SAY IT TO MY FACE!

APPLYING FACIAL IMAGING TO UNDERSTANDING CONSUMER EMOTIONAL RESPONSE

Curiouser & Curiouser

AUSTRALIAN MARKET & SOCIAL RESEARCH SOCIETY NATIONAL CONFERENCE 2011

AMSRS National Conference 2011

Best Presentation Award

Sponsored by ESOMAR

ESOMAR

AMSRS

Australian MARKET & SOCIAL RESEARCH SOCIETY

INTRODUCTION

Facial expressions are one of the strongest visual methods to convey emotions and one of the most powerful means used by human beings to relate to each other. Facial Imaging, driven by advances in machine-learning software, passively records human emotions from facial expressions recorded via camera. It can be applied to better measure consumer response to marketing stimuli (e.g. advertising, packaging, retail displays) as a robust, repeatable method that

1. Overcomes issues of direct and intrusive questioning;
2. Offers consistent response across cultures and ethnicities;
3. Is accessible across both online and offline platforms; and
4. Evaluates video (e.g.TV commercials) or static (e.g. Press Ads, Logos) stimuli, in either finished or 'storyboard' form.

Conventional research techniques are often restricted by respondents' ability to recall and difficulty in describing subtle and unconscious emotional states. This has led to increasing interest in neuroscience-influenced methods. Many methods (e.g. FMRI, EEG), although a significant step forward in sensitivity of measurement, are higher cost, potentially intrusive, and often difficult to apply across diverse populations or integrate with existing research methods.

Methodologies like Facial Imaging, while still firmly science-based, can be integrated into existing research frameworks making those frameworks more powerful. This offers the potential not only for insights into short-term issues, such as on-shelf brand choice, but also for creating more accurate, consistent measures of emotive response – measures that can be applied strategically, across brands and categories.

The paper covers recent marketing and advertising case study examples of work undertaken by ***nViso,*** a Swiss-based technology company, who have been one of the pioneers of commercial Facial Imaging using machine-learning tools.

BACKGROUND

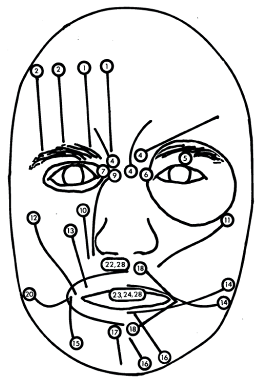
Before describing the operational issues, applications, and future potential of Facial Imaging, it is important to appreciate the historical background and academic context within which the science of understanding and interpreting human facial expressions has developed. The original idea can be traced back to Charles Darwin, upon whose work Paul Ekman et al developed and progressed the theoretical framework by which facial expressions could be consistently coded. More recently, the research undertaken by Sorci et al has explored and tested how the coding approach can be computerised to ensure consistent coding protocols across subjects by different bodies at different times and to enable and facilitate large scale applications, beyond those feasible to manual coding efforts, by introducing machine-learning software applications.

CHARLES DARWIN

As early as 1872, Charles Darwin postulated that facial expressions were universal, across cultures, ethnicities, and even species [1]. Facial expressions were essentially behaviours that evolved as a mechanism of communication. Darwin claimed expressions were both innate (i.e. expressions are not learnt) and universal (i.e. some expressions are the same in all humans and animals). David Matsumoto's work [2] comparing expressions on winning and losing athletes in the 2004 Olympics showed no difference between those at the Summer Olympic Games and blind-from-birth competitors at the Paralympic Games, implying that expressions of happinness in victory and sadness in defeat for example are 'hard-wired' and not learnt from observing others' behaviour and expressions.

PAUL EKMAN ET AL.

The most prominent work on facial expressions to date is that conducted by Dr. Paul Ekman. Dr. Ekman, ironically, originally set out to disprove Darwin's theory of universality but, in the end, concluded that to all intents and purposes Darwin was right. Ekman's work initially postulated the universality of 6 emotions (7 if we include neutral), they being: happiness, surprise, fear, anger, disgust, and sadness (+neutral). In 1978, Ekman and Friesen developed the Facial Action Coding System (FACS) to taxonomize every human facial expression. FACS measurement units are called "Action Units" (AUs) representing the muscular activity of the face. FACS is the leading global standard for facial expressions classification. See Figure 1 below for illustration.



The diagram to the left highlights illustrates the parts of the face that are captured under the Facial Action Coding System (FACS). FACS was developed Ekman and Friesen (1978) is a comprehensive set of coding standards developed in order to objectively capture the richness and complexity of facial expressions. FACS is based on the anatomy of the human face, and codes expressions in terms of component movements, called "action units" (AUs). Ekman and Friesen defined 46 AUs to describe each independent movement of the face. FACS measures all visible facial muscle movements, including head and eye movements, and not just those presumed to be related to emotion or any other human [6].

FIGURE 1: FACIAL ACTION CODING SYSTEM

SORCI ET AL

Work undertaken at the Swiss Federal Institute of Technology (EPFL) over the past decade has resulted in Ekman's work being translated into a machine-learning environment, as opposed to the 'manual' use by human expert facial coders. In essence, the work of Dr. Matteo Sorci et al has resulted in an artificial intelligence which precisely detects facial muscles. These facial muscle movements have been encoded based on Ekman's FACS. Combining artificial intelligence and the FACS taxonomy facilitates the machine-learning system which then decodes facial behaviour, allowing emotional response to various stimuli to be picked up consistently and continuously in real time.

The purpose of their work was to develop real interacting human-computer systems, where algorithms written by humans should be able to capture, mimic, and reproduce human perceptions of facial expressions.

A key issue considered in building such systems was the definition of facial expression measurements to study and quantify facial behaviour. The work undertaken focused on two major approaches in psychological research which are message and sign judgement [3].

The main task of message judgement is the inference of the displayed facial behaviour, in terms of inferred emotion. As indicated by Cohn [3], among the different descriptors available, those of Ekman [4] have been largely used in the past 20 years. Ekman proposed the use of the 6 basic emotions (happiness, surprise, fear, disgust, sadness and anger) that are universally displayed and recognized from facial expressions [5].

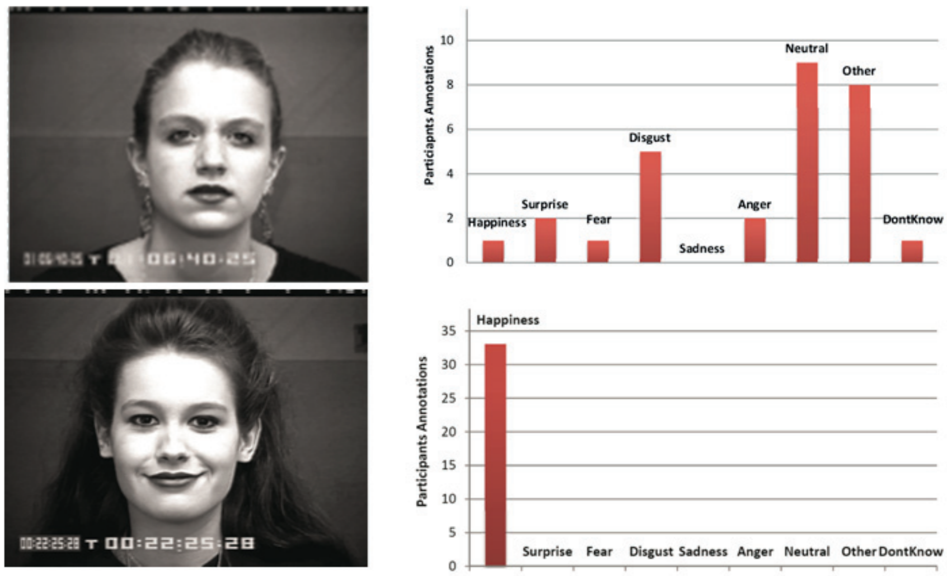
In sign judgement approaches, facial movements are used to describe facial behaviour. Among the various methods the Facial Action Coding Systems (FACS) [6][7] is the most comprehensive and widely used. The FACS is a human-observed based system designed to detect subtle changes in facial features and associates facial expression changes with actions of the muscles that produce them. For example, a nasolabial furrow, running down from the nostrils outward beyond the corners of the lips, can be judged as "sadness" in a message-judgement and as a facial movement that raises the cheeks in a sign-judgement approach. In other terms, while message judgement is all about interpretation, sign judgement attempts to be objective. A typical automatic facial expressions recognition system [8][9][10] is based on a representation of each expression, learned from a training set of pre-selected meaningful features. In the learning process, an expert is asked to associate labels to training samples. An expert should be someone having a strong knowledge of the problem, in order to ensure the correctness of what is trying to be reproduced.

Three important questions arose from this fundamental hypothesis of this "learning by examples" technique:

- Can one expert be representative of humans' perception?
- How to get and use the experts' strong knowledge?
- How to represent the visual information used by the experts?

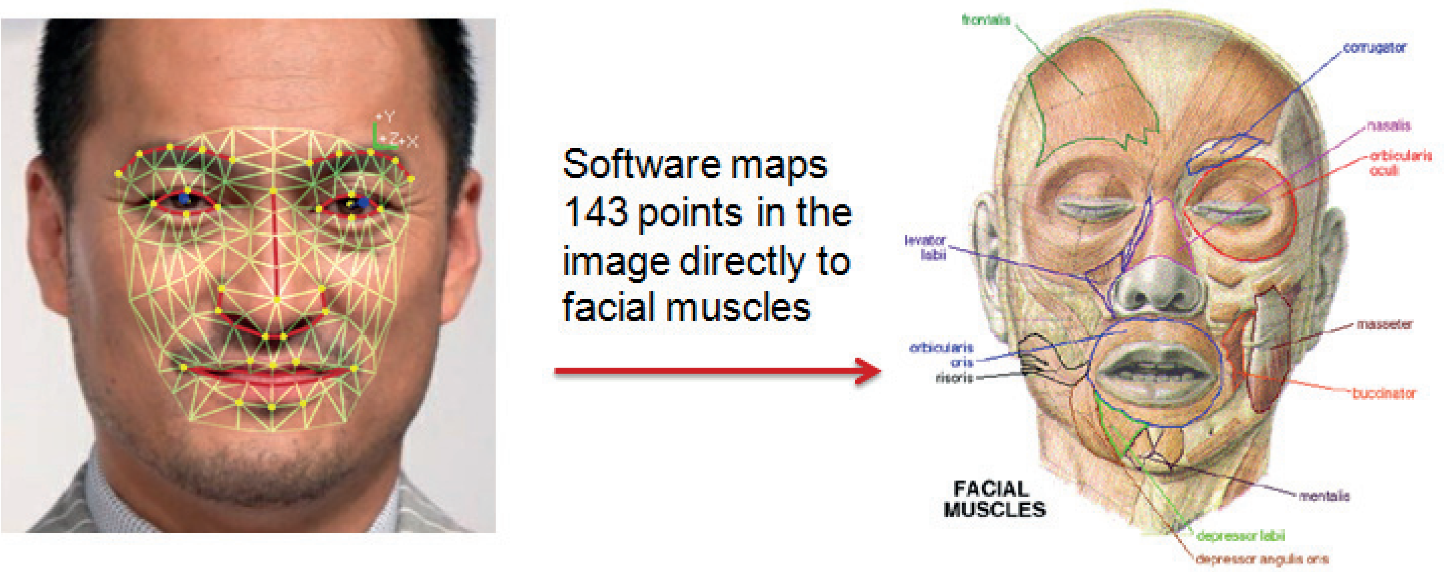
To illustrate the challenges faced, data collected through a recent web-based survey show that the perception (i.e. labelling) of a human facial expression by a human observer can be affected by a subjective component, which results in a lack of a unique ground-truth [11]. In Figure 2 below, the lower of the two facial expressions is almost universally seen as being happy by those perceiving it. However, the more 'ambiguous' expression in the upper photograph generates a much wider perceptions of emotional expressions from those perceiving it.

Figure 2: Examples of Heterogeneous and Homogeneous Judgements in Data Collected by Survey



It is this variability in expert assessments that machine-learning can help reduce. Hence most recent attempts in the representation of visual information for facial expression have focused on reproducing the set of rule descriptors suggested by the FACS system. Based on this system, a facial expression can be linguistically described in terms of measures that can be extracted from the face. These measures can be considered as the mathematical representations of local facial features. In the last decade, discussions in psychophysics and cognitive psychology [12][13][14][15] have shown that face recognition and perception of emotions rely on featural1 and configural2 information. Human's visual perception of a face involves the processing of both local facial measures and their holistic spatial layout. The implication of these findings is that an automatic system, aiming at interpreting faces needs to extract and make use of these two sources of information as well.

Figure 3: Software Maps 143 Points in the Image Directly to Facial Muscles



The literature on the research and development undertaken since the turn of the 21st century is extensive and listed in the appendix of references. To expand on the development of the modelling approach used in the algorithms here would warrant an extended paper in itself. In brief, the application of a set of discrete choice models (DCM) was used to fine-tune the predictive ability of the system. The descriptiveness of the model was improved by sequentially introducing complementary set of features. The estimation of the three proposed models, showed the correctness of the chosen sets of features, revealing the best fitting behaviour of the final third (and most complex) model. Figure 3 above shows, diagrammatically, how software can apply the modelling conventions to generate a 'spider' map which maps 143 inter-related points directly to the movements in the facial muscles as picked up by the camera.

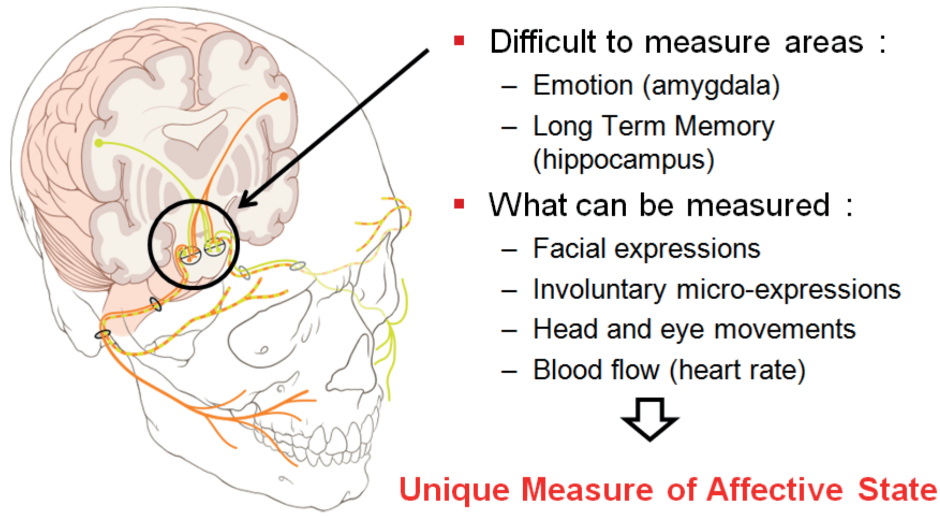
FACIAL EXPRESSIONS AND THE LINK TO THE BRAIN

It will come as no surprise that one of the most prominent, promoted, and debated innovations in the market research world in the past decade is the introduction of neuroscience and biometric based applications to measuring consumer emotions. Facial Imaging, as captured by machine-learning software, is essentially a biometric method whose 'biological manifestations' (i.e. facial expressions) are strongly linked to the limbic system in the brain, which is notoriously hard to measure.

The limbic system of the brain, sketched in Figure 4, is the primary part of the brain that regulates emotion. The amygdala is one of the main areas to process emotional content of behaviour and memory. The hippocampus converts objective versions of events from short term to long term memory. Together the hippocampus and the amygdala can shape memories, combining the emotional version of the amygdala with the more objective version of the hippocampus. Because of the effect of the amygdala on the memory, the emotional state of the learner will change the subjects' perception of the memory. Quantifying the impact of emotion in memory necessarily involves these two regions in the brain. To date, these areas and their interaction are difficult to measure and most of the methods are not able to define robust measures related to them. Facial Imaging provides a way to compute a set of unique measures to help decipher the consumer affective state. The face, in this setting, can be considered as a window on the brain thanks to the facial nerve that controls the muscles of facial expression as a consequence of impulses sent through the amygdala.

Figure 4 below perhaps illustrates better the link between these deep-seated parts of the brain and the facial nerves and muscles that generate not only the basic facial expressions but the innate involuntary micro-expressions, usually undetected by the untrained eye. These deep-seated areas of the brain also influence head and eye movements and blood flow, detected by technologies such as eye-cameras and biometric sensor belt that pick up skin sweat, heart rate, and respiration

Figure 4: Facial Nerves and the Link to the Brain



A BRIEF COMPARISON SUMMARY OF NEURO/BIOMETRIC METHODS

Neuro/biometric Neuro/Neuro/biometric methods take on wide variety forms in both how they operate and what they actually measure. Furthermore, their nature of operation and interpretation varies considerably from one to another. To put facial imaging into perspective Figure 5 summarises the key current neuromarketing or neurometric methods, including those which are strictly-speaking, biometric.

METHOD	DESCRIPTION	FEATURES
FMRI	Records increased oxygen levels in blood flows from deep within the brain	High spatial resolution to detect and locate activity in specific areas of the brain at all depths
EEG	Measures tiny electrical signals produced by the brain	Able to detect activity between left and right pre-frontal cortex, which represent positive and negative effects respectively. More sensitive to areas near surface of the brain
EYE-TRACKING	Based on view that when someone's visual attention is fixed on a stimulus, they cognitively process information at the same time	Precise measurements of stimuli viewed, lengths of viewing and directions of movement Equipment now of low intrusion, some non-intrusive
BIO-SENSORS	Degree of arousal, created by stimuli by using sensors that pick up skin sweat, heart rate, and respiration	Measures key physiological reactions felt to be indicators of emotion
FACIAL CODING	Based on Ekman's theories on universality of expressions and micro-expressions that are unconscious measurements of emotional response	Experts examine and interpret real-time or recorded interviewee responses to stimuli, direct questioning looking for movements in facial muscles to explain underlying emotional reactions
FACIAL IMAGING (FACS)	Based on Ekman's theories on universality of expressions and micro-expressions that are unconscious measurements of emotional response	Applies machine-learning technology and integrates FACS coding system to provide emotional measurement via artificial intelligence

It is worth just specifically contrasting Facial Imaging against Facial Coding. Facial Coding is the method of having a human trained in FACS to annotate the movement of the facial muscles. Facial Imaging is the method of teaching a machine to understand a face and estimate automatically what is going on. If correctly programmed, Facial Imaging offers scalability and ongoing consistency and repeatability of measurement to enhance the depth and expand the breadth of insight generated from applying the highly skilled Facial Coding technique.

FACIAL IMAGING –IN OPERATION

Assuming the reader accepts the current state of the theory behind the application of Facial Imaging, the next issues to be addressed are

1. How do we make it work?

2. What does the output look like?

3. What potential applications does it have in the sphere of marketing and social research?

Before work began on aggregating, analysing, and interpreting the data collected via machine-learnt facial imaging, the operational and executional features needed to be developed and tested. This is no different from any other form of traditional market or survey research; accuracy and robustness of the data are fundamental pre-requisites for valid interpretation. These are basic hygiene factors that cannot be ignored.

Data collection via webcam, whether online in home/work or offline/online in a central location or theatre setting, is not without its challenges. These can be categorized in three broad groups:-

1.Technical, e.g. programming, validation, comparison to ‘expert’ coding

2.Practical, e.g. band width, lighting, eyewear, facial hair

3 .Data processing, e.g. volume, visualization, delivery

Clearly, all these needed to be addressed and overcome before any analytical approaches were developed.

TECHNICAL CHALLENGES

The challenges of matching, then surpassing manual ‘expert’ coding have been expounded above both from an academic and a practical perspective.

Many technical challenges have been addressed and solved in order to make the Facial Imaging approach robust to out-of-the-lab applications. The first challenge is to give a computer the ability to recreate a basic “vision” of a face that humans take for granted. This is actually a two-fold challenge involving on one side the creation of mathematical models and algorithms to make the interpretation of a face automatic and on the other side running these algorithms in real-time.

Figure 6 below shows the ‘spider’ mask superimposed on a subject’s face. This is the key measurement tool that the machine-learning software superimposes on the respondents’ faces to record the emotional reactions to the visual stimuli. The mask consisting of 143 points around the main facial components is a representation of a face from the computer’s point of view and provides a direct link to key muscles on the face. This is essentially the tool which in identifying and then linking the key muscle points on the face allows expressions to be captured and coded into the FACS system. From this the application of validated metrics allows the processing, aggregation, and interpretation of the facial imaging data collected.

Key to the viable application of the method is its ability to consistently measure the same facial expressions over the full range of physiologies. In other words, we need to show that a particular set of movements displaying a given emotion or set of emotions will be recorded on one respondent the same way as any other respondent displaying the same emotions, despite superficial physical differences, resulting from differences in age, sex, skin colour, or ethnicity.

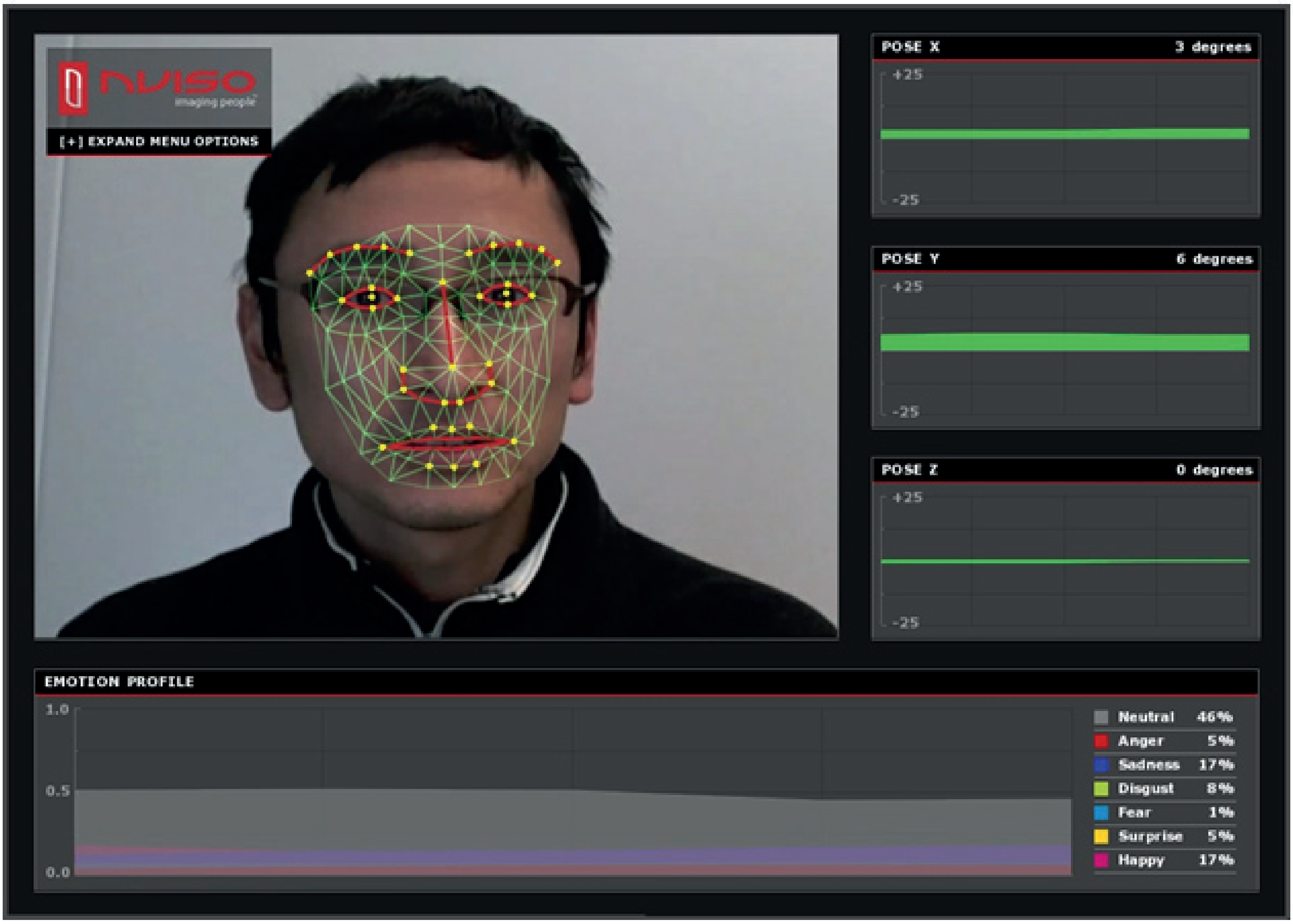


Figure 6: 'Spider' Mask Automatically Superimposed over Respondent's Face

PRACTICAL CHALLENGES

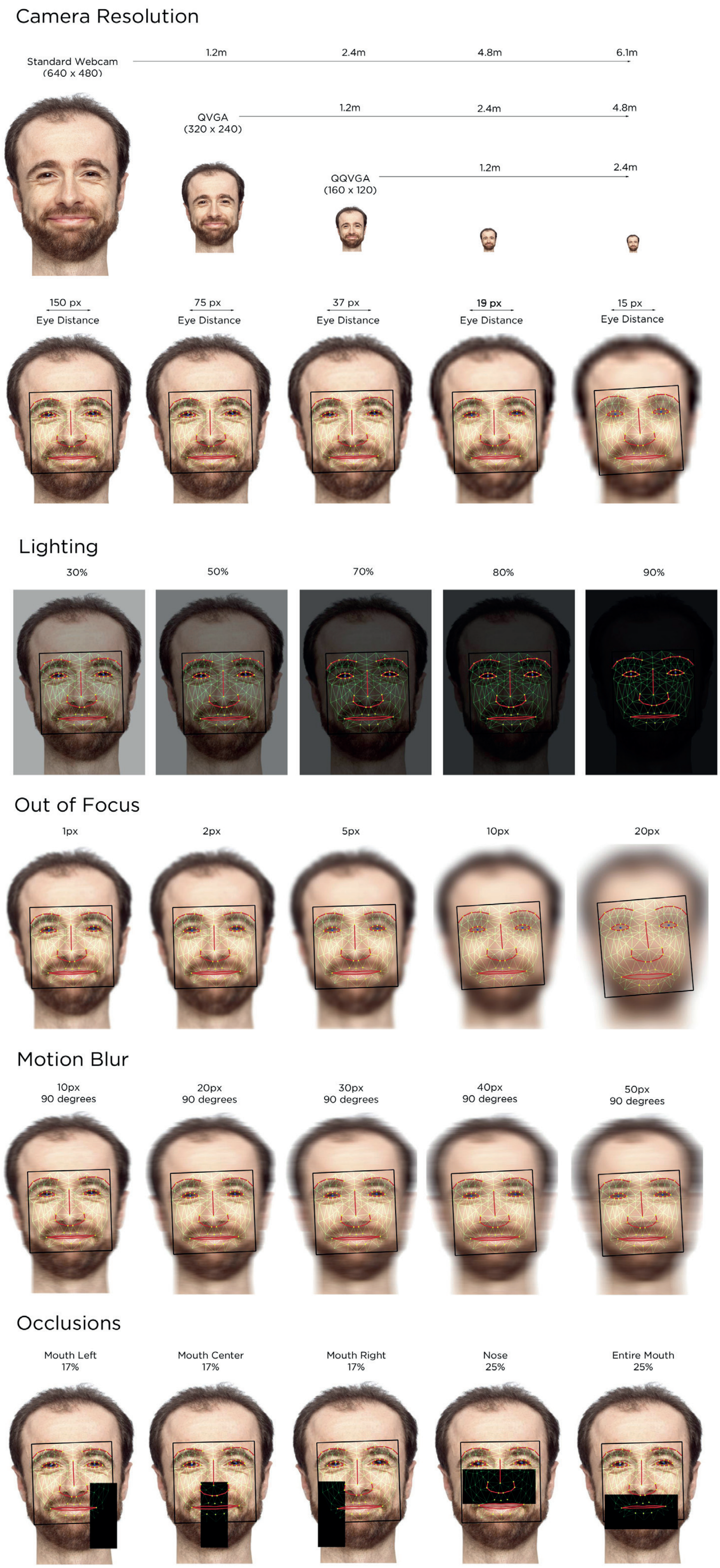
In reality, most of the images captured will not be in 'ideal' studio like lighting environments, where the face of the respondent is perfectly lit and stable to ensure flawless capture. Certainly, in the case of in home interviewing, the range of environments, light sources, posture, and the like are extremely heterogeneous. Nevertheless, these are the 'natural' environments which the consumers/respondents inhabit and where they lead their lives. So it is important that the technology works in an effective and non-intrusive fashion in these settings.

Computer scientists have been studying automated facial analysis for the past 20 years and this daunting task has become easier. The technology now exists to overcome the aforementioned practical issues in capturing expressions in natural environments with sufficient precision to allow automated analysis. Computing power and algorithms are now advanced enough to make this viable at present and into the future. The software has to be sensitive and powerful enough to capture the hundreds of measurements points that track 43 facial muscles in real-time under demanding natural environments.

Furthermore, it has to deal with common features of peoples' physical appearance. It is reasonable to request that respondents do not speak, eat, or smoke while watching the stimulus material. These activities distort the muscles whose movements display the respondents' emotional responses. However, the software has to be robust enough to overcome challenges that are everyday such as the presence of spectacles/eyewear and facial hair.

Figure 7 below demonstrates the range of factors including camera resolution, lighting conditions, out-of-focus, and motion blur and appearances with which the software has to cope. In all these cases, the 'spider' mask is able robust in real world environments to be able to successfully catch the emotional responses to a set of TV commercials shown.

Figure 7: Typical Range of Environments Where Facial Imaging has to Operate



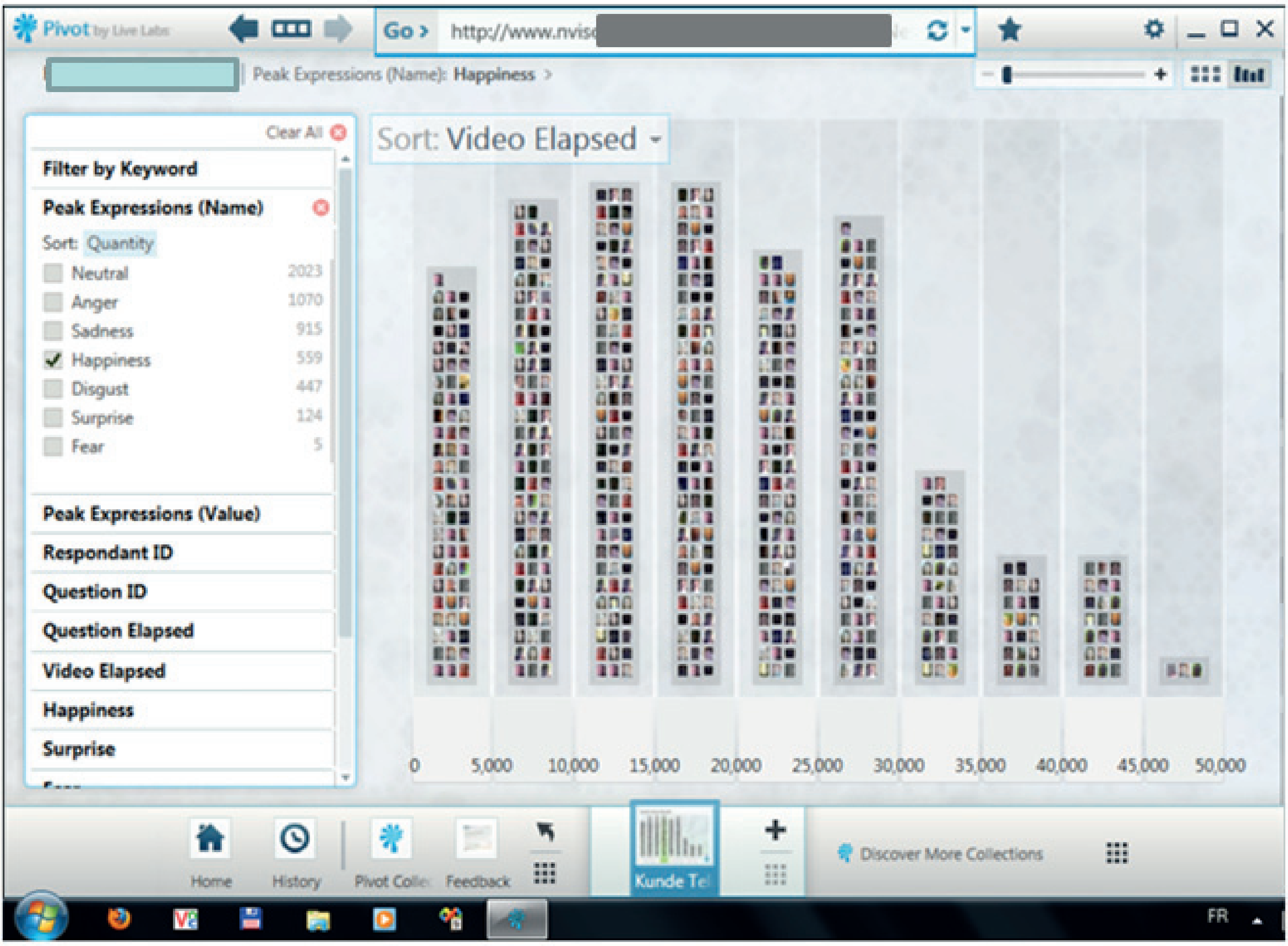
DELIVERY CHALLENGES

Not surprisingly when a sample of 200 respondents is filmed second by second over a typical TV commercial and 143 movements are being monitored and recorded, the processing task is immense.

For 1 minute of video, a set of 200 respondents will generate over 100,000 images to process. However this type of automatic system does not have the intelligence of a human observer and will be prone to error in the presence of certain events and conditions that must be filtered to ensure reliable and robust data. For example, when someone picks up their mobile phone during the interview, or goes to the kitchen to take drink, these images must be discarded from analysis. A data-cleaning step is thus required to remove artefacts from the automatically processed data and this can be performed automatically or manually depending on the desired robustness of the data.

In the automatic case, special algorithms learn how to detect events such as drinking, head not looking at the screen, and talking that can deal with a good proportion of common errors during processing. In scenarios where data robustness is absolutely key, images can be viewed manually and edited. This approach is aided by special visualization tools that allow thousands of images to be viewed in real-time thanks to a special graphical pivot table, greatly reduced the time needed to view and check images. Figure 8 below shows an example of this pivot table, where the data points, as shown in the stacked bars, are in fact images captured of emotions expressed at a given intensity at specific, analyst-selected points in time.

Figure 8: Pivot Table Generated from Captured Images



RAW DATA FORMAT

In the ensuing sections, we demonstrate by two cases studies, examples of the analyses that can be derived from the measurements collected. In the main, these analyses are built up from data collected at the respondent level on a time period basis, measuring the relative level of each emotion's intensity displayed in a given second's exposure to the stimuli under investigation. Figure 9 below shows a typical raw data format for a respondent. The measures are relative, representing the 'proportion of emotion' and sum to 1.0 across each time period (in this case, each second,)

Figure 9: Typical Raw Data Format

ID	Second	Happiness	Surprise	Fear	Disgust	Sadness	Anger	Neutral
082	0	0.035	0.175	0.055	0.055	0.275	0.020	0.385
082	1	0.035	0.195	0.065	0.060	0.250	0.010	0.385
082	2	0.025	0.135	0.050	0.055	0.290	0.010	0.435
082	3	0.030	0.145	0.050	0.055	0.265	0.010	0.445
082	4	0.035	0.165	0.045	0.055	0.255	0.015	0.430
082	5	0.030	0.145	0.045	0.055	0.270	0.010	0.445
082	6	0.035	0.170	0.060	0.055	0.240	0.005	0.435
082	7	0.045	0.155	0.050	0.055	0.260	0.010	0.425
082	8	0.045	0.155	0.045	0.050	0.235	0.005	0.465
082	9	0.025	0.225	0.055	0.045	0.220	0.010	0.420
...	...	...	...	...	...	...	...	...
082	29	0.325	0.195	0.080	0.095	0.035	0.015	0.255
082	30	0.360	0.195	0.065	0.070	0.025	0.015	0.270
082	31	0.210	0.225	0.035	0.160	0.090	0.070	0.210
082	32	0.265	0.215	0.085	0.085	0.040	0.015	0.295

There are a number of ways this data can be utilised. In the simplest form we can present reporting of using aggregates or averages by respondent. However, early work indicates that indexes representing the emotional increases/decreases and emotional lift generated are preferred as they help better identify key inflection points in continuous video stimuli examples below.

To facilitate the indexing, respondents are exposed to a blank blue screen for 2-3 seconds before being shown the stimulus (e.g. TV commercial) under investigation. This provides a base line prior read for each respondent, upon which subsequent changes in emotional response can be indexed.

To illustrate how these measures are turned into reporting outputs, we demonstrate by example using two case studies based on trials conducted in late 2010.

CASE 1 - HEINEKEN 'WALK IN BEER CUPBOARD'

For this case, a sample of 150 respondents, adults aged 18-64, were recruited for interview online and asked, after appropriate instruction, to view a television commercial. During the viewing part of the interview the respondents, with their permission, were filmed. The commercial was the acclaimed Heineken 'Walk in Beer Cupboard' commercial (those interested can find it on HYPERLINK {"http://www.youtube.com/watch?v=wjkUkqTUHNc"}

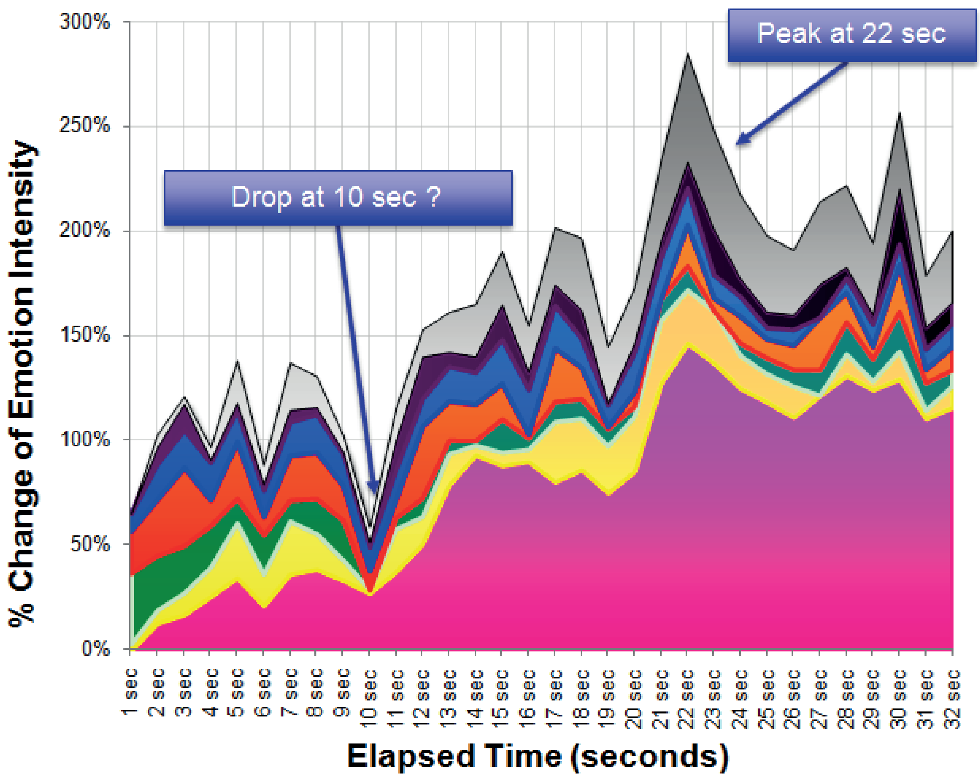
The Facial Imaging software measured the emotional responses, from facial movements via webcam, collecting second by second changes in emotional 'expression' for the 150 respondents.

Figure 10 below shows both how the emotional intensity overall rises and falls from the prior baseline over the 30 second period. It not only shows the overall levels, comparable to arousal/activity measurements via a skin sensor (e.g. via galvanistic skin response), EEG or fMRI but adds more colour by displaying the relative intensity levels of the key emotions measured.

Figure 10: Walk in Beer Cupboard - Emotional Intensity Profile

WHAT EMOTIONS DOES CAMPAIGN EVOKE?

This chart (real data) shows the second-by-second variation in intensity and mix of emotions among respondents watching a TV Commercial



Thus we are able to see what is happening with respect to emotional response, as and when it occurred, on a second by second basis. There is generated quantifiable information on "amount" of emotional change evoked. This commercial turned out to be highly emotional in contrast to others tested which showed lower peaks. In other words, there is a clear guide to build and decline trends for differing emotions; there is a precise read on "tipping points" where response changes emotional direction.

Furthermore, due to the scalability factor implicit in the software application, it becomes potentially feasible to cover much larger samples reasonably cost-effectively. Overall measures of statistical accuracy aside, the key benefit, if applicable, is the ability to look at sub-groups and market segments. Figure 11 below shows a simple gender comparison with clear evidence that women lose interest a short way into the ad as their 'build rate' falters between 5 to 10 seconds, whereas the male build steadily increases up to around 22 seconds. (In this case, this may not present a problem depending on the target, as we can see the male's interest captured and sustained.)

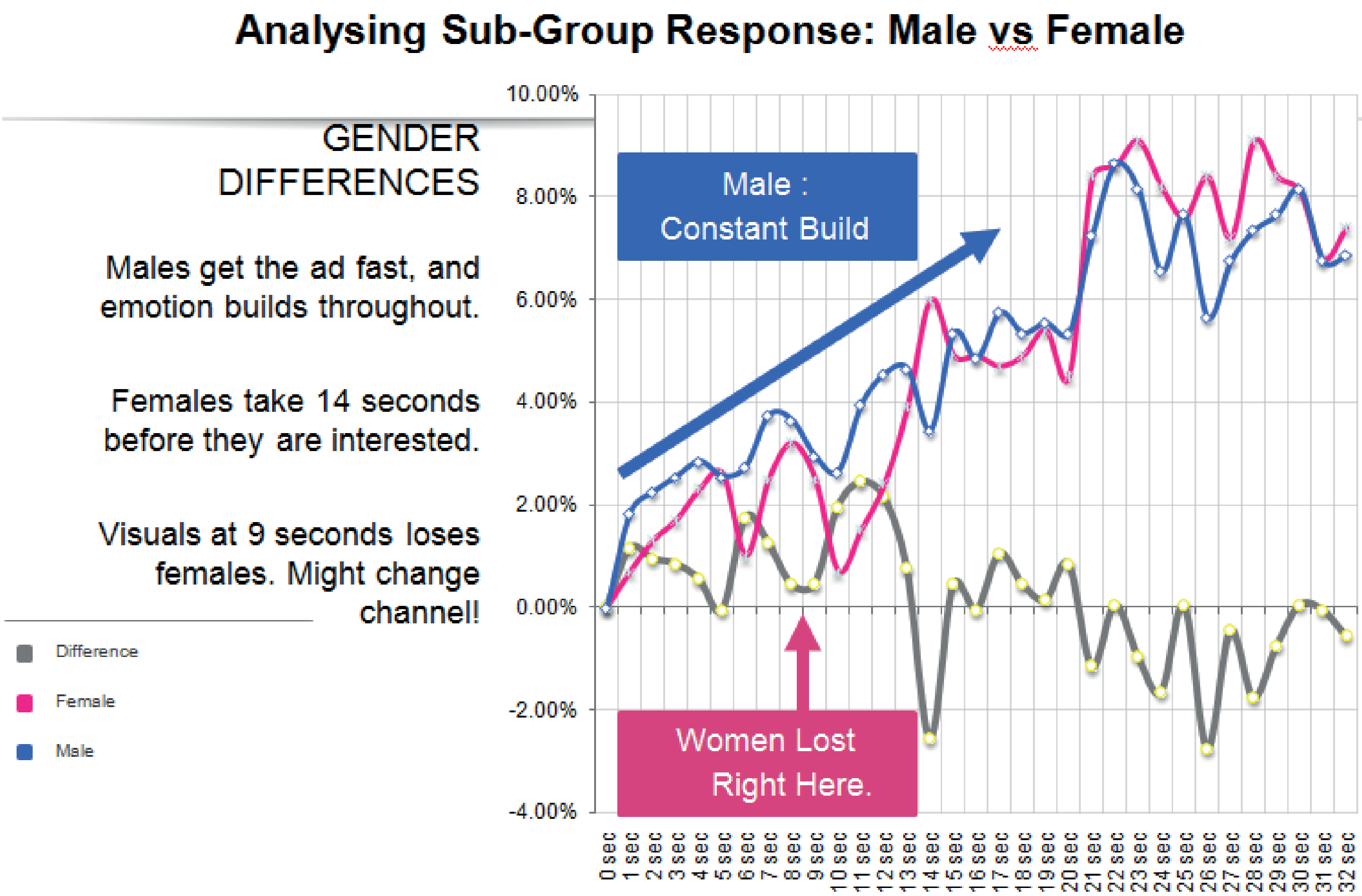


Figure 11: Walk in Beer Cupboard - Gender Comparison

Based on this data, a summary of findings might, depending on objectives, typically show:-

1.This was a good ad, but took time to build in the first 8 seconds. Target consumers could be lost (channel switching).

2.The exact scene where a sub-group (in this case, females) 'turned off' was identified. If this was important for targeting the issue can be addressed specifically and precisely.

3.If there was a need to create a 15sec version we could help decide what was most important to retain.

Furthermore as an aside, taking a global outlook, it is worth noting that this commercial has been run in several countries in each country's respective language. Had international comparisons/re-assurances been needed, it is not inconceivable to conduct the entire test without verbal/written questioning in all key markets to ensure consistency of response/appreciation and so forth.

CASE 2 - HSBC 'DOLPHIN'

For this case, a sample of 150 respondents, adults aged 18-64, were recruited for interview online and asked, after appropriate instruction, to view a television commercial. During the viewing part of the interview, the respondents, with their permission, were filmed. The commercial was the HSBC 'Dolphin' commercial (those interested can find it on {HYPERLINK http://www.youtube.com/watch?v=uml27Mr1IhA.}

This is a very different commercial to the upbeat and amusing Heineken 'Walk in' ad. It starts off rather bleakly before switching mood, introducing some 'drama', and then signs off with the bank logo. Figure 12 shows the emotional build trend for 3 of the relevant emotions, and noticeably the first third of the commercial evokes the sadness the scene-setting would suggest.

INTENSITY TREND

This commercial evokes a different range of emotions than Heineken – starting with sadness and surprise.

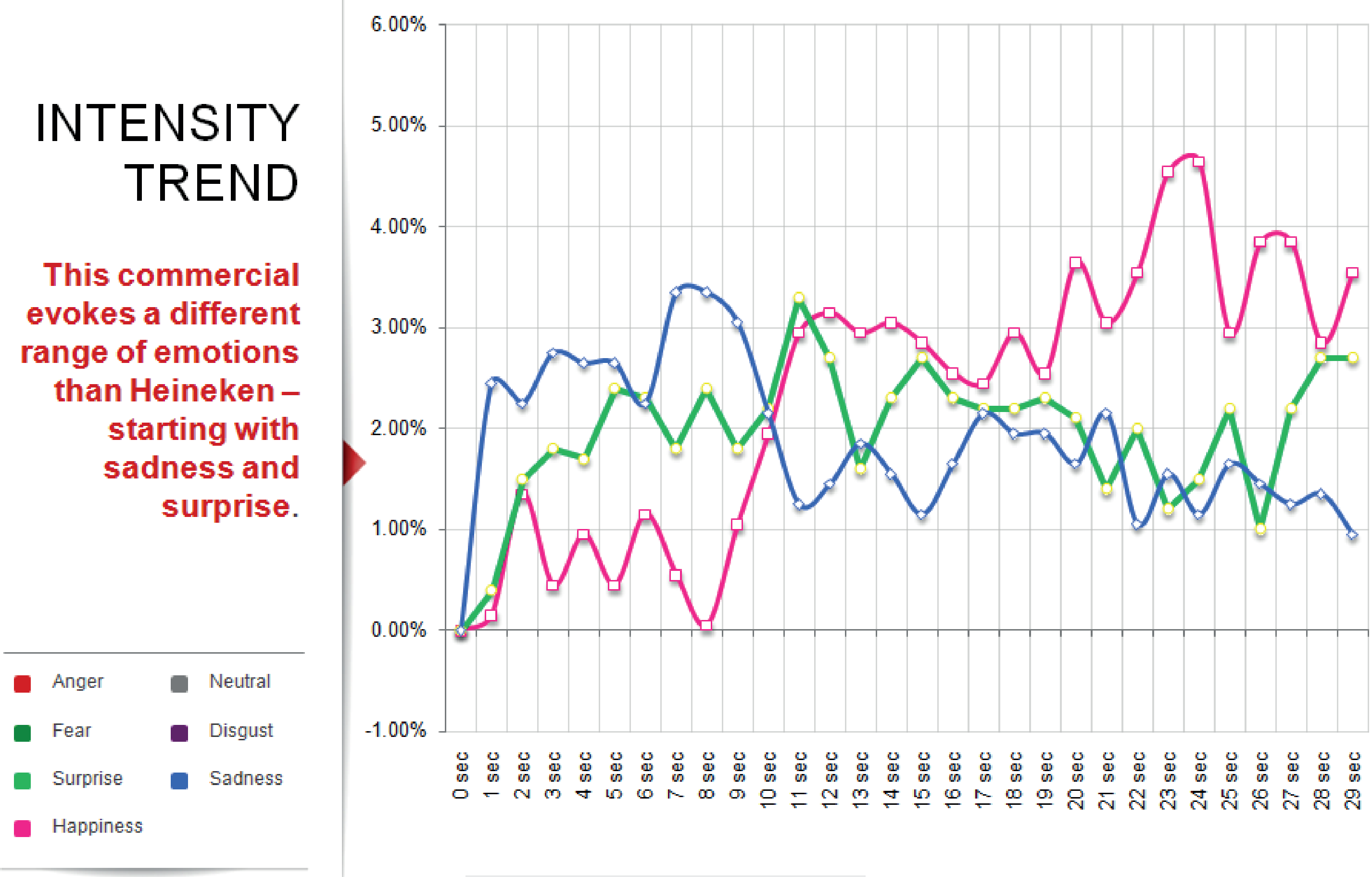


Figure 12: Dolphin - Emotional Intensity Trend

How this increases consumer understanding to aid marketing and advertising decisions, is illustrated, in one way, by looking at the happiness emotional response and aligning it by frames.

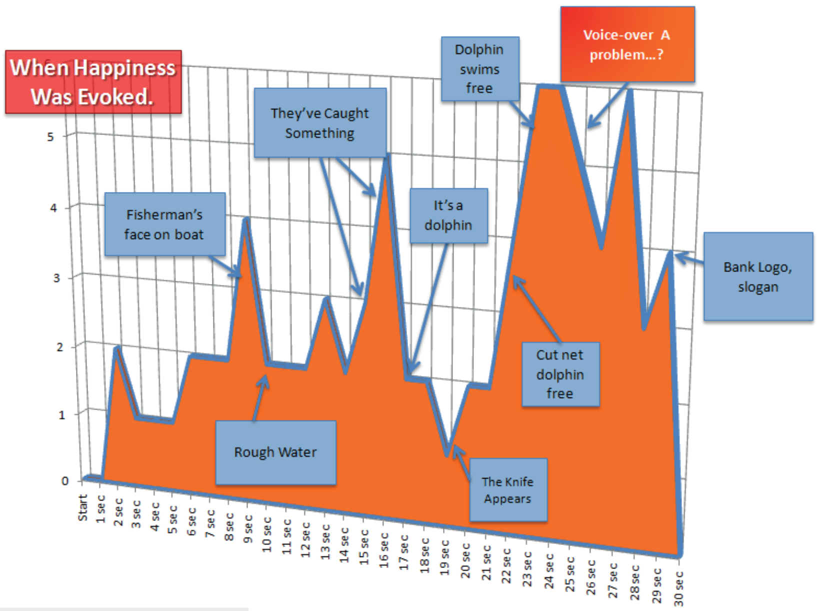


Figure 13 - Dolphin - Frame Analysis

Again, facial imaging software allows us to measure precisely what and how the elements in the stimulus, in this case the commercial, are contributing to its impact. In this example, which showed scenes of a fishing boat accidentally capturing a dolphin, we are able to see which visuals evoked a strong "happiness response" and check that people were demonstrating the intended feelings (Figure 13 above).

More importantly, we can see if some elements of the ad are causing unexpected reactions. In this case, when the voice-over came on we lost positive emotion – meaning people may have turned off before seeing the key branded message.

INTRODUCTION

First, it is recognised that despite some very encouraging trials and pilot tests, more experimentation and testing is desirable. Nevertheless, to leverage the strengths of Facial Imaging in enhancing the market research process, and before entering into specifics, we need to recognise that this technology counters some of the key criticisms increasingly being raised as arguments against modern market research, namely:

1. Conventional market research (MR) techniques are predicated on rational models of information gathering and we force respondents to engage their 'rational' brain.

2. MR methods are too slow and insensitive when it comes to emotional research. We too often rely on recall and do not capture fleeting emotional responses in real-time.

3. The need to improve recording nuances and depth of emotional response when respondents cannot describe subtle, often unconscious, reactions.

In bypassing the problems of self-reported emotions by relying on passive methods to collect emotional response, we can overcome these challenges and contribute greatly to the development of market research. The methodology can be applied to a wide range of marketing research issues as briefly outlined below.

EMOTIONAL RESPONSE TO ADVERTISING, LOGOS, & RELATED BRANDING

As has already been illustrated in the two case studies above, Facial Imaging is readily applied to advertising pre-testing when addressing the emotional communication from and response to audio-video stimulus. It can even be argued that such application can completely remove the need for any direct questioning on the advertising in many cases. However, in reality, the likelihood of users moving away from stock questions, and their attached norm banks, is slim. Thus, in reality, it is likely facial imaging will be used to enhance the emotional measurement aspects in these applications.

Notwithstanding the fact that the moments in time during a TV commercial when the brand's logo, tag-line and other 'livery' are displayed are precisely identified on the emotional timeline, Facial Imaging can help better understand the impact of a given brand when it is applied to a specific marketing initiative, whether it is advertising, sponsorship, endorsement or the like. If we believe that one aspect of brand equity, as defined by Kevin Keller [16], is what we know about a brand impacting on how we react to the marketing of that brand, then the technology can help select with which activities a brand should be associated.

CONCEPT SCREENING AND PRODUCT/SENSORY TESTS

Given the nature of how consumers are exposed to stimuli in concept and product testing, there is more work needed to determine best how to exploit the features of Facial Imaging in their evaluation. Experimentation into this application is ongoing and is in the field at the time of writing.

Logistically, concept testing provides the easier challenge. A respondent can be asked to view a finished or rough format concept, in a still position for a given period of time, much in the same way that they would be exposed to a TV commercial. However, unlike the commercial where the 'story' unfolds over a linear timeline, the concept is generally shown all at once. Thus, work is underway on unfolding the display of the concept to develop a time line that will best suit the overall explanation and display of the concept in a robust, consistent, and repeatable fashion. This involves, in the main, breaking down the concept into its visual elements, essentially the photo/drawing and the components of the text, e.g. headline, description of features, benefit, reason to purchase, price/quantity etc. Then those elements are exposed or unwrapped over time much in the same way builds on bullet points in a PowerPoint presentation are revealed. Even at this stage, comparisons are being made to assess whether the entrance of a text element (wipe right, box out, etc) is relevant and if relevant which is most appropriate.

Slightly more complex, but similar in approach to concept testing design, is that of packaging testing. Current designs, under experimentation, involve exposing the packaging in question at different angles and aspects for a fixed period of time. So, for example, respondents may be exposed to a computer graphic the front of the pack for 5 seconds, then to a 'rotating' all round view of the pack for the next 5 seconds, and finally the back of the pack for 5 seconds. As with the concept testing, some form of time-based delineation is needed to ascribe the emotional responses to a particular part or element of the stimuli. Clearly, this will vary considerably by category, format, distribution channel and so forth. In particular, though we feel this could be a very strong approach when paired with eye-tracking research that is often used to assess shelf impact of new versus existing pack designs and formats.

The whole area of product testing and sensory testing is more challenging logistically. At present, facial imaging relies on a reasonably stable facial image being presented with around 30° rotation left or right of centre allowed. In a product testing situation, not only would there be more movement likely but clearly, as far as food products are concerned, there would be facial muscles used in the act of consumption, which needs to be avoided in the uses developed so far. One option under consideration is to allow the respondent to consume the product, then to show short film of the consumption act and monitor the consumer response to that film. This could possibly be achieved by showing the film first, and then introducing the product, then showing the film again, and any differences on a respondent to respondent basis could be inferred by their respective 'shift scores'.

In the case of sensory testing, we have to be aware that, in the main, these exercises are executed in heavily controlled circumstances anyway, designed to prevent any contamination of the process. Whereas, we can see the same challenges on taste-testing as presented by general product testing, we feel that in a controlled environment, filming a consumer responding to various fragrances, scents, and smells presents fewer obstacles. This application will be further explored experimentally in early 2012.

PRODUCT LAUNCH PREDICTION/ SIMULATED TEST MARKETING

The majority of the major simulated test marketing systems undertaken today (e.g. the Novaction and BASES suites, and so forth) contain at least concept evaluation component, usually a product test element and in some cases the advertising concept and other marketing elements. So, clearly, a facial imaging component, if properly applied will likely enhance the protocols. Furthermore, the non-intrusive nature of the software application means it can be introduced as a component of many of these tests without disrupting the process and thus potentially altering the protocols that have been applied over decades to building up the norm-bases on which the predictions are based. In fact, given that companies like BASES were able move from conducting their research in a predominantly mall or door-to-door based environment, at least at the concept stage, to a mainly online one without any significant technical or trend disruption, the addition of a facial imaging component to should be methodically neutral.

WEBSITE AND ONLINE ADVERTISING IMPACT

It is probably safe to say if the technology can be applied to conventional TV commercial pre-testing via an online data capture platform, then anything related to website evaluation and online advertising logically follows on. In fact, in these latter cases emotional responses are collected real-time in about as realistic situation as had ever been attained in the evaluation of any advertising.

Furthermore, we see the opportunity here, maybe removing the 'reality' a bit, for combining two of the new(er) technologies - eye-tracking and Facial Imaging. To a degree websites, e.g. Westpac, Sydney Morning Herald, Qantas etc are relatively static but comprise numerous features so it is useful to know where the respondent is looking/ reading in order to assess and evaluate the emotional responses.

TRACKING ADVERTISING WEAR-OUT

Even the all-time classic ads pall after consistent viewing over time. But in order to help maximise returns not only on the media investment but on the production costs of development, some measure of wear-out would be useful. There may be a time when a TV commercial ceases to deliver significant impact, when it becomes 'part of the furniture', due to continuous exposure, or perhaps just to innovative advertising from competitors or other segments. The objective of monitoring wear out would be to predict or capture the 'tipping point' beyond which ROI on media investment fell below the desired level.

An approach under consideration is to create a panel of respondents each of whom are shown a variety of commercials on a regular, say monthly, basis and to monitor their arousal levels and response profiles. The key measures will be the change in the degree to which the ads raise arousal and provoke response over and above a neutral baseline setting as time elapses. It is expected that due to familiarity certain emotions, e.g. surprise, will diminish faster than others (e.g. happiness, which could be as much linked to a well-liked brand as to its particular advertising.)

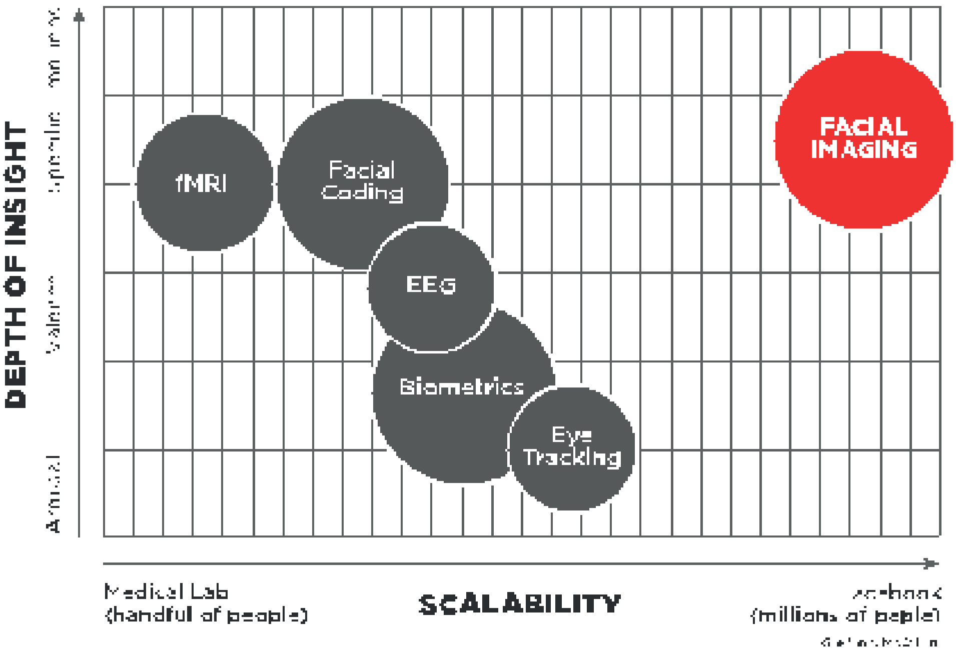
In assessing any decay in response, with ongoing sales levels, trial and repeat purchase levels and so forth, metrics for decay, indicating when a fresh campaign or execution should be considered, can be developed over time.

SCALABILITY

One of the main hurdles in augmenting the existing array of market research tools with the benefits of neuro/biometric applications has been essentially logistical. In the main the equipment needed for many applications is higher cost than that for conventional research, not always highly portable, and often requires specialists or highly trained technicians to operate. Even manual facial coding which has less technology attached can only work when responses are interpreted by dedicated individuals with years of training and experience.

Having said that, recent developments have reduced the size of the equipment and, especially in the area of eye-tracking, brought down the costs involved. Nevertheless, ignoring any technical, methodological, or academic advantages, web-cam based Facial Imaging has a marked advantage in being more scalable both logistically and in terms of cost/investment. When this is coupled with the fact that as of the time of writing, 90+% of new laptops and notebooks are already fitted with web-cams, and these are becoming standard is most smart phones, tablets, and other devices, it is clear that the ability to augment or introduce facial imaging as part of the research design is already practical and cost-effective.

Figure 14: Comparison of Scalability Potential amongst Neuro/biometric Methods



POTENTIAL CONTRIBUTION IN APPLYING FACIAL IMAGING

In essence, Facial Imaging brings four important logistical and methodological benefits to the conduct of research both operationally an in application. In summary these are

Scalability – Facial Imaging can be used with large or small samples more readily and generally more cost-effectively than currently existing neuro/biometric methods, and so offer the benefits of higher statistical accuracy and the ability to analyse more in-depth.

Non-Intrusiveness – The respondent does not need to do anything except watch, so the method, per se, has less impact on the nature of response

International Applicability – Facial Imaging markedly reduces the challenges of translation issues and reduces problems associated with comparing scales across cultures

Adaptability – Relatively easy to integrate with conventional studies, as well as other neuro/biometric approaches across a range of applications, online or offline

LOOKING AHEAD

There are also broader, longer term issues to the industry as a whole that facial imaging could help address. A recent article [17] in ESOMAR's Research World Connect online magazine asked "What if we could conduct surveys without asking any questions, what if we could directly assess 'feelings' without intricate scales or complex biometric or neuroscience equipment?" New technologies, such as Facial Imaging, offer up exactly that prospect, and – utilised cleverly – they might just be what is needed to ensure conventional market research is not marginalised but can maintain a role in the centre of the marketing decision-making process.

Despite many challenges, questionnaires remain the bedrock of most quantitative market research. While social media research, MROC's, neuroscience tests, and a host of others all offer useful alternatives, sometimes asking a sample of people is simply the easiest, most cost-effective and most reliable way to tap into consumers' mind-sets. Yet response rates are declining and, as marketing focuses more and more on the emotional, the problem of getting 'accurate' response to survey questions becomes more acute. Surveys are often too long, a bit dull, and too lacking in the subtlety necessary to uncover hidden, non-rational response. Most researchers are very aware of these issues and, as a result, debate on methods of emotional measurement and ways of making surveys more engaging rage throughout the industry.

There are three key considerations

1 .There are simply quite a few limits on how engaging a survey can actually be. Making it really "fun" takes up a lot of creative effort and is seldom fully successful. Design effort might often be better spent on keeping a survey short and simple, rather than on inventing 'engaging' approaches.

2. There is no foolproof way of getting people to read words, make choices between images or move objects around a screen that does not involve heavy use of their frontal/pre-frontal cortex. While some survey procedures can be shown to be better than others at evoking subconscious reaction, none are really direct measures of emotional response. They are, in one way or another, all implicit measures that are subject to conscious filtering and cannot capture direct emotional response.

3. If we can embed passive collection measures that let us collect hard data on the soft emotions, we can focus more of our surveys on getting more accurate information on what people are actually doing in their day-to-day lives. Surveys can be shorter, easier to answer and yet at the same time can start to provide much clearer answers to key questions on the real links between emotions and behaviour.

Without wanting to be unnecessarily alarmist, it may be worth considering our role as marketing advisors if we do not attempt to adopt some of these new technologies. As a simple example, it is already possible to take a video of a commercial or visual concept from any marketing company, show it online in 10 countries without needing to worry about questionnaire design, translation or coding and then automatically analyse the results and generate a fast report comparing emotional engagement, types of emotions aroused and which images evoke most emotion. There are several other technology companies around the world, outside the market research sphere, working to make various kinds of direct, scientifically objective measurement of emotion more accessible and useable. Whatever we as researchers may think, this kind of data-collection and analysis is going to become increasingly popular, and survey research will have to adapt.

This could perhaps frighten some market researchers. After all, if an advertising agency or end-user can take their marketing material to a technology company and get back the kinds of outputs described it does challenge the role of research in the field of consumer emotions. While it is perfectly possible to get a huge amount of information about a commercial, packaging or product concept by “videoing faces”, you'll get a lot more by embedding the Facial Imaging into a wider online survey and analysing it in-depth. The opportunity inherent in comparing more accurately assessed emotional reaction to brands and marketing stimuli with behavioural and demographic information is huge.

Essentially all this means that Facial Imaging holds up the promise of solving some fundamental issues with collection of emotional response, for instance that it can:

1.Overcome the issues of “rationalised” or superficial response to survey questions

2. Deliver consistent response across cultures and ethnicities

3. .Work consistently across all sorts of sectors, subjects and media types with little modification.

So, while it is felt that the application of Facial Imaging software can help assess the emotional content of a commercial or movie, this may represent only the tip of the iceberg in the use of such methods. The potential to provide a better way to integrate and enhance survey data is, perhaps, the real reason that approaches like Facial Imaging are important in research – they represent tools that potentially take us a lot closer to that holy grail of a synthesised, all-embracing explanation of why people buy products / services and why some ads engage and others do not. And that, more than anything else, is what we, as researchers, are supposed to be doing. Rather than spelling the end for the survey then, scientific approaches like Facial Imaging could provide just the kind of support that survey research requires in order to remain relevant in the new age of marketing.

Commercials used in the Case Studies:

Heineken - Walk in Beer Cupboard - http://www.youtube.com/watch?v=wjkUkqTUHNc

HSBC - Dolphin, http://www.youtube.com/watch?v=uml27Mr1IhA

[1] C. Darwin, The Expression of the Emotions in Man and Animals. J. Murray, London. (1872)

[2] D. Matsumoto, B Willingham, The thrill of victory and the agony of defeat: Spontaneous expressions of medal winners of the 2004 Athens Olympic Games. Journal of Personality and Social Psychology.(2006) 91(3). 568-581

[3] J. F. Cohn, Foundations of human computing: facial expression and emotion, in: ICMI '06: Proceedings of the 8th international conference on Multimodal interfaces, ACM, New York, NY, USA, (2006), pp. 233–238.

http://dx.doi.org/http://doi.acm.org/10.1145/1180995.1181043 doi:http://doi.acm.org/10.1145/1180995.1181043.

[4] P. Ekman, An argument for basic emotions, Cognition & Emotion 6 (3) (1992) 169–200.

[5] P. Keltner, D. Ekman, Facial expression of emotion, in: Handbooks of emotions, M.Lewis & J.M.Havilland, (2000), pp. 236–249.

[6] P. Ekman, W. V. Friesen, Facial Action Coding System Investigator's Guide, Consulting Psycologist Press, Palo Alto, CA, (1978.)

[7] P. Ekman, W. Friesen, J. Hager, Facial action coding system, Research Nexus, Network Research Information, Salt Lake City,UT, (2002.)

[8] Y.-L. Tian, T. Kanade, J. Cohn, Facial expression analysis, in: S. L. . A. Jain (Ed.), Handbook of face recognition, Springer, (2003.)

[9] P. M., M. Bartlett, Machine analysis of facial expressions, in: K. D. . M. Grgic (Ed.), Face Recognition, Vienna, Austria: I-Tech Education and Publishing,(2007), Ch. 20, pp. 377–416.

[10] B. Fasel, J. Luettin, Automatic facial expression analysis: A survey, Pattern Recognition 36 (1) (2003) 259–275.

[11] M Sorci, G. Antonini, J.-P. Thiran, M. Bierlaire, Facial Expressions Evaluation Survey, iTS (2007).

[12] M. J. Farah, K. D. Wilson, M. Drain, J. N. Tanaka, What is special" about face perception?, Psychol Rev 105 (3) (1998) 482–498.

[13] A. Schwaninger, J. S. Lobmaier, S. M. Collishaw, Role of featural and configural information in familiar and unfamiliar face recognition, in: BMCV '02: Proceedings of the Second International Workshop on Biologically Motivated Computer Vision, Springer-Verlag, London, UK, (2002), pp. 643–650.

[14] R. Cabeza, T. Kato, Features are also important: Contributions of featural and configural processing to face recognition, Psychological Science 11 (2000) 429–433.

[15] M. Meulders, P. D. Boeck, I. V. Mechelen, A. Gelman, Probabilistic feature analysis of facial perception of emotions, Journal Of The Royal Statistical Society Series C 54 (4) (2005) 781–793.

[16] K L Keller, Strategic Brand Management, Prentice-Hall, 1998

[17] A Gordon, The Face as the Ultimate Survey Tool, Research World Connect June (2011)