Deep Breath:

Uncovering the missing dimension in consumer emotions

Graham Page - EVP, Commercial Research & Innovation,

iMotions, Copenhagen, Denmark

Presented at ESOMAR Congress 2026

Abstract

As the research industry has become more focused on automated, efficient data collection and analysis via software and Gen AI, human involvement has diminished, which paradoxically makes it harder to understand how people truly feel, rather than just what they say. AI-based tools do exist to implicitly measure verbal sentiment or facial expressions: but these tools only measure certain aspects of people's emotional experience. In particular, it has been difficult to measure emotional arousal - so tension, or excitement, or anxiety – from remote participant video or stated sentiment. This paper outlines an entirely novel approach developed by iMotions, which provides this missing dimension of emotion: by measuring body movements in research participant video, it is possible to estimate respiration rate, which is a key indicator of emotional arousal. This has been validated against lab-standard physical respiration bands, and shown to work at scale with market research panellists. It has been applied to research across thousands of participants to uncover new emotional insights, and can help clients optimise ad campaigns, edit entertainment content to maximise audience engagement, and remove pain points from user experiences. By increasing the range of emotions that can be measured implicitly from participant video, this approach allows researchers, (but also automated tools such as LLMs), to go beyond what viewers say, to better understand how they really felt in the moment, and provide better guidance for marketing decisions as a result.

Introduction

Innovation in market research in recent years has become almost synonymous with one thing: the more efficient generation of data and findings. The templating of research protocols has minimised the number of design decisions required; the rise of automated research platforms has cut increasing numbers of human steps out of the process of data collection and tabulation; and now the adoption of Large Language Models (LLMs) and other forms of Generative AI technology has greatly accelerated the generation of research summaries and reporting, and even brought into question the need for real respondents. However, as data has become faster and cheaper to generate, paradoxically the ability to generate **insight**, particularly about consumer motivation, has become more difficult. By cutting people out of the interview and analysis process, it is harder for research professionals to go beyond what is said, to see what is really meant. LLMs, by definition, take what people say at face value, and can't see beyond the literal statements to get at the unspoken meaning.

This is particularly problematic when it comes to what arguably is the core of the Insights Industry's job, which is to understand what really motivates people to buy, or watch, or vote. It is not enough to understand what people do, and who does it, or even why they say they do it. We need to understand how people feel, or how they expect certain purchases or behaviours to make them feel, as emotion is a fundamental driver of behaviour that many consumers fail to realise is at work in their choices.

Researchers have long used observational methods that allow them to go beyond what people say, such as tone of voice, body language and facial analysis. However, this has become increasingly challenging as global research has shifted online, as the human steps have been progressively eliminated from the interview process. In this context, researchers

(and now, LLMs) know what people said, but how do they know what they really meant? In response, many researchers in the last decade or so have turned to automated analysis of participant video, using algorithmic tools to assess facial expressions, body posture and verbal sentiment.

However, these tools only measure certain aspects of people's emotional experience, as there are only so many emotions that play across the face. In particular, it has been difficult to measure emotional arousal - so tension, or excitement, or anxiety – from remote participant video or stated sentiment.

The tension in the build up to a dramatic moment in an ad, or the frustration at a poor online experience, can be missed solely by looking at expressions, or after-the fact recollections. This is a key issue, as Arousal is one of the fundamental dimensions of emotion that is studied in academic investigation, and is a key component of many well-known models such as Russell's Circumplex model of emotion[1], or the Vector model of emotion[2,3] (see Fig. 1)

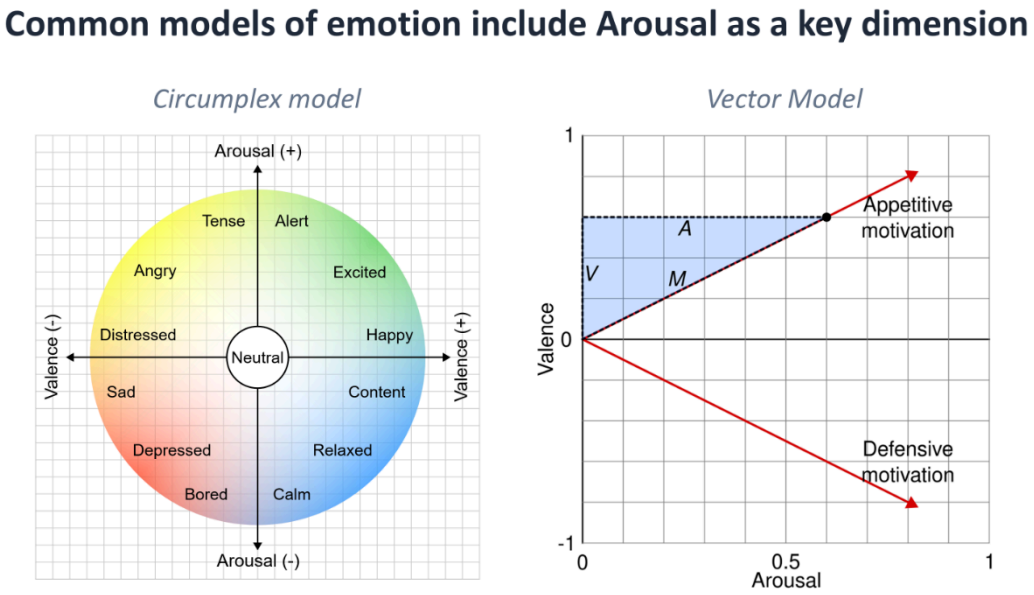


Figure 1. Russell's Circumplex model of emotion (1980)[1], and the Vector model (e.g. Bradley et al 1992)[2]. Emotions are characterised along two fundamental dimensions: Valence (positive or negative affect) and physiological Arousal (intensity).

In academic settings, arousal is measured via physiological signals which indicate the state of the nervous system such as Galvanic Skin Response (GSR - the electrical conductivity of the skin), heart rate, or respiration rate. This presents obvious challenges when applied to large scale market research, as such methods require electrodes or devices to be applied to participants' bodies. While some small-scale research panels exist for this work (such as those offered by Mediaprobe) these still require participants to apply a sensor or hold a device during fieldwork, and as such these approaches are limited in how quickly they can be applied, or the scale and speed at which research can be conducted.

However, there are additional signals available in remotely-collected participant video that can provide some of this missing insight. Our paper details how iMotions has developed a groundbreaking approach to measuring emotional arousal via the webcam-based

measurement of respiration rate, which unlocks this additional dimension of emotional response for use at much larger scale by the industry.

Video – based measurement of respiration

iMotions has built software over many years which allows the signals from different types of biosensors to be brought together and analysed in one platform. As such we have lots of experience in both direct bio-signal measurement, and measurement of signals from respondent video (historically for facial expression detection). iMotions' Affectiva facial expression technology has been used for almost 15 years by advertisers worldwide to assess audience's emotional responses to ads, entertainment content and experiences. As such this represents a huge corpus of data from which to assess and develop additional signals.

Reviewing such data, it is apparent that there are subtle signals in videos of research participants that can signal more than just facial expression. Changes in skin tone can indicate heart rate, and body movements can provide small but meaningful indications of respiration rate. In particular, it was noted that subtle changes in the shoulders and upper torso have the potential to indicate when viewers breathe in and out.

Based on this insight, iMotions developed a machine-learnt algorithm which isolates such body movements, and from this established a continuous measurement of breathing rate and its variability (see Fig. 2). Given the higher potential for noise in a participant video compared to a direct bio-sensor, quality filters were also developed to isolate the signals related specifically to breathing, and to identify cases where the signal does not conform to a normal breathing pattern.

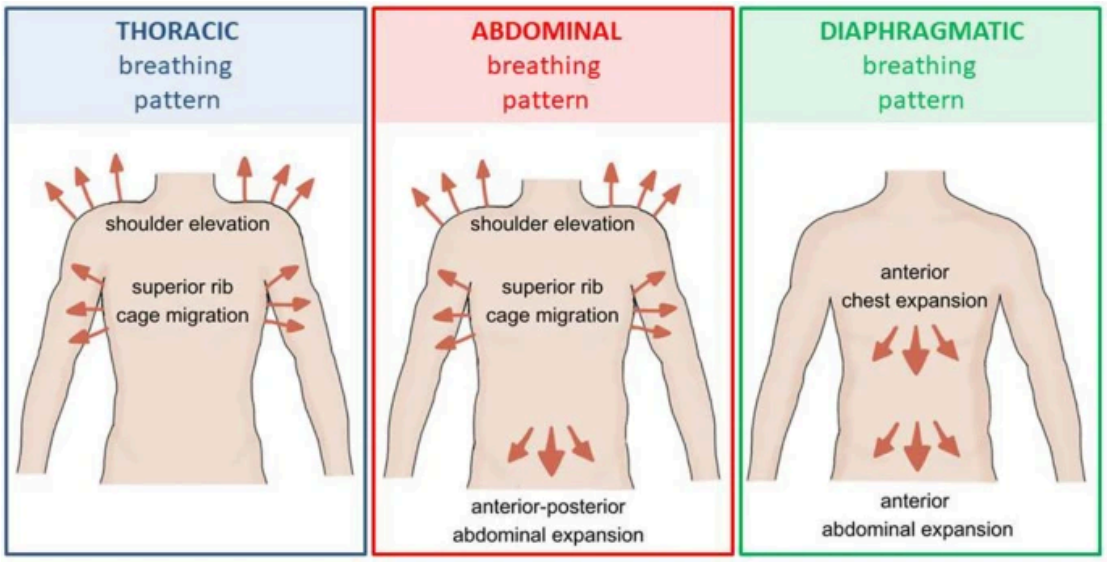


Figure 2. Body signals that indicate breathing actions[4]

Validation

As a vendor of biometric software and hardware, iMotions has unique access to lab-standard technology with which to validate the algorithm. Therefore, two validation studies were conducted: Firstly, to compare the signal from participant video to that derived from gold-standard technology, in this case the Biopac respiration belt, which directly measures the movement of the ribcage via a physical sensor; secondly to establish whether good data could be collected from research panellists in the wild[5].

In the first experiment, 10 participants were recruited and asked to watch a number of videos with varying emotional content – one baseline of neutral stimulus, and then a series of funny or emotional ads, and positive or disgusting images. In each case, responses were recorded via webcam, and respiration rate and number of cycles was also measured via a Biopac belt.

When comparing the results, despite the small number of participants, highly significant correlations ($r>0.8$) were found between the respiration metrics (respiration rate, & number of breathing cycles) from the webcam method and the Biopac belt (for example, see Fig. 3), strongly suggesting that the webcam-based method can se used as an accuracy proxy for direct measurement of respiration.

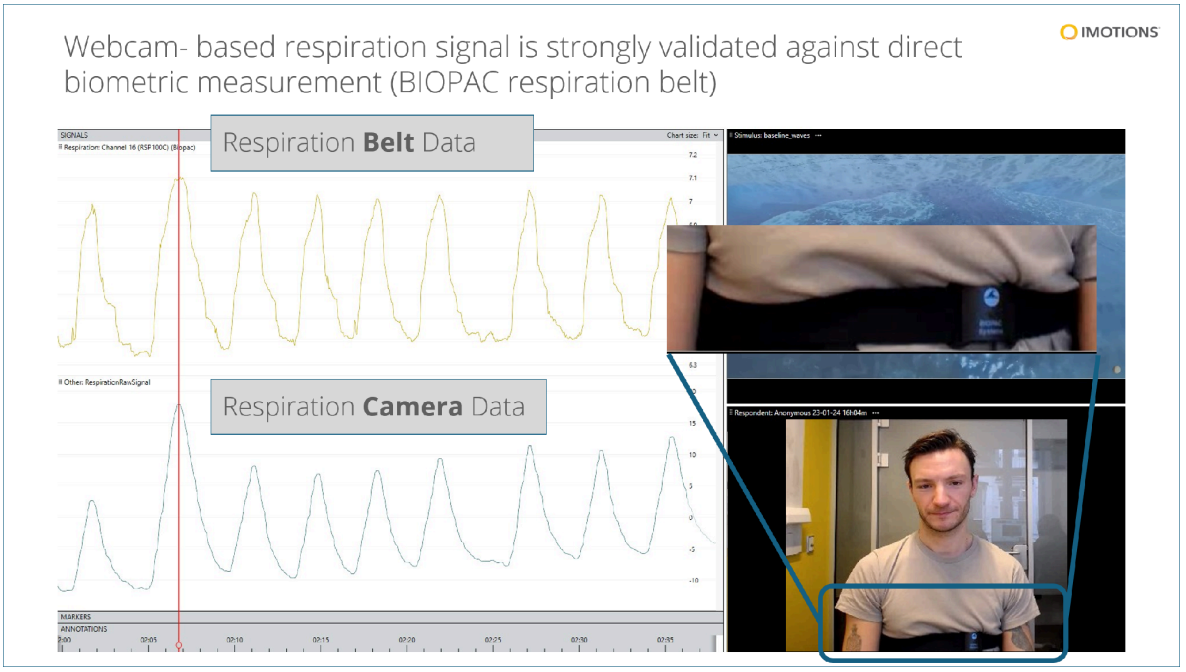


Figure 3. Comparison of respiration -rate data from one participant, collected via webcam and Biopac sensor.

The second experiment focused on the scalability of the method: can data be collected effectively from typical research panellists? To investigate, a larger scale study was completed, using research participants recruited from the Prolific panel. Again, participants were shown a variety of videos, and their response recorded (with consent) via webcam, with respiration metrics calculated based on those recordings. Of 329 qualified participants who reached the point in the interview where they were asked for consent to be filmed, 77% of them consented to be recorded and continued through the positioning check to be exposed to the video, and 205 completed the study in total. This opt-in and completion rate compares well with other comparable statistics obtained, for instance when asking participants for consent to be recorded for facial expression analysis.

Equally importantly, the webcam respiration also successfully distinguished between the different types of videos shown to participants. The stimuli consisted of a neutral baseline video, two humorous ads, and videos of positive and disgusting images taken from International Affective Picture System. Significant differences in respiration rate were observed, both between the neutral baseline video and the positive and humorous content (higher breathing rate observed) and between the positive stimuli and disgusting images (lower breathing rates observed for negative stimuli).

This suggested that not only can the webcam algorithm successfully measure breathing rate, but this can also be achieved at scale with research panellists, and there is meaning in the resulting measurement of breathing rate that can successfully differentiate emotional responses to visual stimuli such as ads or images.

As a qualitative example, Fig. 4 shows the average respiration rate over time for a sample of participants exposed to 4 videos. The first and third video are neutral baselines, showing calming undersea imagery (video 1) and nature scenes (video 3). Video 2 is first-person user-generated content of parkour (specifically a woman walking along the edge of a high-rise building), and video 4 is footage of a baby laughing. Both video 2 and 4 would be expected to be more arousing, and we do see clear evidence of higher respiration rate to those videos. The response to video 4 is also accompanied by clear positive facial expressions (smiles), but crucially, the response to video 2 is not. This is a more anxiety-inducing video, and while there are small bumps in facial expressiveness, there is a much clearer response in breathing rate, suggesting that as expected, this approach can pick up emotional responses that do not show on the face.

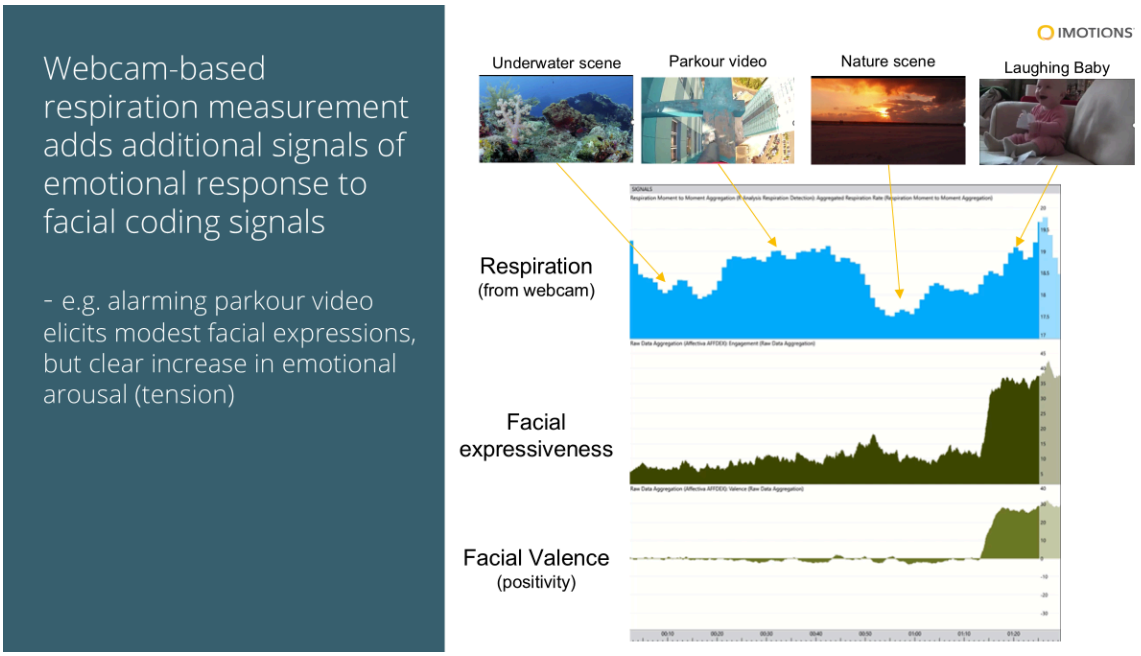


Figure 4. Webcam – based respiration and facial measurements in response to 4 videos. Anxiety-inducing parkour video generates clear respiration (arousal) change, but limited facial expression change.

Implementation & data quality safeguards

Given the successful validation of the method, the Webcam respiration algorithm has been implemented within the iMotions suite of tools, specifically within:

- *iMotions Online Research* – a DIY research platform that allows commercial researchers to set up and run online survey projects, which also include webcam-based eye-tracking, facial expression analysis and now respiration measurement
- The *Affectiva Media Analytics* platform – an enterprise-level research platform that allows the integration of webcam-based audience response measurement into ad-testing and entertainment-testing research tools (extensively used by the ad industry, with over 100k ads tested to date in the database, in over 90 countries).

This has opened up the use of the approach for a variety of applications, both in video-testing but also in wider Concept, UX and NPD work.

Given that most applications in online research are within self-serve or templated products, where speed is of the essence and human intervention is kept to a minimum, and where individual respondent videos are unlikely to be reviewed manually, the algorithms have also have additional quality safeguards implemented as part of the process. At the data collection stage, participants are given clear guidance in how to sit, and are given a framing oval to confirm that they are visible to the camera, and at the analysis stage, additional data cleaning is carried out. For example, data which provides respiration rates outside the typical human boundary of 7-30 breaths per minute is discarded, as is data which does not conform to the expected waveform for respiration (this is a threshold which can be adjusted by the user in iMotions Online). Likewise, data which exhibits large swings in breathing rate, or minimal movements is also excluded, as this is typically a signal of an artefact produced by participants moving around or moving their shoulders out of shot.

This ensures that even when data is collected on mobile devices, only valid data is included in the analysis, and motion artefacts are kept to a minimum.

Insights created

While technically impressive, the value of the respiration metric is only apparent in the additional insights that it can bring. Since implementation, the tool has been applied widely across projects run using the iMotions online platforms, and this has revealed many cases where new insights have been uncovered. While client confidentiality limits our ability so share certain examples, the following cases give a flavour of the insights this approach enables.

a. Uncovering excitement

Understanding emotional responses is a critical element of movie trailer research and iMotions has researched several thousand trailers and episodes. They often have highly complex and emotional storylines, and so widening the palette of emotional measurement is of clear benefit. Some years ago, we researched the trailer for the last James Bond movie “No time to Die”, which includes the usual mix of action, drama, humour and stunts. Viewers love the trailer, and show very clear facial reactions, but there were always a couple of moments which produced surprisingly little facial expression. Testing the trailer again, and using the respiration algorithm added clear colour to these moments, Figures 5 and 6 show the respiration rate of viewers, and also some key facial signals – “Valence”, or overall positivity, and “brow furrow” often indicating tension in the context of trailers. During a key action and fight scene, facial

expressions were slightly negative, but there was a clear spike in respiration rate, suggesting the action was indeed engaging viewer as we might expect. Similarly, later in the trailer there is a further action sequence including a dramatic motorcycle stunt – and while there was relatively little facial reaction to this, there was a clear increase in respiration rate. So, as we might hope, measuring respiration can give a clear signal about emotional engagement in the absence of facial responses.



Figure 5. Respiration and facial responses or fight sequence in James Bond trailer (key moment in red on the timeline)

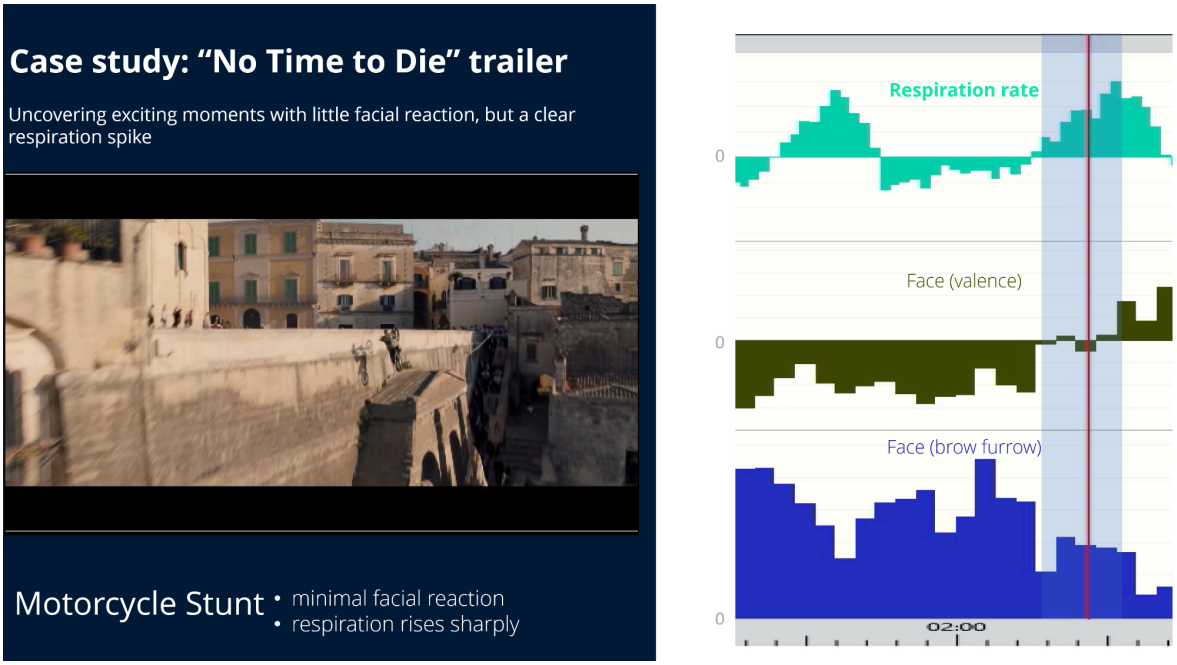


Figure 6. Respiration and facial responses to stunt and action sequence in James Bond Trailer (key moment in red on the timeline)

b. Gasp moments

Respiration does not have a linear and positive relationship with engagement. Generally speaking, faster breathing is associated with excitement, or tension, but there are other characteristics associated with additional responses – for instance, holding one's breath is part of the biological startle response, and can be an indicator of shock. An ad we recently tested for Columbia outdoor clothing illustrates this perfectly. It is a very intense commercial, with lots of fast cuts, threatening and occasionally violent and graphic imagery designed to convey how harsh the outdoors can be (and of course, why you therefore need Columbia clothing). Early parts of the ad, where the threat and intensity build, are associated with a clear increase in breathing rate. But then at a key moment (see Fig. 7), we see a sharp drop in respiration accompanied by a mouth open response – an almost perfect example of participants gasping slightly. The reason? At this moment, the ad recreated the moment where Aaron Ralston (of *127 Hours* & "I cut my own arm off to escape being trapped" fame) is trapped by a rock (again). Respiration measurement gives real clarity to the emotional power of this moment – inevitably the most memorable scene in a very intense spot.



Figure 7. Gasp moment (marked in red) at shocking scene in Columbia ad

c. Converging signals: Effort and difficulty

While communication and content testing is always fertile ground for bio-sensing technology, there are clearly other areas where measurement of arousal is important. In a recent study we were able to demonstrate a clear interaction between advertising exposure and subsequent effort in a brand-location task at a shelf front. All participants were shown a virtual shelf and asked to locate a specific brand (CeraVe), but half the participants were shown an ad for the brand first, whereas the other half saw the ad after the shelf task. Participants were recorded using their webcam, and eye-tracking facial analysis and respiration measurement were conducted on the recordings. The

hypothesis was that exposure to the ad would increase the saliency of the brand and packaging, and make the search task easier – which is indeed what we found: Participants who saw the ad first showed greater visual attention on the target brand during the task than those who did not. However, there was also evidence that they found the task less onerous or stressful, as respiration rate was lower during the task amongst those exposed to the ad first, and they also showed lower levels of brow furrow, which is often associated with effort in this scenario (see Fig. 8). In this case, both the expression and respiration data converged to indicate lower effort during the task, and the improved attention to the brand showed a positive outcome of this lower effort. This both illustrates the potentially subtle but important role of point-of-sale advertising, and shows how respiration and other video signals can tell us something about how much effort we are requiring of consumers at retail.

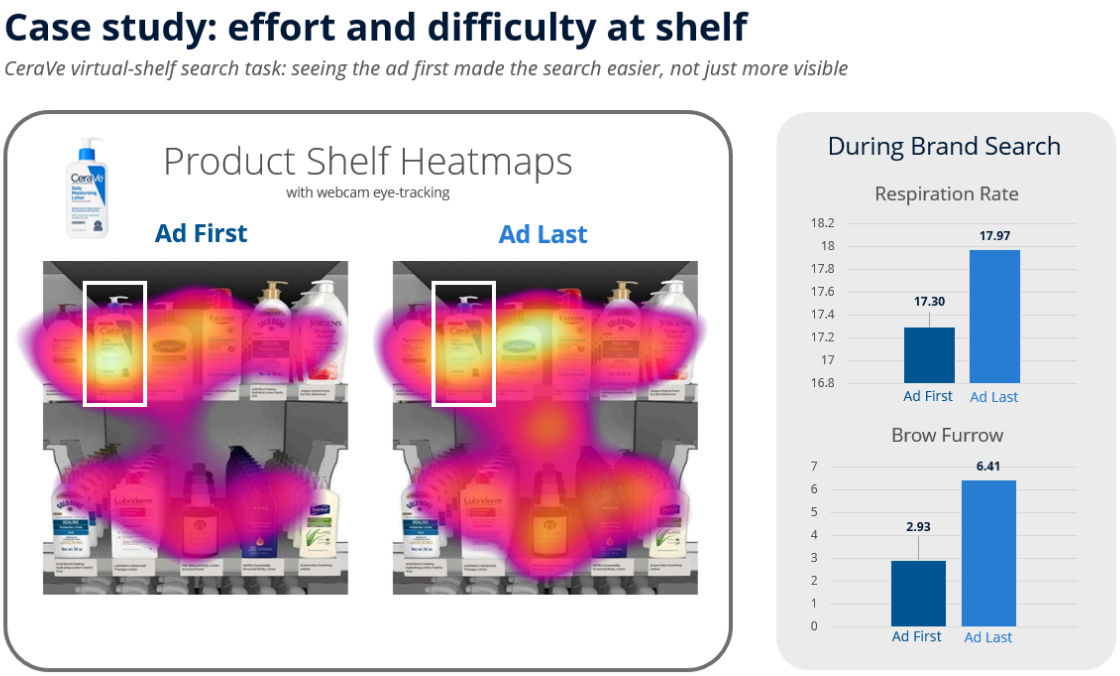


Figure 8. Eye tracking, facial expression data (brow furrow) and respiration rate data during brand-search shelf task (target brand: CeraVe). Participants split into 2 groups, Ad First (shown brand ad before task) and Ad Second (saw brand ad after the task).

6. Applications

This novel approach to bio-signal measurement has many applications, as already suggested in the case studies above, and has already begun to be used at scale by many iMotions clients. This respiration signal can be applied whenever there is a need to go beyond what is said in an interview or survey, to dig deeper into consumer reactions that participants may be unaware of, unable to remember, or feel reluctant to discuss. It has real potential to add a further dimension to the measurement of consumer reactions “in the moment”, without requiring them to self-monitor and interrupt or distort their experience. By simply requiring a webcam and viewer consent, this can be applied along side survey

instruments and alongside existing currencies of direct measurement, such as facial expression analysis. Thus far, we have seen it to be particularly useful in:

a. *Advertising and brand content research*: adding understanding of the emotional flow of video narrative, and whether viewers are truly engaged in key moments, and how this leads to stronger or weaker ad performance.
b. *Creator content research*: this is particularly relevant for content which may be considered more edgy or contentious, and over which brand owners may need more objective measures of responses than available through online metrics of likes, clicks and shares.
c. *Entertainment content testing*: the addition of direct arousal measurement is a significant step forward given the need to keep viewers engaged over long periods, and the wide emotional palette that different genres draw from. Tension and immersion are key elements that are hard to measure from dial-testing or recall, which take viewers out of the experience. This is a potentially powerful alternative to GSR measurement, which is difficult to scale beyond small research panels.
d. *Qualitative research*: Alongside facial and voice signals, respiration measurement offers another signal to assess participants' non-verbal signals, which is increasingly difficult to acquire in online fieldwork. It also offers a further potential signal to tune the responses of Generative-AI moderators, and may allow them to go beyond the face value of participant speech, and identify alternative meanings, or when to dig deeper with further prompting.
e. *User Experience*: Anticipating when users become frustrated with functionality is a critical element of UX research and webcam respiration offers a potential means to spot rising tension and frustration. This adds fidelity to user recordings and may help pinpoint triggers of frustration that occurs before users seem "blow up" or abandon tasks. It may also more clearly identify moments of surprise (and combined with facial expressions when these are good or bad!)
f. *Retail*: As illustrated above, identifying effort and pain points is critical in digital retail research, and respiration measurements is a means to identify when decision-making flows, or where it sticks and purchase are potentially lost.

7. Summary: This is not the AI you were looking for…

… but it may be the AI that you need. Not all technological innovations in research involve Generative AI. While LLMs are gaining the most attention, this may mean researchers are missing other novel approaches which add **depth** to insights, rather than just **efficiency**. In an era where the human role in the research process is being compressed, video-based respiration measurement can add back some of the signals that automated tools which focus just on words may miss. As ever, this is not a replacement for other tools, but it adds an additional modality, which, taken with other signals such as facial expression or tone of voice, and alongside stated answers and surveys, can provide a more rounded view of human reactions. It is an easily implemented, bio-based signal that can help humans and machines go beyond what is said, to what is felt, and get to the truth of consumer experience.

References

1. Russell, James (1980). "A circumplex model of affect". *Journal of Personality and Social Psychology*. **39** (6): 1161–1178.
2. Bradley, M. M.; Greenwald, M. K.; Petry, M.C.; Lang, P. J. (1992). "Remembering pictures: Pleasure and arousal in memory". *Journal of Experimental Psychology: Learning, Memory, and Cognition*. **18** (2): 379–390.
3. Watson, D.; Tellegen, A. (1985). "Toward a consensual structure of mood". *Psychological Bulletin*. **98** (2): 219–235.
4. Image source : Sikora, M., Mikołajczyk, R., Łakomy, O. *et al.* (2024). “Influence of the breathing pattern on the pulmonary function of endurance-trained athletes.” *Sci Rep* **14**, 1113.
5. Gregoret, L.; Seernani, D.; Wilson, J.M. (2024). “In-Lab and Remote webcam-based Respiration: A promising candidate for neuromarketing?” *NeuroPsychoEconomics Conference, May 2024.*