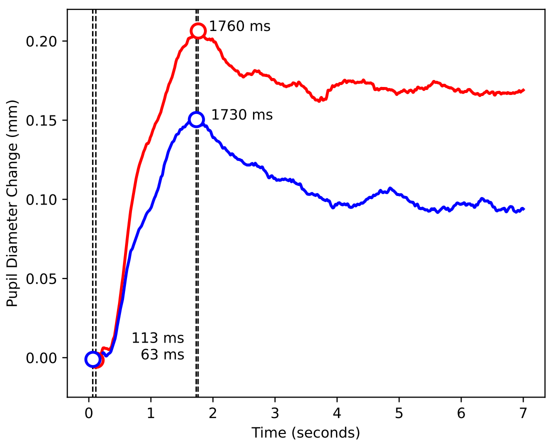

Fig. 3 *Experts' mean of the pupil diameter changes (in millimeters) during the 7-second timeline of positive (in red) and negative (in blue) face stimuli.*

All curves (Figures 3, 4 and 5) also illustrate larger subjects' pupil changes when evaluating newborn face images after a painful procedure (positive stimuli) than at painless rest (negative stimuli) for all sample groups of subjects. Using multivariate analysis of variance [12] (with the first principal components that retain at least 95% of the corresponding

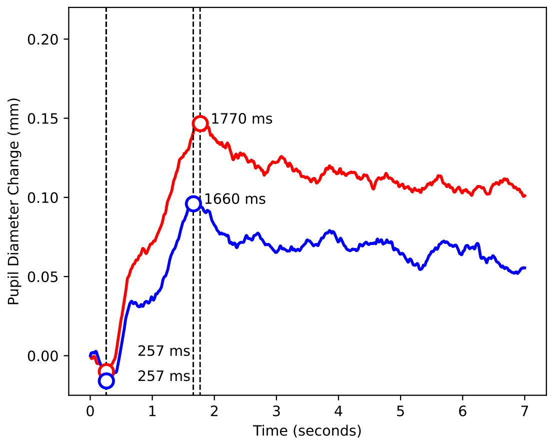

Fig. 4 *Parents' mean of the pupil diameter changes (in millimeters) during the 7-second timeline of positive (in red) and negative (in blue) face stimuli.*

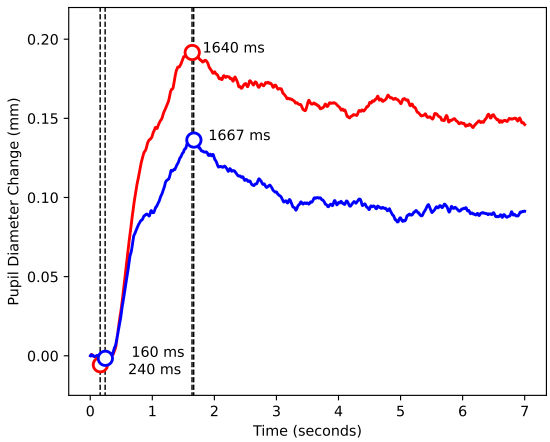

Fig. 5 *Non-experts' mean of the pupil diameter changes (in millimeters) during the 7-second timeline of positive (in red) and negative (in blue) face stimuli.*

total variance), these differences were, however, statistically significant ($p < 0.05$) for experts only. In fact, the Pearson correlation (r) analyzes of the timeline pupil diameter changes indicate strong statistically similarity between all these curves and sample groups. Quantitatively, for positive stimuli, $r = 0.975$ between experts and parents, $r = 0.989$ between experts and non-experts, and $r = 0.968$ between parents and non-experts. Analogously, for negative stimuli, $r = 0.921$ between experts and parents, $r = 0.980$ between experts and non-experts, and $r = 0.920$ between parents and non-experts.

These results (Figures 3, 4 and 5) based on the recording, monitoring, and signal processing of the pupil diameter changes of all subjects have identified that a short exposure to a facial expression is sufficient to assess the presence or absence of pain in newborns. The initial decrease in the pupil diameter in all curves immediately after the neonatal face trial on set is related to a well-known phenomenon called light reflex [15, 16, 17, 18]. The striking pattern of these pupil size changes, though, is the common steep increase from the baseline followed by a slower decrease for both positive and negative stimuli and all sample groups of subjects. Since pupil dilation is positively associated with increased cognitive load [19, 20, 17] and such information-processing load has happened here essentially in the first 2 seconds of face stimuli exposure, with no statistical differences between the sample groups of subjects, our findings show that recognizing facial expression of pain in newborns can be characterized as a fast specific visual attention task that is not driven entirely by clinical experience, but cannot be carried out effortlessly either.

Conclusion: This work discloses that as minimal an exposure time as two seconds is sufficient for us to make our visual pain inference from newborn face images. Experience on this matter, either professional or personal, increases accuracy in such inference, but requires the same first 2 seconds of face stimuli exposure. Since neonatal pain assessment is a vital human decision making, our experimental results suggest that it might be possible to create a more distinctive or selective approach for the analysis of facial expression of newborns with educational implications to neonatal pain diagnosis and training in clinical practice.

Acknowledgment: This work has been supported by the Brazilian agencies FAPESP (2018/13076-9), CNPq and CAPES.

R. N. Orsi, L. P. Carlini, T. M. Heiderich and C. E. Thomaz (*FEI, Brazil*), G. V. T. Silva, J. C. A. Soares, R. C. X. Balda, M. C. M. Barros and R. Guinsburg (*UNIFESP, Brazil*)

E-mail: rafaelorsi@fei.edu.br

References

1 R.V.E. Grunau and K.D. Craig: Pain expression in neonates: Facial action and cry, *Pain*, 1987, 28, 395-410.
2 K.S. Deyo, K.M. Prkachin and S.R. Mercer: Development of sensitivity to facial expression of pain, *Pain*, 2004, 107, 16-21.
3 J.B. Wilmer, L. Germine, C.F. Chabris, et al.: Human face recognition ability is specific and highly heritable, *Proceedings of the National Academy of Sciences of the United States of America*, 2010, 107(11), 5238-5241.
4 E. McKone and R. Palermo: A strong role for nature in face recognition, *Proceedings of the National Academy of Sciences of the United States of America*, 2010, 107(11), 4795-4796.
5 P. Hummel, M. Puchalski, S.D. Creech and M.G. Weiss: Clinical reliability and validity of the N-PASS: neonatal pain, agitation and sedation scale with prolonged pain, *Journal of Perinatology*, 2008, 28(1),55-60.
6 T.M. Heiderich, A.T.F.S. Leslie, and R. Guinsburg: Neonatal procedural pain can be assessed by computer software that has good sensitivity and specificity to detect facial movements, *Acta Paediatrica*, 2015, 104(2), e63-e69.
7 L.P. Carlini, J.C.A. Soares, G.V.T. Silva, et al.: A visual perception framework to analyse neonatal pain in face images, *Lecture Notes in Computer Science*, 2020, 12131, pp. 233-243.
8 F.G. Tamanaka, L.P. Carlini, T.M. Heiderich, et al.: Neonatal pain assessment: A Kendall analysis between clinical and visually perceived facial features, *Computer Methods in Biomechanics and Biomedical Engineering: Imaging Visualization*, 2022, pp. 1-10.
9 A.D. Keedwell and J. Dénes, *Latin squares and their applications*, Elsevier, 2015.
10 A. Olsen, *The Tobii I-VT fixation filter*, 2012.
11 M.C.M. Barros, C.E. Thomaz, G.V.T. Silva, et al.: Identification of pain in neonates: The adults' visual perception of neonatal facial features, *Journal of Perinatology*, 2021, 41, pp. 2304-2308.
12 R.A. Johnson and D.W. Wichern, *Applied Multivariate Statistical Analysis*, New Jersey: Prentice Hall, 1998.
13 R.C.X. Balda, R. Guinsburg, M.F.B. Almeida, et al.: The recognition of facial expression of pain in full-term newborns by parents and health professionals, *Archives of Pediatrics Adolescent Medicine*, 2000, 154(10), pp. 1009-1016.
14 J.C.A. Soares, M.C.M. Barros, G.V.T. Silva, et al.: Looking at neonatal facial features of pain: Do health and non-health professionals differ?, *Jornal de Pediatria*, 2022, 98(4), pp. 406-412.
15 E.H. Hess and S.B. Petrovich: Pupillary behavior in communication, *Nonverbal Behavior and Communication*, 1987, pp. 327-348.
16 J. Beatty and B. Lucero-Wagoner: The pupillary system, *Handbook of Psychophysiology*, 2000, 2, pp. 142-162.
17 T. Partala and V. Surakka: Pupil size variation as an indication of affective processing, *International Journal of Human-Computer Studies*, 2003, 59(1-2), pp. 185-198.
18 M.M. Bradley, L. Miccoli, M.A. Escrig, et al.:The pupil as a measure of emotional arousal and autonomic activation, *Psychophysiology*, 2008, 45(4), pp. 602-607.
19 D. Kahneman and J. Beatty: Pupillary diameter and load on memory, *Science*, 1966, 154, pp. 1583-1585.
20 J. Hyona, J. Tommola, and A.M. Alaja: Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks, *The Quarterly Journal of Experimental Psychology*, 1995, 48(A), pp. 598-612.