

Physiological heatmaps: a tool for visualizing users' emotional reactions

François Courtemanche¹  · Pierre-Majorique Léger² · Aude Dufresne³ · Marc Fredette⁴ · Élise Labonté-LeMoine¹ · Sylvain Sénécal⁵

Received: 26 January 2017 / Revised: 4 August 2017 / Accepted: 8 August 2017
© The Author(s) 2017. This article is an open access publication

Abstract Practitioners in many fields of human-computer interaction are now using physiological data to measure different aspects of user experience. The dynamic nature of physiological data offers a continuous window to the users and allows a better understanding of their experience while interacting with a system. However, in order to be truly informative, physiological signals need to be closely linked to users' behaviors and interaction states. This paper presents an analysis method that provides a direct visual interpretation of users' physiological signals when interacting with an interface. The proposed physiological heatmap tool uses eyetracking data along with physiological signals to identify regions where users are experiencing different emotional and cognitive states with a higher frequency. The method was evaluated in an experiment with 44 participants. Results show that physiological heatmaps are able to identify emotionally significant regions within an interface better than standard gaze heatmaps. Applications of the method to different fields of HCI research are also discussed.

Keywords Physiological computing · Eyetracking · Affective computing, heatmap, user experience

✉ François Courtemanche
francois.courtemanche@hec.ca

¹ Tech3lab, HEC Montréal, 3000 Côte Sainte-Catherine, H3T 2A7, Montréal, QC, Canada

² Department of Information Science, HEC Montréal, 3000 Côte Sainte-Catherine, H3T 2A7 Montréal, QC, Canada

³ Department of Communication, University of Montréal, 2900 Edouard-Montpetit, H3T 1J4, Montréal, QC, Canada

⁴ Departement of Statistics, HEC Montréal, 3000 Côte Sainte-Catherine, H3T 2A7, Montréal, QC, Canada

⁵ Departement of Marketing, HEC Montréal, 3000 Côte Sainte-Catherine, H3T 2A7, Montréal, QC, Canada

1 Introduction

Physiological measures are increasingly used in many different areas of human-computer interaction (HCI) to infer knowledge about the affective and cognitive states of users. For example, they have been used in video game studies to measure boredom [28] and game experience [47]. Intelligent tutoring systems are now taking advantage of the dynamic nature of physiological signals to improve the adaptation of pedagogical interventions to user needs during learning sessions [45]. Researchers in consumer neuroscience are also using physiological signals to better understand the influence of computerized interfaces on brand recognition [59]. In information systems research, physiological signals have been used, among other things, to improve understanding of the mental processes of business users in multitasking contexts [43]. In the field of affective computing, physiological signals are one of the main data types used to develop affect detection systems [15]. However, the addition of physiological measures to the analysis toolbox of these fields requires expert knowledge in physiological computing (e.g., artifacting, baselining, synchronization, feature extraction, etc.) [25]. Among the different challenges, one of the most important is connecting the recorded signals to the users' actions. As noted by Ganglbauer [26] and Pantic [50], the main obstacle to the use of physiological signals is their reduced informative value when they are not specifically associated with user behavior. Physiological data would be much more valuable if interpreted in terms of interaction states.

To meet this challenge, most researchers are focusing on finding ways to measure physiological signals and interaction states synchronously. Kivikangas et al. [38] have developed a triangulation system to interpret physiological data from video game events. Conati [16] uses physiological affect detection along with task tracing to provide an explicit representation of the cause of affective user states in a virtual learning environment. Bailey and Iqbal [2] have developed task recognition and physiological methods to study cognitive interruption costs for intelligent user notification systems. Dufresne et al. [24] have proposed an integrated approach to eyetracking-based task recognition as well as physiological measures in the context of user experience research. While these works have produced interesting results, they are not easily transferable to new contexts of use because they are based on non-generic internal information from the interactive system (e.g., video game logs, application events, or areas of interest).

In this paper, we propose a new generic method for the visual representation and interpretation of physiological measures in HCI: physiological heatmaps. The tool implementing our approach for physiological heatmaps has been presented in [27]. This paper presents the full technical details and the underlying physiological computing challenges. As illustrated in Fig. 1., gaze heatmaps are used in eyetracking research and in industry as intuitive representations of aggregated gaze data [48]. Their main use is to help researchers and HCI experts answer the question: "Where in the interface do people tend to look?" [67]. Using the same rationale, physiological heatmaps make it possible to map users' physiological signals—and physiologically inferred emotional and cognitive states—onto the interface that they are using. Physiological heatmaps can then help answer the question: "Where in the interface do people tend to feel something?" Therefore, they can provide useful information in the combined interpretation of users' physiological signals and behavior, while they interact with an interface or a system.

The paper is organized as follows. Section 2 describes the standard method used to create gaze heatmaps and outlines the different parameters affecting how they appear. Section 3 details how gaze heatmaps are extended and modified to create physiological heatmaps. Section 4 describes the experiment that was performed to validate the proposed method. The results are presented in section 5 and discussed in section 6. Concluding remarks are given in section 7.

2 Gaze heatmaps

Heatmaps were originally introduced by Pomplun et al. [55] and now represent one of the most popular ways of visualizing aggregated eyetracking data [48]. At the simplest level, heatmaps are a visual representation of user-gaze distribution over an interface (see Fig. 1).

Although heatmaps seem relatively intuitive and straightforward, Bojko [6] indicates several issues that must be cautiously considered to ensure the validity of their interpretation. Heatmaps creation involves a number of arbitrary parameters for which selected values can have a significant influence on their final rendering. This section describes the three main steps for generating a gaze heatmap: accumulation, normalization, and colorization and outlines the related parameters.

2.1 Preliminary definitions

As illustrated in Fig. 2, the human field of view (FOV) is represented by an ellipsoid with a major horizontal axis of 180° and a minor vertical axis of 130° [22]. However, the image captured by the FOV is not perceived with a homogenous acuity. The FOV consists of three concentric areas of decreasing acuity: the foveal area ($0\text{--}2^\circ$), the parafoveal area ($2\text{--}5^\circ$), and the peripheral area ($>5^\circ$). Most photoreceptor cells are clustered in the fovea, at the center of the retina.

Therefore, sharp objects such as words or image details, can only be processed with full acuity in the foveal area. Images become gradually blurrier as we move from the fovea into the peripheral area. For example, acuity at 5° is of 50%, and the “useful” visual field extends to about 30° [22]. The essential goal of eye movements is then to place visual information of

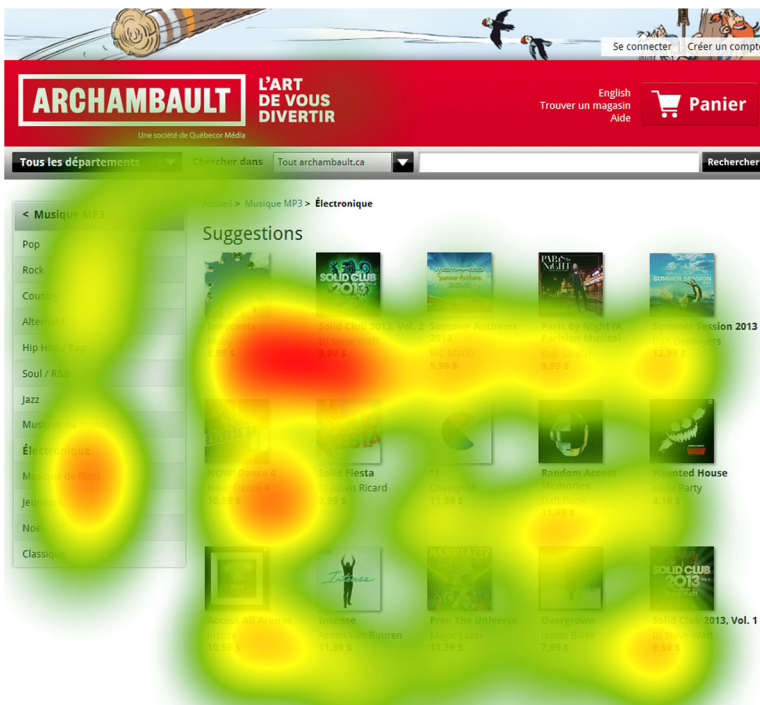


Fig. 1 Gaze heatmap. Red regions indicate a higher frequency of users' gazes

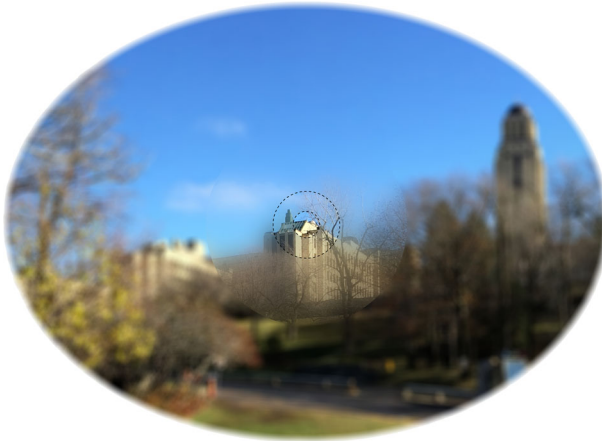


Fig. 2 Field of view. From center to end: foveal area, parafoveal area, and peripheral area

interest onto the fovea. To do so, the two main types of eye movements are fixations and saccades. The role of fixations is to stabilize the retina over a stationary object of interest to allow information encoding [22]. Representing 90% of the viewing time [32], the average duration of fixations is between 150 ms and 600 ms. Saccades are rapid eye movements that occur between fixations, in order to reposition the retina on a new location [22]. Fixations and saccades are detected in the raw eyetracking data using different types of algorithms based on dispersion, velocity, and/or acceleration information [30].

The first thing to consider when creating heatmaps is whether to use raw gaze data or pre-processed fixations. Most eyetracking analysis programs use fixations because they are less likely to include noisy data and they reduce the amount of data to process. Bojko [6] illustrates the effects of using raw data and advocates the use of fixations. In the present study, eye fixations were used as the basic data for creating heatmaps.

2.2 Accumulation

The first step in heatmap rendering is the creation of a blank map with the same dimensions as the image stimulus ($n \times m$ pixels). Then f is defined as a fixation at a location (x, y) and p as a pixel at a location (i, j) on the blank map. For each fixation (f), all the pixels (p) are attributed an intensity level through a scaling function (s) and a weight function (w), as modelled by this equation:

$$\text{Intensity}(p) = s(f, p) * w(f) \quad (1)$$

The scaling function, sometimes referred to as a point spread function (PSF), represents the probability that pixel p will be perceived by the user during fixation f . Different types of scaling functions can be used:

- No scaling: $s = 1$
- Linear scaling: $s = \frac{1}{\sqrt{(x-i)^2 + (y-j)^2}}$
- Gaussian scaling: $s = e^{-\frac{((x-i)^2 + (y-j)^2)}{2\sigma^2}}$

Most eyetracking analysis programs use a Gaussian scaling function [30]. A low variance (σ^2) will create many narrow hot spots surrounded by cold areas on the rendered heatmap. Inversely, a high variance will induce a uniformly colored heatmap with a few sparse hot spots. Wooding [67] suggests using the size of the fovea projected onto the image stimulus to determine the variance parameter. In this research, we used Wooding's approach following the implementation described by Blignault [5]. The σ constant is expressed in terms of the full width of the distribution at half maximum (FWHM):

$$\text{FWHM} = 2.3548 \times \sigma$$

Blignault [5] suggests defining FWHM as to represent 40% of the maximum parafoveal visual span:

$$\sigma = 0.17 \times 5^\circ \text{ visual span}$$

Although using the visual span represents a more biologically valid way to determine the variance parameter, the size of the visual span remains arbitrary. As noted by Wooding [68], the true width of the Gaussian distribution should depend on the area over which a fixation can be said to exist and is therefore task dependent.

In gaze heatmaps, the weight function (1) represents the eyetracking metric used to compute intensity. The four main metrics are fixation count, absolute fixation duration, relative fixation duration, and participant percentage (see Bojko [6] for a discussion of the pros and the cons of each metric). The heatmaps presented in this paper were created using absolute fixation duration. Intensity values falling on the same pixel are then summed to produce a height map (Fig. 3a).

On a practical level, most heatmap implementations do not calculate the accumulated intensities for all the pixels of an image stimulus for each fixation. Such a standard approach would lead to an algorithm of complexity $O(mn^2)$ for an image of $n \times n$ pixels and m fixations. The two main alternatives are to truncate the Gaussian kernel [51] (e.g., beyond 2σ) or to use a cone or a cylinder approximation [67]. Duchowski et al. [23] propose a GPU-based algorithm that makes it possible to extend the Gaussian kernel to image borders for smooth online heatmap rendering. In this research, the Gaussian kernel for gaze heatmap was truncated at a distance representing twice the parafoveal visual span (10°). This choice was based on the assumption that pixels beyond this region are less likely to be perceived by users during a fixation.

2.3 Normalization

In order to outline the most prominent parts of an interface, the height map is not rendered in its entirety. The second step in heatmap construction is therefore to choose what proportion of the height map will be used in the final visualization. A parameter $t \in [0, 1]$ is defined to

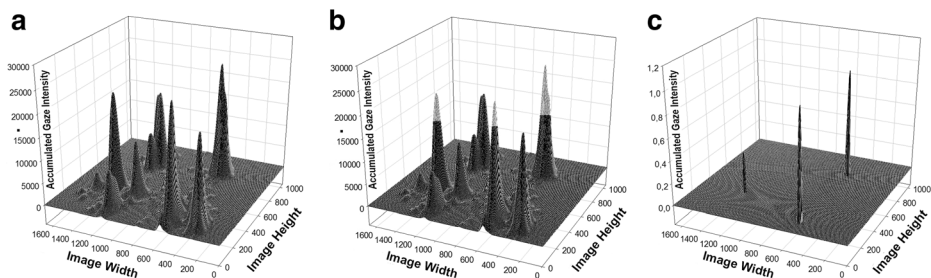


Fig. 3 Height maps: **a** raw accumulation, **b** threshold, and **c** normalization

compute the threshold under which intensity values are neglected. For example, $t = 0.2$ implies that only pixels with an accumulated intensity superior to 20% of the maximum intensity will be considered. The value of the t threshold can be seen as representing the water level in a flooded valley. As illustrated in Fig. 3c, only the non-immersed parts of the height map (over the threshold level) are used in the final rendering. In most eyetracking programs, the threshold can be modified in order to capture the range of values that are of interest to a particular study. Too high a threshold will result in a heatmap with few colored regions, and too low a threshold will result in a fully covered heatmap with poorly differentiated areas.

The height of the highest peak largely depends on the number of subjects (i.e., more subjects imply more fixations). Therefore, in order to compare heatmaps, it is necessary to normalize the height map before the colorization step. The most standard approach is to give the maximum accumulated intensity the value of 1. In this research, we implemented normalization for each pixel at location (i, j) using a min-max equation:

$$\text{Normalized } (p_{ij}) = \frac{I_{ij} - I_{\min}}{I_{\max} - I_{\min}} \quad (2)$$

In this equation, I_{\max} stands for the maximum accumulated intensity value and I_{\min} stands for the minimum intensity higher than the t threshold. As illustrated in Fig. 3b, the non-immersed parts of the height map are rescaled in the range of $[0, 1]$.

2.4 Colorization

The last step in creating a heatmap is colorization. The main idea is to change the stimulus image to reflect the relative variations in the height map. Colorization can be applied directly on the image stimulus or on a semi-transparent superimposed layer. The former solution has the disadvantage that the resulting visualization at a specific pixel depends both on the initial color of the pixel and on the color mapping [68].

Height variations can then be mapped to different color properties resulting in various types of visualizations. The most commonly used visualizations are heatmaps, luminance maps, and contrast maps [30]. For instance, heatmaps are created using a rainbow gradient linearly interpolating between blue-green, green-yellow, yellow-red, and red. Although the rainbow gradient is currently the most popular visualization, it has many flaws [8]. For example, it has no perceptual ordering and it obscures small details by only showing apparent changes at color boundaries. A good alternative is to map accumulated intensity onto the brightness values for a single hue gradient. One should choose a mapping that is consistent with the visualization objective. Breslow et al. [12] have demonstrated that multicolored scales (e.g., rainbow) are best suited for identification tasks (i.e., determining absolute values using a legend) and brightness scales (single hue) are best suited to compare relative values. In this research, colorization was applied on a semi-transparent layer using a brightness gradient. The implementation uses an RGB interpolation based on the color ramping tables provided in [10].

3 Physiological heatmaps

As stated, the main goal for creating physiological heatmaps is to answer the question: "Where in the image do people tend to feel something?" In this new visualization method, the users' gaze now serves as a means of transferring data in order to map

physiological responses onto the image stimulus. Following proper data manipulation, the color variations now represent the physiological signals' distribution over the interface. More precisely, it represents the relative intensity of the signals at the time users were looking at different parts of the interface. This section describes how the three main steps of gaze heatmap construction (accumulation, normalization, and colorization) are adapted, along with a prior synchronization step, to create physiological heatmaps.

3.1 Preliminary definitions

The purpose of creating physiological heatmaps is to allow for a more informative and contextualized visualization of physiologically inferred emotional and cognitive states of users. The physiological signals used in heatmap rendering should therefore be selected in a way that optimally represents the psychological construct of interest (e.g., emotion, cognitive load, attention, flow, etc.). We frame this problem using the concept of psychophysiological inference. Cacioppo et al. [13] describe psychophysiological inference using the following equation:

$$\Psi = f(\Phi) \quad (3)$$

where ψ is the set of psychological constructs and Φ is the set of physiological variables. The f relationship can then be declined in four ways: 1) one-to-one: a psychological state linked in an isomorphic manner to a physiological variable, 2) one-to-many: a psychological state reflecting various physiological variables, 3) many-to-one: various psychological states related to a single physiological variable, or 4) many-to-many: multiple psychological states linked to multiple physiological variables. Most research using physiological signals to infer a psychological construct are based either on the first or the third relationships. For the sake of clarity and concision, the creation of physiological heatmaps will be described using the one-to-one relationship. Section 6.1 explains how the proposed method can be easily used to map emotions or other psychological constructs resulting from a more robust inference process (e.g., one-to-many).

3.2 Synchronization

As depicted in Fig. 4a, eyetracking and physiological data are recorded while a user is interacting with an interface. However, it is not possible to start all the devices at the exact same time, and as a result, recordings are asynchronous and each data stream has its own specific time frame.

There are various methods to synchronize concurrent physiological and behavioral recordings, but there are two main approaches: direct synchronization and indirect synchronization. In the first, eyetracking markers are sent in real time to each piece of physiological recording equipment (e.g., to an electrocardiogram or an electroencephalogram). This method requires coding in order to adapt the devices' API and it can be laborious as the number of concurrent recording devices rises. Moreover, it can make it impossible to change the eyetracking parameters afterwards (e.g., the fixation detection algorithm). In indirect synchronization, all the devices are started independently and data files are exported and reconciled after recording (Fig. 4b). The latter approach was adopted in this research as it leaves more freedom for data manipulation (as required in the steps described in the next subsections). The specific implementation is described in section 4.4.

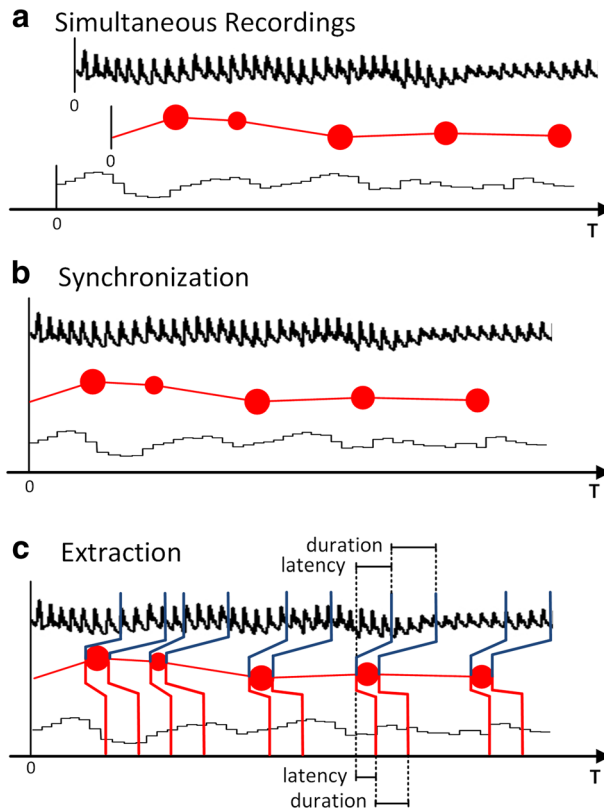


Fig. 4 Data processing. This figure has been created for the sake of illustration. The upper signal represents blood volume pressure (BVP) and the lower one represents pupil size. a) simultaneous recordings of eyetracking data and two physiological signals b) recordings synchronisation c) multiresolution extraction windows optimized for each physiological signal

3.3 Accumulation

In a physiological heatmap the intensity eq. (1) now represents the intensity of a user's physiological signals during a fixation. Therefore, the weight function needs to be modified accordingly. To do so, a specific segment of the user's physiological signal is mapped to each of her/his fixations on the image stimulus. We call this process the extraction step. It is implemented using a feature-based approach initially developed in the fields of affect detection and machine learning [53]. This method consists of computing different statistical features of the signal (e.g., mean, standard deviation, or max) over a certain time period. These features are then used as inputs in machine learning algorithms to predict a psychological construct (e.g., emotion, attention, or cognitive load). In our context, statistical features are used to compute the weight function and are calculated over segments of physiological signals, starting at each user's fixation onset. In this specific research, the physiological heatmap weight function has been implemented to represent the mean of the signal using the following equation:

$$W(f) = \frac{\sum_{i=s}^e p_i}{e-s} \quad (4)$$

where P_i stands for a physiological data point at time i , and s and e are respectively the start and the end of the segment associated with fixation f . Weight functions representing other features (e.g., min, max, or standard deviation) can be defined in a similar manner.

However, a segment's start and end points cannot be selected simply by using the onset and the duration of the related gaze. Because emotion and cognition require physiological adjustments stemming from multiple response patterns, different physiological signals present various durations and latencies for a given stimulus [40]. In the field of physiological computing, this phenomenon is known as signal asynchronicity [65]. For example, in response to a stimulus, heart rate may change more rapidly than electrodermal activity, but more slowly than pupil size. In our context (Fig. 4c), *latency* is defined as the time elapsed between a fixation onset and the beginning of a related physiological reaction (s in (4)). *Duration* is defined as the time elapsed between the start and the end of the physiological reaction ($e - s$ in (4)). It is therefore necessary to use specific extraction windows that are optimized in terms of latency and duration for each physiological signal and for the psychological construct of interest. Otherwise, the weight function would not capture the signal variation due to the processes occurring during a specific fixation. The windows used in this research were empirically optimized following a multiresolution feature extraction approach described in previous research [17]. As for gaze heatmaps, the scaling function's Gaussian kernel (1) is truncated in order to speed up computations. However, as emotional information is highly salient, we set the truncation threshold of physiological heatmaps at three times the parafoveal distance (15°). Coy and Hutton [18] report effects of extrafoveal emotional information on saccadic movements at distances up to 12° of fixation's center.

3.4 Normalization

As physiological signals are subject to significant interpersonal variations, absolute values cannot be used to compare data from multiple subjects. Moreover, physiological signals from the same person may vary according to the context (e.g., because of the position of sensors) and the time of day [53]. Therefore, physiological signals need to be corrected to account for the subject's baseline [65]. To do so, the result of the weight function (4) for each fixation is corrected using the following z-score equation:

$$W'_i = \frac{W_i - \mu}{\sigma} \quad (5)$$

where μ and σ are respectively the mean and the standard deviation of the weight function's result (4) for all of a subject's fixations. The z-score normalization was chosen for two main reasons. First, it makes it possible to account for user differences in terms of baseline physiological levels (μ) and physiological sensitivity (σ). The weight function's outputs of different users can then be summed up in a congruent way. Second, it makes it possible to obtain a higher contrast between the "physiologically significant" areas of an interface and the neutral ones. In gaze heatmaps, every fixation has a positive intensity value and increases the height map. In accordance with the new weight function (5), a positive W' represents a value above the mean, and a negative W' , a value below the mean. Physiologically unimportant fixations ($W' < 0$) are not considered in the accumulation step; only important physiological activity makes a contribution. Therefore, the resulting height map has a topology that is different from that of a gaze heatmap. The accumulated height map is also normalized using the same threshold function (2) as described in section 2.3 to allow for comparison of different physiological heatmaps.

3.5 Colorization

The colorization of physiological heatmaps is based on the same principle as the colorization of standard gaze heatmaps (see section 2.4). However, in the former, more than one psychological construct can be mapped at the same time onto an image stimulus for analysis purposes. For example, it can be used to compare regions of high cognitive load versus regions of negative emotional valence. Therefore, it is not possible to use a unique rainbow gradient, as mappings of different constructs would be indiscernible from one another. Using rainbow gradients of different colors would be confusing, when comparing multiple maps at the same time, because they are not perceptually ordered [8].

On the other hand, an increase in luminance is a stronger perceptual cue and is consistent across hues. We can see in Fig. 5 that the difference between segments A and B is easy to grasp using only the *Rainbow 2* gradient: the redder the hue, the more intense.

However, it is much more difficult to visually determine whether this difference is the same as the one obtained with *Rainbow 1*, as it requires comparing gradations of blue with gradations of red. On the contrary, it is easier to compare segments A and B using *Luminance 1* and *Luminance 2*. The differences are consistent across both representations as they are based on the same luminance gradient, even though they are not depicted by the same color (blue and orange). We therefore suggest representing psychological constructs with luminance gradients of different colors (see Fig. 14 for an example). Furthermore, we recommend using complementary colors (i.e., opposed colors on the color wheel) [44].

4 Experimental validation

This section presents the experimental validation that was conducted to evaluate the proposed physiological heatmap method. The main goal was to assess the ability of physiological heatmaps to identify the most emotionally engaging parts of different interfaces. Section 4.1 describes the signals that were used to create physiological heatmaps. Sections 4.2 and 4.3 respectively present the apparatus and image stimuli used in the experiment. Section 4.4 describes the synchronization method used to coregister eyetracking and physiological data.

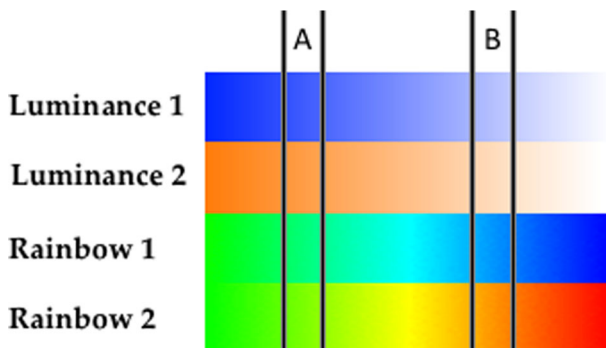


Fig. 5 Luminance vs rainbow gradients. Luminance 1 and 2 are varying only in luminance from a range of 0 to 255 with a fixed hue (blue and orange). Rainbow 1 and 2 are varying in hue, starting from the same green to blue and red

4.1 Physiological signals

The Circumplex model of affect [56] describes emotions using the two dimensions of valence and arousal. Valence is used to contrast states of pleasure (e.g. happiness) and displeasure (e.g. anger), and arousal to contrast states of low arousal (e.g. calm) and high arousal (e.g. excited). Based on this model, two types of physiological heatmaps were produced to visualize users' emotions: arousal and valence heatmaps. Arousal heatmaps were created using three physiological signals: electrodermal activity (EDA), pupil diameter (PD) and electroencephalography (EEG). Valence heatmaps were created using the FaceReader 5 (Noldus, Netherland) facial expressions analysis software [69]. As will be discussed in section 6.1, three signals were used in order to assess the impact of the choice of physiological signals on the final heatmap rendering. A third party software was also used in order to illustrate the versatility of the approach.

Electrodermal activity (EDA) measures electrical conductance changes of the skin near eccrine glands (i.e. sweat glands). Although the main function of eccrine glands is thermo-regulation, it has been shown that EDA measured at the plantar and palmar locations are also sensitive to psychological stimuli [9]. More precisely, it is known to be highly correlated with a person's arousal level [20].

The main function of pupil dilation is to dynamically adapt to changes in ambient illumination (pupillary light reflex) [3]. However, research showed that variations in pupil diameter (PD) also respond significantly to cognitive and emotional stimuli [37, 41, 62]. Recent data suggests that the pupil's response is significantly modulated by emotional arousal regardless of hedonic valence [1, 11].

Electroencephalography (EEG) measures the brain's electrical activity at the scalp level using small electrodes [54]. The measured activity is generated by the simultaneous activation of millions of neurons in different parts of the brain. Many psychological states and reactions can be inferred from the temporal and frequential analysis of the EEG signal. Of particular interest to this work, it has been shown that frontal activation is related to the arousal dimension of emotion [58]. As evidence suggests that alpha power (8-13 Hz) is inversely proportional to underlying cortical processing, decrease of alpha power is usually used as a measure of activation [19].

4.2 Participants and apparatus

Fifty participants (29 female, 21 male) were recruited for this experiment over a period of three weeks. Data from six participants was rejected due to equipment malfunction. Therefore, data from 44 participants was used in the analyses. All participants had normal or corrected-to-normal vision and were pre-screened for glasses, laser eye surgery, astigmatism, epilepsy, and neurological and psychiatric diagnoses. Most participants were either undergraduate or graduate students from *HEC Montréal*. Informed consent was obtained from each participant and a 20\$ gift certificate compensation was given upon experiment completion. Stimuli were presented on a 22' (508 mm × 356 mm) LG LED monitor with a resolution of 1680 × 1050 pixels and a refreshing rate of 60 Hz. Average distance between participants' eyes and the monitor was of 660 mm.

A Tobii X-60 (Tobii Technology AB) eyetracker was used to record subjects' eye movement and pupil diameter using a sampling rate of 60 Hz. A nine points calibration was performed for all participants and was repeated until sufficient accuracy was achieved. As indicated by [21], current video-based eye trackers have a spatial resolution of up to 0.01°/2 kHz. However, such a high-resolution level is experimentally hard to achieve due to subjects' variability. In this research,

sufficient accuracy was defined as ~ 1 cm (0.868° of visual angle) around the center of the calibration points. The Tobii implementation of the I-VT fixation filter algorithm [57] was used to extract fixations from the eyetracking data (minimum fixation duration = 60 ms). The following parameters were used for fixation merging: maximum angle between fixations was set to 0.5 degrees, and maximum time between fixations was set to 75 ms.

EEG activity was measured with a 32-electrode array geodesic sensor net, using the Netstation acquisition software and EGI amplifiers (Electrical Geodesics, Inc). The vertex (recording site Cz) was the reference electrode for recording and the common average reference was calculated and applied later on. Impedance was kept below $50\text{ k}\Omega$ and a sampling rate of 256 Hz was used to record the data. An Independent Component Analysis (ICA) was applied to attenuate the movement of eye blinks and saccades in the EEG data [36]. The ICA used a selection of 200 s of training data located at 1000s into the recording. An automatic artifact rejection was used to exclude epochs with voltage differences over $50\text{ }\mu\text{V}$ between two neighboring sampling points and a difference of over $200\text{ }\mu\text{V}$ in a 200 ms interval. These steps were performed using the Brain Vision Analyzer 2 software (Brain Products).

Electrodermal activity was recorded with a wireless MP-150 Biopac amplifier (Biopac MP) using two electrodes placed on the palm of the non-dominant hand.

Videos of the participants' faces were recorded using the Media Recorder 2 software (Noldus, Netherland) at a resolution of 800×600 and a frequency of 30 frames per second. Videos were processed in FaceReader 5 (Noldus, Netherland) a posteriori to produce valence inference at a frequency of 30 inferences per second.

4.3 Image stimuli

The image stimuli presented to participants were composed of a set of standardized pictures from the International Affective Picture System (IAPS) [42] displayed on a single image with a gray background (see Fig. 6).

The images were created following the next four experimental conditions. The main idea behind conditions 1, 2, and 3 was to elicit different level of arousal or valence level while keeping the other dimension fixed.

- Condition 1: images from condition 1 were designed to elicit negative emotions varying along the arousal dimension. Each image contained three pictures amongst which one was

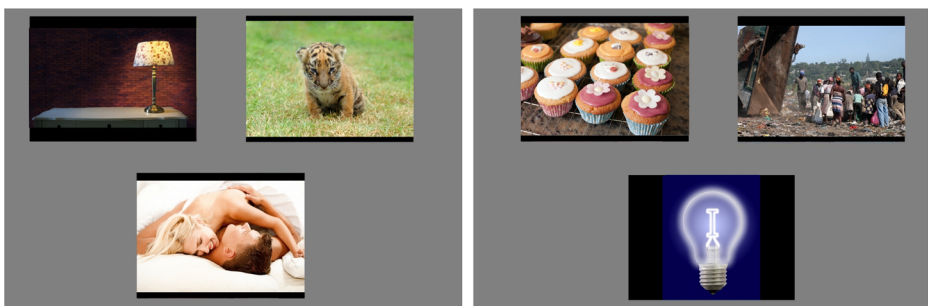


Fig. 6 Left: image stimulus from condition 2 composed of one neutral picture (top-left) and two non-neutral pictures of positive valence with medium arousal (top-right) and high arousal (bottom). Right: image stimulus from condition 3 composed of one neutral image (bottom) and two non-neutral pictures of medium arousal and positive (top-left) and negative valence (top-right)

- always neutral (low arousal with neutral valence). The two non-neutral IAPS pictures used for each image had negative valence (depicted in blue in Fig. 7). Three image stimuli (nine IAPS pictures) were created.
- Condition 2: images from condition 2 were designed to elicit positive emotions varying along the arousal dimension. Each image contained three pictures amongst which one was always neutral (low arousal with neutral valence). The two non-neutral IAPS pictures used for each image had positive valence (depicted in red in Fig. 7). Three image stimuli (nine IAPS pictures) were created.
 - Condition 3: images from condition 3 were designed to elicit emotions varying along the valence dimension with a fixed arousal level (medium). Each image contained three pictures amongst which one was always neutral (low arousal with neutral valence). The non-neutral IAPS pictures were of negative and positive valence (depicted in green in Fig. 7). Six image stimuli (18 IAPS pictures) were created.
 - Condition 4: images from condition 4 were designed to elicit many contrasted emotional reactions. Two images contained four pictures located in different quadrants of the Circumplex model of affect. The first image from condition 4 contains 4 pictures with medium arousal and medium valence contrast. The second image from condition 4 contains 4 pictures with high arousal and high valence contrast.

The presentation sequence is illustrated in Fig. 8. A vanilla baseline method [35] was employed during the 60 s rest periods in between condition blocks in which randomly colored squares were presented for 6 s while participants had to passively watch and count the number of white squares.

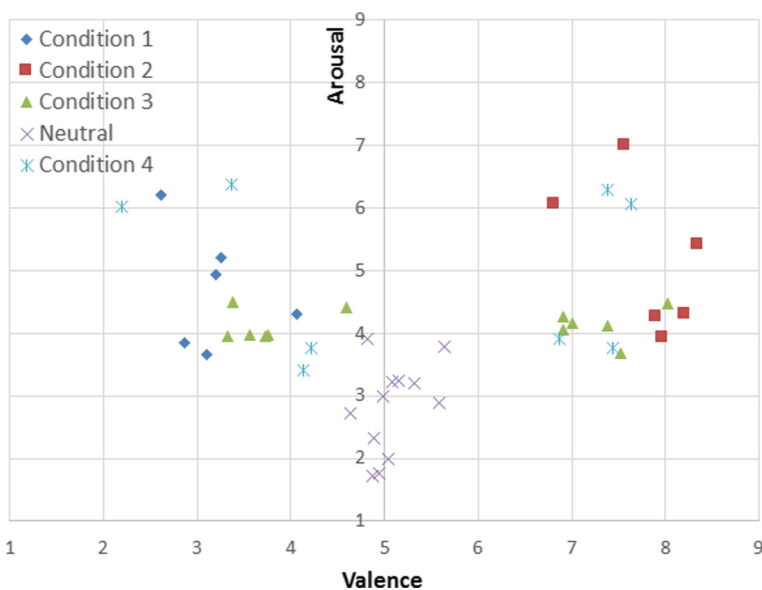


Fig. 7 Non-neutral pictures from conditions 1 and 2 vary along the arousal dimension with either a negative or positive valence. Non-neutral pictures from condition 3 vary along the valence dimension with a medium arousal. Neutral pictures have a low arousal and a neutral valence. The first image from condition 4 contains 4 pictures with medium arousal and medium valence contrast. The second image from condition 4 contains 4 pictures with high arousal and high valence contrast

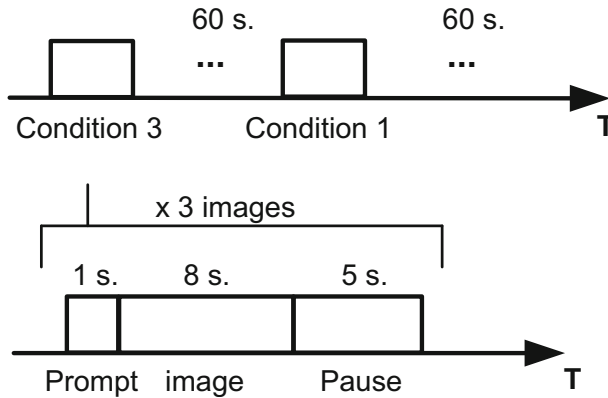


Fig. 8 Presentation sequence. Three images were presented within each condition block. After the presentation of a visual prompt (green cross at the center of the screen), each image was shown for a duration of 8 s followed by a 5 s pause. Blocks were separated by a 60 s rest period to ensure that physiological signals returned to baseline level

A total of 14 images were presented to participants (6 images in condition 3, 3 images in condition 1, 3 images in condition 2, and 2 images in condition 4). There were 2 blocks assigned to condition 3 and 1 block to conditions 1, 2 and 4. The presentation order of condition 1, 2, and 3 was randomized. As images from condition 4 were purposefully design to elicit strong emotional reactions, it was always the last block to be presented. This was done in order to avoid emotional contamination with the other stimuli. Data from conditions 1, 2, and 3 was analyzed together and data from condition 4 was analyzed separately. A practice image containing three pictures (neutral, and high arousal with positive and negative valence) was presented before the beginning of the experiment. As the IAPS collection contains outdated pictures, a pre-screening based on interjudge agreement was conducted by eight research assistants in order to avoid presenting pictures which no longer had the same emotional interpretation or contained anachronisms. Otherwise, prior experience with IAPS has shown that some pictures may induce laughter instead of the intended emotion. That explains why some neutral pictures that were selected are slightly higher in the IAPS arousal scale (see Fig. 7).

The IAPS pictures used in the image stimuli were: condition 1: (1111, 9331, 7018), (5390, 2752, 2457), (2683, 9001, 7001), condition 2: (1710, 1441, 7025), (5621, 1460, 7175), (4676, 1920, 7014), condition 3: (4700, 9341, 7236), (2154, 9110, 7004), (1740, 1945, 7006), (2035, 9090, 7010), (2341, 6010, 7012), (1410, 2490, 6150), condition 4: (1304, 7405, 2491, 1601), (6212, 8499, 2280, 2388), and practice (7009, 9183, 4697). Correlations were calculated between pictures' luminance and emotional ratings in order to account for possible confounds (especially with regards to pupil size measures). No significant correlation was found between luminance and arousal for pictures in conditions 1, 2, and 4 and between luminance and valence in conditions 3 and 4.

4.4 Synchronization

Indirect synchronization was implemented using a third party device (Syncbox from Noldus, Netherland) to send TTL (Transistor-Transistor Logic) signals every 10 s to each recording device. The data streams were then synchronized by aligning their TTL

markers. As equipment manufacturers strongly recommend using only one computer per measurement tool to guaranty their specified precision level, multiple computers were employed. Therefore, another synchronization step is required to account for clock drifting. The elapsed time between two subsequent TTL signals recorded by a device is slowly drifting (forward or backward) compared to the expected 10,000 ms interval. The discrepancy between two recording devices can rise up to 100 ms over a time period of 1 h. A linear regression method [29] was therefore used to correct clock drifting effects in the different data streams, and ensure that synchronization remained throughout the entire time course of a recording session.

5 Results

As illustrated in Fig. 9, the data used in the analyses was obtained by comparing the volume of the sections of the height maps that are above the normalization threshold (light gray) with the standardized arousal and valence ratings of the underlying pictures. The standardized values of valence and arousal used were the average ratings for men and women provided by the IAPS manual.

A data point used for analysis is then composed of two numbers: height map volume and IAPS rating. One height map was generated for each image stimulus, per subject and per measure (i.e., arousal_EDA, arousal_Pupil, arousal_EEG, positive valence, negative valence, and gaze). Each type of height map was evaluated separately. In this context, for a given signal and a given construct, a participant would generate four data points for a stimulus composed of four IAPS images.

Data was analyzed following two angles. First, we wanted to compare the relative performance of gaze and physiological heatmaps. We measure performance using the Pearson correlation

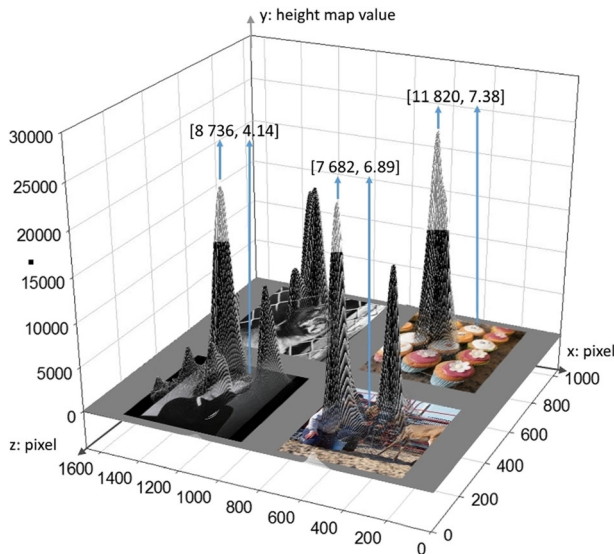


Fig. 9 Data example of positive valence stimuli. The top right picture has a standardized valence of 7.38 and its associated height map volume is 11,820. Four data points are generated from this image stimulus for a given subject: [0, 3.37], [11,820, 7.38], [8736, 4.14], and [7682, 6.89]. The top left image has a value of 0 since the portion of the height over it is below the threshold and wouldn't be displayed in the rendered physiological heatmap

coefficient between the two numbers of a sample of data points. Second, we wanted to compare the performance of physiological heatmaps applied to stimuli having a regular “emotional density” with stimuli having a high “emotional density”. We define emotional density as the number of contrasted and emotionally significant regions in an image. For example, stimuli from conditions 1, 2 and 3 contain two non-neutral pictures and have a regular density. The last two stimuli from condition 4 have a higher emotional density as they contain four non-neutral pictures.

5.1 Performance

The experimental design expected 792 data points from the stimuli varying along the arousal dimension ($44 \text{ participants} \times 18 \text{ pictures in conditions 1 and 2}$). However, 778 data points were obtained for gaze, EDA, and pupil heatmaps and 772 data points were obtained for EEG heatmaps. Missing data are due to artefacts in the physiological signals during a fixation or to the fact that some participants did not look at all the pictures in the presented image stimuli. The correlations were all significantly different than 0 (all p -values < 0.001). The p -values were corrected by assessing the effect of each measure on the arousal values by using linear regression model accounting for the potential correlation between each repeated measure coming from the same subject. Fig. 10 presents the correlation coefficients of arousal pictures.

The experimental design expected 528 data points for positive valence stimuli ($44 \text{ participants} \times 12 \text{ positive pictures in condition 3}$) and 528 data points for negative valence stimuli ($44 \text{ participants} \times 12 \text{ negative pictures in condition 3}$). Both positive and negative picture sets included the neutral pictures of a stimulus image. 513 data points were obtained for positive valence and 512 for negative valence. Fig. 11 presents the correlation coefficients obtained for valence stimuli.

5.2 Scalability

The two images of condition 4 containing four pictures were used to test the scalability of the method. In the context of physiological heatmaps, we define scalability as the ability of the

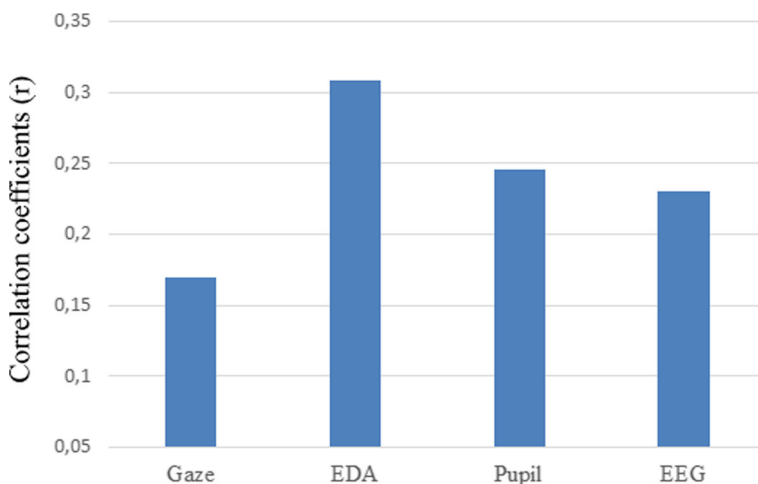


Fig. 10 Arousal results. Pearson correlation coefficients (r) between arousal ratings and heigh map volume are as follows: gaze = 0.170, EDA = 0.308, pupil = 0.246, and EEG = 0.230. All p -values are below 0.001

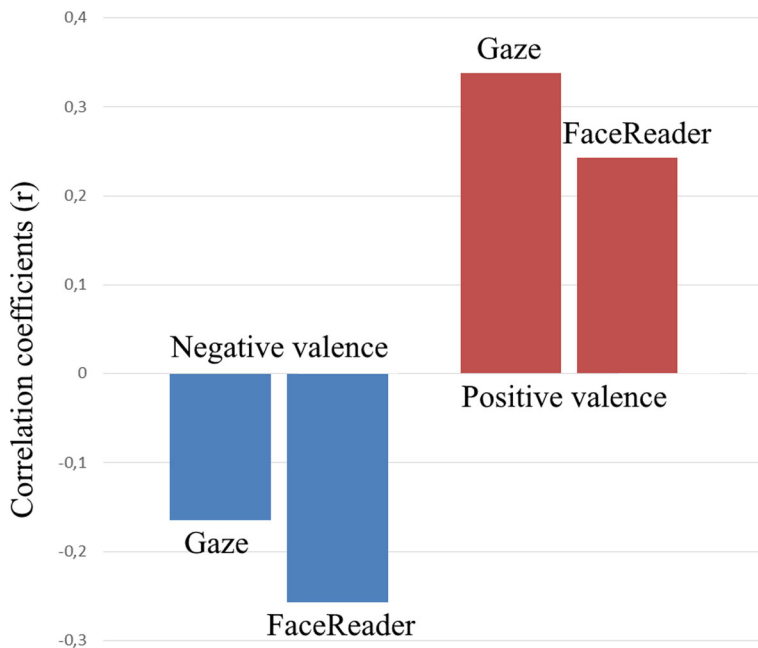


Fig. 11 Valence results. Correlation coefficients (r) between valence ratings and heigh map volume are as follows: Negative valence: gaze = -0.165 and FaceReader = -0.257 . Positive valence: Gaze = 0.338 and FaceReader = 0.243 . All p -values are below 0.001

method to maintain performance while the emotional density of stimuli increases. The experimental design expected 352 data points for arousal and valence stimuli from these two images ($44 \text{ participants} \times 4 \text{ pictures}$). After artefact rejection, 331 data points were obtained for gaze heatmaps, 331 for EDA heatmaps, 331 for pupil heatmaps, and 329 for EEG heatmaps. Table 1 presents the correlations results for arousal heatmaps.

Arousal heatmaps based on electrodermal activity and pupil size were significantly correlated with arousal ratings. Arousal heatmaps based on gaze and EEG were not significantly correlated with arousal ratings. No correlation remained significant for valence heatmaps.

6 Discussion

The proposed physiological heatmap method is designed as a new tool to help HCI and UX practitioners answer questions such as: "Where in the image do people tend to feel something?" or "Where in the interface do people tend to experience more cognitive load?" The experimental results indicate that the method can answer the first question better than standard gaze heatmaps. The results

Table 1 Correlation results for arousal heatmaps

	r	p -value
Gaze	-0.065	0.255
EDA	0.151	<0.01
Pupil	0.382	<0.001
EEG	0.101	0.088

are discussed according to inference validity (section 6.1) and spatial resolution (6.2). Section 6.3 presents an application of the method as well as a real case scenario: emotional salience maps.

6.1 Inference validity

Physiological heatmaps are designed to visualize users' physiological signals according to higher level affective and cognitive states. As such, the accuracy of the method depends on the strength of the underlying psychophysiological inference (see section 4.1). For example, a cognitive load heatmap based on pupil size and heart rate would be more accurate than a heatmap based on respiration rate, as the two former signals are known to be correlated more closely with cognitive load. Results presented in section 5.1 show that the accuracy of the arousal heatmaps is higher when based on EDA ($r = .308$) than on pupil size (.246), which in turn is better than EEG (.230). This result is in line with the psychophysiological literature [14]. When implemented in a one-to-one fashion, the f relationship (3) better predicts arousal than pupil size or EEG. As mentioned, the current study uses only one signal at a time for psychophysiological inference to simplify the description of the framework and the full pipeline of the proposed method. However, as the regulation of emotions relies upon the sympathetic and parasympathetic activity of the autonomic nervous system, it requires physiological adjustments stemming from multiple response patterns [40]. Hence, a one-to-many f relationship would produce a more precise and robust psychophysiological inference. To do so, the most efficient approach would be to first train a machine learning model to recognize the psychological construct of interest using a set of different relevant physiological signals. The model could then be used to generate an inference for each gaze using different extraction windows per signal (see section 3.3). These inferences would then serve as input for the heatmap creation process. Furthermore, as illustrated by the use of the FaceReader (Noldus) facial expression software to produce valence heatmaps, the proposed method can also be used to create physiological heatmaps of emotional or cognitive states based on commercial inference engines (e.g., Eyeworks,¹ B-Alert²). Finally, besides being related to inference validity, the choice of physiological signals is also task-dependant. Some contexts of use can tolerate a moderate level of sensor intrusiveness (e.g., laboratory experiment), as other contexts require a minimal level of intrusiveness to preserve the interaction's authenticity (e.g., marketing studies).

6.2 Spatial resolution

The choice of the signals underlying the psychophysiological inference also has an impact on the spatial resolution of the method in one of two main ways. First, as illustrated in Fig. 12, a signal with a low temporal resolution (i.e. long duration and latency) can produce overlaps between extraction windows and subsequent gazes. Although the extraction windows of gazes 1 and 2 are not overlapping (which is not possible for a given signal) gaze 2 occurs during gaze 1's window. However, as gaze 2's effect on the user's physiology has a certain latency, it would have limited interference with gaze 1's associated signal. Furthermore, when emotionally engaging elements of an interface are far apart, low temporal resolution is less critical. For example, gazes 1 and 2 are associated to the same emotional element and contribute to the same emotional state. The elements in Fig. 12's interface are at such a distance that extraction windows of gazes on the first element cannot overlap with gazes on the second element. Nonetheless, when used on an interface with

¹ <http://www.eyetracking.com/Software/Cognitive-Workload>

² <http://www.advancedbrainmonitoring.com/eeg-software/>

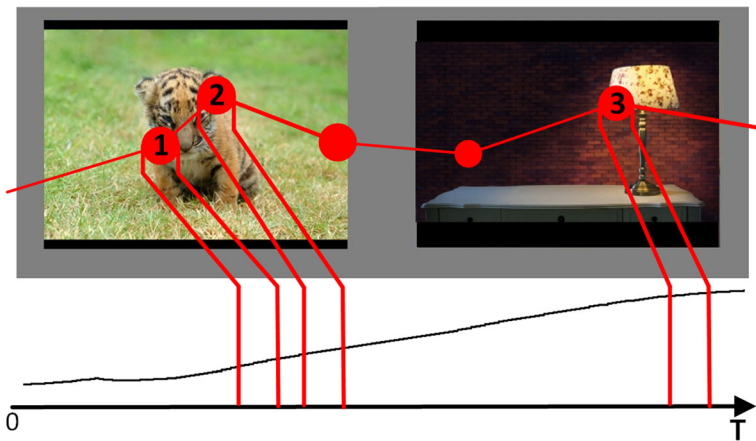


Fig. 12 Physiological signals' temporal resolution vs heatmaps spatial resolution. This figure has been created for the sake of illustration. The represented signal is electrodermal activity (EDA)

many emotionally engaging elements that are close to each other, we recommend using physiological signals that have a higher temporal resolution (e.g., EEG, pupil size).

The second effect of signals' temporal resolution is related to the law of initial values, stating that a "change of any function of an organism due to a stimulus depends, to a large degree, on the prestimulus level of that function" [66]. The use of this law in psychophysiology is subject to debate and it is instead recommended to discuss the principle of initial values [34]: a correlation between the prestimulus baseline of a function and the direction and intensity of a reaction is generally observed. For example, in Fig. 12, the signal's segment associated with gaze 3 is also affected by gazes 1 and 2. The carryover effect of low temporal resolution signals on the psychophysiological inference is discussed in more depth in [65]. We therefore suggest using signals that have a fast return to baseline. As future work, we are working on a modification of the weight function (4) that takes into account the initial value of the gazes' extraction window in order to attenuate the carry over effect.

Conditions 1, 2 and 3 of the experiment contained stimuli representing interfaces having a standard emotional complexity (two contrasted regions and a neutral), and for which the approach showed good results (both for valence and arousal). The fourth condition of the experiment was designed to test the scalability of the proposed method. The two stimuli contained 4 pictures with none being neutral. In most HCI contexts, an interface with 4 (or more) strongly contrasted emotional regions could be considered "emotionally dense". The results presented in section 5.2 show that the heatmaps based on EEG and EDA suffered from the increased spatial resolution. No significant correlation was found between EEG heatmaps and arousal ratings, and the correlation of EDA heatmaps was reduced nearly by half ($r = 0.308$ to $r = 0.151$). Gaze heatmaps were not significantly correlated with arousal ratings. However, the heatmaps based on pupil size remained significantly correlated with arousal ratings and with a stronger relationship ($r = 0.246$ to $r = 0.382$). These results suggest that using a physiological signal that has a good temporal resolution, while being adequately related with the psychological variable of interest, allows physiological heatmaps to be used with emotionally complex interfaces. In addition to the aforementioned future work on carry over effects, using a psychophysiological inference based on multiple signals should also increase the spatial resolution of the method.

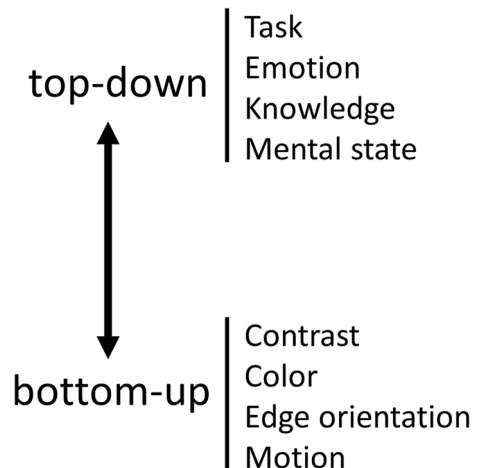
6.3 Emotional salience

Eye movements and shifts of attention in natural behavior are explained following two hypotheses (see Fig. 13). The bottom-up hypothesis explains eye movements according to low-level stimulus properties such as contrast, color, edge orientation, or motion [39, 52]. Based on a visual saliency map [33], gazes are directed to the regions with the highest saliency. Following the top-down hypothesis, attention and gazes are driven by high level factors such as emotions, mental state, semantic context, and task-related elements [63, 64]. In pure and controlled contexts, each model can account for subjects' gaze patterns relatively well. However, no unique model can achieve the same results in a rich and real world environment.

For example, visual saliency maps fully explain attention shifts only in scene-free viewing without task demands [7]. In complex social scenes, subjects first look at peoples' eyes regardless of visual saliency [4]. Using different types of activities (scene memorization, pleasantness rating, visual search, and free viewing), Mills et al. [46] found a direct task effect on both spatial and temporal characteristics of eye movements during scene perception. Research also shows that visual saliency and affective saliency are in competition for gaze control in stimuli that contain emotional information [31, 49].

For HCI practitioners, one of the main consequences of the interplay of bottom-up and top-down factors in eye movements is the ambiguity of gaze heatmaps interpretation. In a pure scene-free viewing context, a gaze heatmap would resemble the stimulus's saliency maps and could be used to identify the most salient regions of an interface. Used on a pure affective stimulus (e.g., IAPS) the gaze heatmap would look more like the affective saliency map. Affective saliency maps are made by having subjects click on the most emotionally engaging parts of an image. A description of how to create manual affective saliency maps can be found in [61]. However, in a real HCI context, where many low-level and high-level factors are present at the same time, the interpretation of gaze heatmaps is more problematic. In most cases, users have to use a visually rich interface in order to execute a task in which they experience different types of emotions and cognitive states. The resulting gaze heatmap is likely to reflect the mean effects of all these factors and cannot be used to analyze one specific dimension of the interaction without ambiguity. However, the proposed physiological heatmap method can be used to disentangle specific factors underlying eye movements on an interface. For example, when choosing a set of signals associated

Fig. 13 Saliency models



with emotions, the proposed method can be used to automatically create physiological heatmaps that will resemble the affective salience maps.

To illustrate this point, we created an arousal heatmap based on pupil size with data collected in a previous study. In this experiment, participants had to look at a hotel's page on a social networking site (SNS) and decide if they wanted to make a reservation. The page was created for the purpose of the study. There was no time limit and average viewing time was 3 min. As shown in Fig. 14 (right part), the most intense regions of the gaze heatmap are on text in the comments section.

However, the arousal heatmap indicates a clear hotspot on the number of stars rated by customers (Fig. 14 - left part). Text areas usually accumulate large amounts of intensity in gaze heatmaps, as they take more time and more gazes to process than pictorial information. On the contrary, reading the number of stars is a quick and transient action that goes unnoticed in a gaze heatmap, even though it may be the most interesting aspect for users. Therefore, in the context of this task, the arousal heatmap can help to better understand how the users' emotional experience affects the way in which they interact with the hotel's SNS.

7 Conclusion

This paper presents a novel method for visual interpretation of physiological data. The proposed physiological heatmap tool enables the representation of the relative distribution of users' physiologically inferred emotional or cognitive states on a given interface.

Physiological heatmaps' capacity to give a better picture of users' emotional reactions than gaze heatmaps was tested in an experiment involving 44 subjects. Results show that physiological heatmaps are more strongly correlated to emotionally significant regions than are standard gaze heatmaps, in terms of arousal using electrodermal activity, pupil size, or electroencephalographic data. By selecting physiological signals related to specific psychological variables (e.g., emotion, flow, cognitive load), physiological heatmaps can help HCI practitioners better understand how different aspects of user experience are related to interface and interaction design.

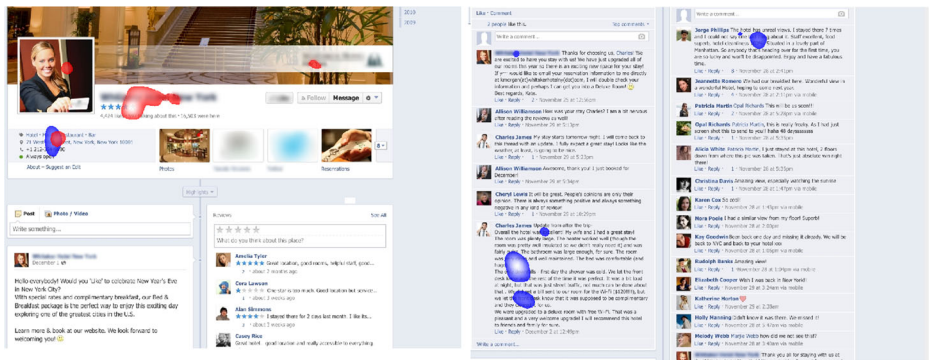


Fig. 14 Emotional salience. Gaze (blue) and arousal (red) heatmaps were created using the parameters described in section 3. Data from 18 participants were used to create both heatmaps. The most intense regions of the arousal heatmap is show in the left part and the most intense regions of the gaze heatmap is show in the right part. The interface was created manually using Photoshop (Adobe) in order to look like a generic social networking site. Parts of the image have been blurred after the experiment to hide copyrighted visual elements (e.g. name of the hotel)

The experimental validation also showed that the spatial resolution of the method is related to the temporal resolution of the selected signals. Only pupil size-based physiological heatmaps successfully achieved scalability and were able to sustain a higher spatial resolution. In future work, the signal extraction step will be enhanced in order to address the physiological carry over effect underlying this limitation. In the current state of the tool, many physiological heatmaps can be displayed at the same time to represent different psychological variables. However, the rendered heatmaps are independent and occlude each other. Another interesting next step would be to use multivariate color representation techniques [60] to create a third semantic rendering from the blending of two overlapping heatmaps (e.g. positive valence + high arousal = excitement). It is also worth noting that the heatmaps presented in this paper were created using only one physiological signal at a time for sake of clarity and concision while describing the method. In order to level the full potential of the method, heatmaps should be based on the result of an inference process applied to many different physiological signals at once.

Finally, we described a concrete use case with data collected in another study on a hotel's web page. The proposed method was able to provide useful information that was unavailable using eyetracking only. There are many other HCI applications for which physiological heatmaps could be used, from improving user experience, to identifying cognitively overwhelming regions of a training simulator interface. Eyetracking and gaze heatmaps have proven to be useful tools to help grasp where users are looking on an interface; physiological heatmaps have been designed to help understand why. As such, our goal is to enrich the toolbox of UX and HCI practitioners in order to help them to better understand their users.

Acknowledgements Authors want to thank Vanessa Georges for manuscript revision and artwork editing and Emma Campbell for the experiment preparation. The authors also want to thank the research assistants who administered the experiment. This work is supported by the SSHRC (Social Sciences and Humanities Research Council) and the FRQNT (Fonds de recherche du Québec - Nature et technologies).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Aboyoun DC, Dabbs JM (1998) THE HESS PUPIL DILATION FINDINGS: SEX OR NOVELTY? *Soc Behav Personal Int J* 26(4):415–419. doi:[10.2224/sbp.1998.26.4.415](https://doi.org/10.2224/sbp.1998.26.4.415)
2. Bailey BP, Iqbal ST (2008) Understanding changes in mental workload during execution of goal-directed tasks and its application for interruption management. *ACM Trans Comput-Hum Interact* 14(4):1–28. doi:[10.1145/1314683.1314689](https://doi.org/10.1145/1314683.1314689)
3. Beatty J, Lucero-Wagoner B (2000) The pupillary system. In: Cacioppo LGT, Berntson GG (eds) *Handbook of psychophysiology*, 2nd edn. Cambridge University Press, Cambridge
4. Birmingham E, Bischof WF, Kingstone A (2009) Saliency does not account for fixations to eyes within social scenes. *Vis Res* 49(24):2992–3000. doi:[10.1016/j.visres.2009.09.014](https://doi.org/10.1016/j.visres.2009.09.014)
5. Blignaut P (2010) Visual span and other parameters for the generation of heatmaps. Paper presented at the Proceedings of the 2010 Symposium on eye-tracking research \&\#38; Applications, Austin, Texas
6. Bojko A (2009) Informative or misleading? Heatmaps deconstructed. In: Jacko J (ed) *Human-computer interaction, New trends*, vol 5610. Berlin Heidelberg, Lecture Notes in Computer Science. Springer, pp 30–39. doi:[10.1007/978-3-642-02574-7_4](https://doi.org/10.1007/978-3-642-02574-7_4)
7. Borji A, Itti L (2014) Defending Yarbus: eye movements reveal observers' task. *J Vis* 14(3):29

8. Borland D, Taylor RM (2007) Rainbow color map (still) considered harmful. *Computer Graphics and Applications*, IEEE 27(2):14–17. doi:[10.1109/MCG.2007.323435](https://doi.org/10.1109/MCG.2007.323435)
9. Boucsein W (2012) *Electrodermal activity*, 2nd edn. Springer, Berlin
10. Bourke P (1996) Colour Ramping for Data Visualisation http://paulbourke.net/texture_colour/colourspace/. Accessed August 2014
11. Bradley MM, Miccoli L, Escrig MA, Lang PJ (2008) The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology* 45(4):602–607
12. Breslow LA, Ratwani RM, Trafton JG (2009) Cognitive models of the influence of color scale on data visualization tasks. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 51(3):321–338. doi:[10.1177/0018720809338286](https://doi.org/10.1177/0018720809338286)
13. Cacioppo JT, Tassinary LT, Bernston G (2007) Psychophysiological science : interdisciplinary approaches to classic questions about the mind. In: Cacioppo JT, Tassinary LG, Bernston GG (eds) *Handbook of psychophysiology*, Third edn. Cambridge University Press, New York, pp 1–18
14. Cacioppo JT, Tassinary LT, Bernston G (2007) *Handbook of psychophysiology*, 3rd edn. Cambridge university Press, New York
15. Calvo RA, D'Mello S (2010) Affect detection: an interdisciplinary review of models, methods, and their applications. *Affective Computing*, IEEE Transactions on 1(1):18–37. doi:[10.1109/t-affc.2010.1](https://doi.org/10.1109/t-affc.2010.1)
16. Conati C (2011) Combining cognitive appraisal and sensors for affect detection in a framework for modeling user affect. In: Calvo RA, D'Mello SK (eds) *new perspectives on affect and learning technologies*, vol 3. Explorations in the learning sciences, instructional systems and performance technologies. Springer New York, pp 71–84. doi:[10.1007/978-1-4419-9625-1_6](https://doi.org/10.1007/978-1-4419-9625-1_6)
17. Courtemanche F, Dufresne A, LeMoyné É (2014) Multiresolution feature extraction during psychophysiological inference: addressing signals Asynchronicity. In: da Silva HP, Holzinger A, Fairclough S, Majoe D (eds) *Physiological computing systems*. Lecture Notes in Computer Science. Springer, Berlin Heidelberg, pp 42–55. doi:[10.1007/978-3-662-45686-6_3](https://doi.org/10.1007/978-3-662-45686-6_3)
18. Coy AL, Hutton SB (2012) The influence of extrafoveal emotional faces on prosaccade latencies. *Vis Cogn* 20(8):883–901
19. Davidson RJ (1988) EEG measures of cerebral asymmetry: conceptual and methodological issues. *International journal of neuroscience* 39(1-2):71–89. doi:[10.3109/00207458808985694](https://doi.org/10.3109/00207458808985694)
20. Dawson ME, Schell AM, Filion DL (2007) The Electrodermal system. In: Cacioppo JT, Tassinary LG, Bernston GG (eds) *Handbook of psychophysiology*, Third edn. Cambridge University Press, New York, pp 159–181
21. Dimigen O, Sommer W, Hohlfield A, Jacobs AM, Kliegl R (2011) Coregistration of eye movements and EEG in natural reading: analyses and review. *J Exp Psychol Gen* 140(4):552–572. doi:[10.1037/a0023885](https://doi.org/10.1037/a0023885)
22. Duchowski A (2007) *Eye tracking methodology : theory and practice*. Springer, London
23. Duchowski AT, Price MM, Meyer M, Orero P (2012) Aggregate gaze visualization with real-time heatmaps. Paper presented at the proceedings of the symposium on eye tracking research and applications. Santa Barbara, California
24. Dufresne A, Courtemanche F, PromTep S, Sénécal S (2010) Physiological Measures, Eye Tracking and Task Analysis to Track User Reactions in User Generated Content. In: *measuring behavior*, Eindhoven, The Netherlands. pp 218–222
25. Fairclough SH (2009) Fundamentals of physiological computing. *Interact Comput* 21(1–2):133–145
26. Ganglbauer E, Schrammel J, Deutsch S, Tscheligi M (2009) Applying psychophysiological methods for measuring user experience: possibilities, challenges and feasibility. Paper presented at the user experience evaluation methods in product development (UXEM'09) workshop at the 12th IFIP TC13 conference on human-computer interaction
27. Georges V, Courtemanche F, Senecal S, Baccino T, Fredette M, (2016) Leger P-M UX heatmaps: mapping user experience on visual interfaces. In: *Proceedings of the CHI conference on human factors in computing systems*, 2016. ACM, pp 4850–4860
28. Giakoumis D, Tzovaras D, Moustakas K, Hassapis G (2011) Automatic recognition of boredom in video games using novel biosignal moment-based features. *IEEE Trans Affect Comput* 2(3):119–133
29. Hairston WD (2012) Accounting for timing drift and variability in contemporary Electroencephalography (EEG) systems. Army research lab. Aberdeen proving ground MD
30. Holmqvist K, Nyström M, Andersson R, Dewhurst R, Jarodzka H (2011) *Eye tracking: a comprehensive guide to methods and measures*. Oxford University Press
31. Humphrey K, Underwood G, Lambert T (2012) Salience of the lambs: a test of the saliency map hypothesis with pictures of emotive objects. *J Vis* 12(1):22
32. Irwin D (1992) Visual memory within and across fixations. In: Rayner K (ed) *eye movements and visual cognition*. Springer series in neuropsychology. Springer New York, pp 146–165. doi:[10.1007/978-1-4612-2852-3_9](https://doi.org/10.1007/978-1-4612-2852-3_9)

33. Itti L, Koch C (2000) A saliency-based search mechanism for overt and covert shifts of visual attention. *Vis Res* 40(10–12):1489–1506. doi:[10.1016/S0042-6989\(99\)00163-7](https://doi.org/10.1016/S0042-6989(99)00163-7)
34. Jennings LR, Gianaros PJ (2007) Methodology. In: Cacioppo JT, Tassinary LG, Bernston GG (eds) *Handbook of psychophysiology*, Third edn. Cambridge University Press, New York, pp 812–833
35. Jennings JR, Kamarck T, Stewart C, Eddy M, Johnson P (1992) Alternate cardiovascular baseline assessment techniques: vanilla or resting baseline. *Psychophysiology* 29(6):742–750. doi:[10.1111/j.1469-8986.1992.tb02052.x](https://doi.org/10.1111/j.1469-8986.1992.tb02052.x)
36. Jung T-P, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ (2000) Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clin Neurophysiol* 111(10):1745–1758. doi:[10.1016/S1388-2457\(00\)00386-2](https://doi.org/10.1016/S1388-2457(00)00386-2)
37. Just MA, Carpenter PA (1976) Eye fixations and cognitive processes. *Cogn Psychol* 8(4):441–480. doi:[10.1016/0010-0285\(76\)90015-3](https://doi.org/10.1016/0010-0285(76)90015-3)
38. Kivikangas M, Nacke L, Ravaja N (2011) Developing a triangulation system for digital game events, observational video, and psychophysiological data to study emotional responses to a virtual character. *Entertainment Computing* 2(1):11–16. doi:[10.1016/j.entcom.2011.03.006](https://doi.org/10.1016/j.entcom.2011.03.006)
39. Koch C, Ullman S (1987) Shifts in selective visual attention: towards the underlying neural circuitry. *Matters of Intelligence*. Springer, In, pp 115–141
40. Kreibitz SD (2010) Autonomic nervous system activity in emotion: a review. *Biol Psychol* 84(3):394–421. doi:[10.1016/j.biopsycho.2010.03.010](https://doi.org/10.1016/j.biopsycho.2010.03.010)
41. Laeng B, Sirois S, Gredebäck G (2012) Pupillometry. *Perspect Psychol Sci* 7(1):18–27. doi:[10.1177/1745691611427305](https://doi.org/10.1177/1745691611427305)
42. Lang PJ, Bradley MM, Cuthbert BN (2008) International affective picture system (IAPS): affective ratings of pictures and instruction manual. Technical report B-3. University of Florida, Gainesville, FL. Doi: citeulike-article-id:7208496
43. Léger P-M, Sénécal S, Courtemanche F, Guinea A Od, Titah R, Fredette M, Lemoyne ÉL (in press) applying eye fixation-related potentials to information systems research: demonstration of the method during natural IS use and guidelines for research. *Journal of the Association of Information Systems*
44. Lujin W, Giesen J, McDonnell KT, Zollner P, Mueller K (2008) Color Design for Illustrative Visualization. *Visualization and Computer Graphics, IEEE Transactions on* 14(6):1739–1754. doi:[10.1109/TVCG.2008.118](https://doi.org/10.1109/TVCG.2008.118)
45. McQuiggan S, Lee S, Lester J (2007) Early prediction of student frustration. In: Paiva AR, Prada R, Picard R (eds) *Affective computing and intelligent interaction, Lecture notes in computer science*, vol 4738. Springer, Berlin Heidelberg, pp 698–709. doi:[10.1007/978-3-540-74889-2_61](https://doi.org/10.1007/978-3-540-74889-2_61)
46. Mills M, Hollingworth A, Van der Stigchel S, Hoffman L, Dodd MD (2011) Examining the influence of task set on eye movements and fixations. *J Vis* 11(8):17
47. Nacke LE, Grimshaw MN, Lindley CA (2010) More than a feeling: measurement of sonic user experience and psychophysiology in a first-person shooter game. *Interact Comput* 22(5):336–343. doi:[10.1016/j.intcom.2010.04.005](https://doi.org/10.1016/j.intcom.2010.04.005)
48. Nielsen J, Pernice K (2012) *Eyetracking web usability*. New Riders, Berkeley, California
49. Niu Y, Todd RM, Kyan M, Anderson AK (2012) Visual and emotional salience influence eye movements. *ACM Transactions on Applied Perception* (TAP) 9(3):13
50. Pantic M, Nijholt A, Pentland A, Huanag TS (2008) Human-Centred intelligent human-computer interaction (HCI2): how far are we from attaining it? *International Journal of Autonomous and Adaptive Communications Systems* 1(2):168–187. doi:[10.1504/IJAACS.2008.019799](https://doi.org/10.1504/IJAACS.2008.019799)
51. Paris S, Durand F (2006) A fast approximation of the bilateral filter using a signal processing approach. In: Leonardis A, Bischof H, Pinz A (eds) *Computer vision – ECCV 2006, Lecture notes in computer science*, vol 3954. Springer, Berlin Heidelberg, pp 568–580. doi:[10.1007/11744085_44](https://doi.org/10.1007/11744085_44)
52. Parkhurst D, Law K, Niebur E (2002) Modeling the role of salience in the allocation of overt visual attention. *Vis Res* 42(1):107–123
53. Picard RW, Vyzas E, Healey J (2001) Toward machine emotional intelligence: analysis of affective physiological state. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 23(10):1175
54. Pizzagalli DA (2007) Electroencephalography and high-density electrophysiological source localisation. In: Cacioppo JT, Tassinary LG, Bernston GG (eds) *Handbook of psychophysiology*, 3rd edn. Cambridge University Press, New York, pp 56–84
55. Pomplun M, Ritter H, Velichkovsky B (1996) Disambiguating complex visual information: towards communication of personal views of a scene. *Perception* 25(8):931–948
56. Russell JA (1980) A circumplex model of affect. *J Pers Soc Psychol* 39(6):1161–1178. doi:[10.1037/h0077714](https://doi.org/10.1037/h0077714)
57. Salvucci DD, Goldberg JH (2000) Identifying fixations and saccades in eye-tracking protocols. Paper presented at the proceedings of the 2000 symposium on eye tracking research & applications, Palm Beach gardens, Florida, USA

58. Schmidt LA, Trainor LJ (2001) Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. *Cognit Emot* 15(4):487–500. doi:[10.1080/02699930126048](https://doi.org/10.1080/02699930126048)
59. Senecal S, Leger P-M, Fredette M, Courtemanche F, Cameron A-F, Mirhoseini S, Paquet A, Riedl R touch screen as input device: does it influence memory retrieval? In: international conference on information systems (ICIS 2013), Milan, Italy, December 16-18 2013. p 8
60. Silva S, Sousa Santos B, Madeira J (2011) Using color in visualization: a survey. *Comput Graph* 35(2):320–333. doi:[10.1016/j.cag.2010.11.015](https://doi.org/10.1016/j.cag.2010.11.015)
61. Simola J, Le Fevre K, Tomiainen J, Baccino T (2014) Affective processing in natural scene viewing: valence and arousal interactions in eye-fixation-related potentials. *NeuroImage*
62. Stanners R, Coulter M, Sweet A, Murphy P (1979) The pupillary response as an indicator of arousal and cognition. *Motiv Emot* 3(4):319–340. doi:[10.1007/BF00994048](https://doi.org/10.1007/BF00994048)
63. Tatler BW, Hayhoe MM, Land MF, Ballard DH (2011) Eye guidance in natural vision: Reinterpreting salience. *J Vis* 11(5). doi:[10.1167/11.5.5](https://doi.org/10.1167/11.5.5)
64. Torralba A, Oliva A, Castelano MS, Henderson JM (2006) Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychol Rev* 113(4):766
65. van den Broek E, van der Zwaag MD, Healey JA, Janssen JH, Westerink JHDM (2010) Prerequisites for Affective Signal Processing (ASP) - Part IV. Paper presented at the 1st International Workshop on Bio-inspired Human-Machine Interfaces and Healthcare Applications - B-Interface 2010, Valencia, Spain
66. Wilder J (1958) Modern psychophysiology and the law of initial value. *Am J Psychother* 12:199–221
67. Wooding DS (2002) Fixation maps: quantifying eye-movement traces. Paper presented at the proceedings of the 2002 symposium on eye tracking research & applications, New Orleans, Louisiana
68. Wooding D (2002) Eye movements of large populations: II. Deriving regions of interest, coverage, and similarity using fixation maps. *Behav Res Methods Instrum Comput* 34(4):518–528. doi:[10.3758/BF03195481](https://doi.org/10.3758/BF03195481)
69. Zaman B, Shrimpton-Smith T (2006) The FaceReader: measuring instant fun of use. Paper presented at the proceedings of the 4th Nordic conference on human-computer interaction: changing roles, Oslo, Norway



François Courtemanche is a researcher at the Tech3lab at HEC Montréal. He holds a PhD in computer science from the University of Montréal and a MSc in computer science from the University of Sherbrooke. He is a member of the Executive Committee and the operation manager of the Tech3Lab. He works in human-computer interactions. His main research focuses on physiological computing, eye tracking, and machine learning. His work has been published in journals such as *Interacting with Computers* and *Applied Artificial Intelligence*.



Pierre-Majorique Léger is a Full Professor of Information Systems at HEC Montréal. He holds a PhD in industrial engineering from École Polytechnique de Montréal and has done post-doctoral studies in information technologies at HEC Montréal and NYU Stern. He is also Invited professor at Henry B. Tippie College of Business (University of Iowa) and Tuck School of Business (Dartmouth University). He is the director of the ERPsim Lab and co-director of Tech3Lab. He is the principal inventor of ERPsim, a simulation game to teach ERP concepts, which is now used in more than 180 universities worldwide and many Fortune 1000 organisations. He has published articles in the Journal of the Association for Information Systems, Journal of Management of Information Systems, Information & Management, Technovation, Computers in Human Behaviour, and many others.



Aude Dufresne is a Full Professor in Communication Sciences at University of Montréal. She holds a PhD in cognitive psychology from the University of Montréal. She works in the fields of interaction and interface design and evaluation, usability evaluation, eyetracking and physiological measures to assess interaction. She has published articles in journals such as Interacting with Computers, Artificial Intelligence in Education and The New Educational Review.



Marc Fredette is an Associate Professor in Management Sciences at HEC Montréal. He holds a PhD in Statistics from the University of Waterloo. He is a member of the Executive Committee of the Tech3Lab and he is in charge of the Consultation Center in Statistics at HEC Montréal. He is also an Associate Editor for Statistical Modelling. He has published articles in journals such as *Biometrika*, *Technometrics*, *Computational Statistics & Data Analysis*, the *Canadian Medical Association Journal*, and *Accident Analysis and Prevention*.



Dr. Élise Labonté-LeMoine is a postdoctoral researcher in Information technology at HEC Montréal, a member of the executive committee of the Tech3Lab and an affiliated researcher with ERPSim Lab. She holds a Ph.D. in Exercise Science from the University of Montreal with a specialty in neuropsychology. Her research has been published in such journals as the *Journal of the Association for Information Systems*, *Business Process Management Journal* and *Computers in Human Behavior*. Her current research centers on how various physical interactions with technology can impact our perceptions, our motivations and our decisions. She is also on the board of directors of Blitz Paradisio, a non-profit dedicated to diminishing the drop out rates in grad school through cross-disciplinary mentorship and writing retreats.



Sylvain Sénécal is Professor of Marketing, RBC Financial Group Chair of E-Commerce, and Co-director of the Tech3Lab at HEC Montreal. His research focuses on online consumer behavior, electronic commerce, and consumer neuroscience. His work has been published in journals such as *International Journal of Electronic Commerce*, *Journal of Retailing* and *Journal of the Association for Information Systems*. Prior to his academic career, he held marketing positions at Arcelor Mittal.