

# Does Facial Coding Generalize Across Cultures? A Spotlight on Asia

Rana el Kaliouby, Evan Kodra, Pankaj Jha, Namita Mediratta

## **Abstract**

Affdex emotion-sensing technology analyzes facial expressions to discern consumers' emotional reactions. Millward Brown and Affectiva partnered to capture emotions at scale, making facial coding an integral part of Millward Brown's advertising pretesting worldwide. We discuss the challenges we faced as we rolled this out globally, and present groundbreaking emotion-sensing algorithms and methodological best practices that address cultural influences, especially in Asia. This work led Unilever, the innovative global brand, to incorporate Affdex worldwide to gain novel insights into ad engagement.

## Unilever's Challenge: A scalable, global and scientifically valid measure of emotion engagement to Ads

Like many Fortune 500 brands, Unilever invests millions in advertising that aims to capture viewers' attention, build brand loyalty, and ultimately drive sales. Key success criteria include the extent ads are emotionally engaging, delivering the desired message, driving branded cut through and persuading the consumer to take action.

To meet this need, marketers increasingly seek ways to measure emotional engagement to augment self-report data and gain insights into ads' emotional impact. This requires technology to capture consumer's natural, unfiltered emotions as consumers interact with content, and to do so globally and at scale. An innovative company, Unilever has embraced facial coding technology from Affectiva to assess viewers' implicit emotion responses to its advertising testing worldwide.

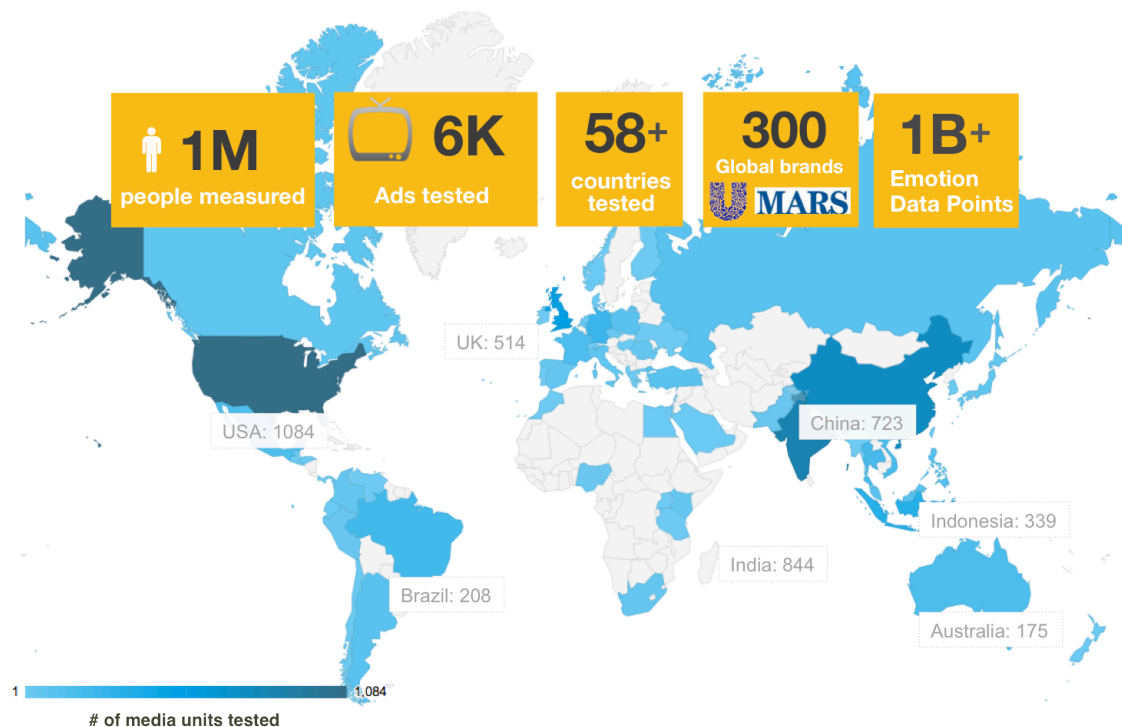


Figure 1: Map showing volume of Afdex facial coding tests by region. The darker regions indicate the markets with the most adoption. To date, we have captured over 1 million facial responses for more than 6000 ad tests in 58 countries for 300 global brands.

This paper recounts the journey that Millward Brown and Affectiva undertook to adopt, scale and operationalize Afdex emotion-sensing technology as an integral part of Link tests, and the insights gained from having tested nearly 2000 Unilever ads worldwide in 2013.

We share the top challenges we faced as we scaled globally and discuss our key learnings, with the spotlight on Asia. We present case studies in Asia including several with Unilever brands in China, India and Indonesia— the largest and fastest-growing adopters of facial coding in Asia. The case studies highlight the value of facial coding in this market, achievable only after addressing the particulars of this market in regard to cultural norms.

## Background

Facial coding has seen increased adoption in marketing research to supplement traditional verbal report measures. Facial expressions have been shown to predict variables related to advertising success, with facial responses correlated with recall (Hazlett & Hazlett, 1999), ad skipping and zapping (Teixeira et al., 2014) and sales performance (McDuff et al., 2013). Affdex facial coding is particularly well positioned to offer content creators and publishers with new insights on the role of emotions in media content testing, personalization and optimization (McDuff et al., 2014). The universality of facial expressions combine with the ubiquity of cameras to make Affdex facial coding an easily understood, technologically accessible, scalable ((McDuff et al., 2012), and unobtrusive means to assess viewers' emotional reactions to media. Figure 1 shows the extent to which Affdex facial coding has been adopted worldwide.

Understanding consumers' emotion response to ads is especially challenging in markets like Asia with its cultural emphasis on conformity and harmony (Tsai et al., 2002). Asian consumers are less likely to express negative feelings about an ad and are more likely to verbally report a positive response, as a result reducing the spread of survey responses. Controversial or provocative ads are therefore less likely to “test” well in Asian markets. Thus, robust technology and study methodologies that enable the capture of consumers' subconscious or visceral emotion responses to content are crucial for this region.

### **Unilever embraces facial coding**

In January 2012, Millward Brown incorporated Affectiva's emotion-sensing technology, Affdex, in its ad testing flagship product - Link. Today, facial coding is incorporated in all Millward Brown Link copy testing worldwide, combining a comprehensive set of evaluative and diagnostic questions with spontaneous and visceral emotion responses via the face.

Unilever was one of the first global clients to adopt facial coding. In January 2013, Unilever paved the way and began to incorporate Affdex with every Link test they undertook. Through 2013, 1855 Unilever ads were tested with Affdex facial coding in 60 countries, for over 175 Unilever brands. Each Unilever ad is norms-tested for emotion assessment against a repository of more than 6000

Affdex facial coding ad tests conducted for clients circling the globe, with more than half in Asia alone.

### **The top three challenges: cultural norms, systems and scale, and global rollout**

In 2014, Affectiva will perform more than 12,000 facial coding tests across its client base—over 400% year over year growth. While we’re excited by this rapid adoption, it did reveal several complex and extensive challenges that we addressed early on.

Our first challenge arose from cultural differences in how people express emotion, which led us to create emotion norms by geographic region. In the following section, we discuss those norms with a spotlight on Asian markets.

Our second challenge was around operations and scale: the Affdex platform requires 100% uptime in the face of fast-moving scaling and fast-changing functional requirements, with “follow the sun” study provisioning, fielding, and reporting requirements at its core.

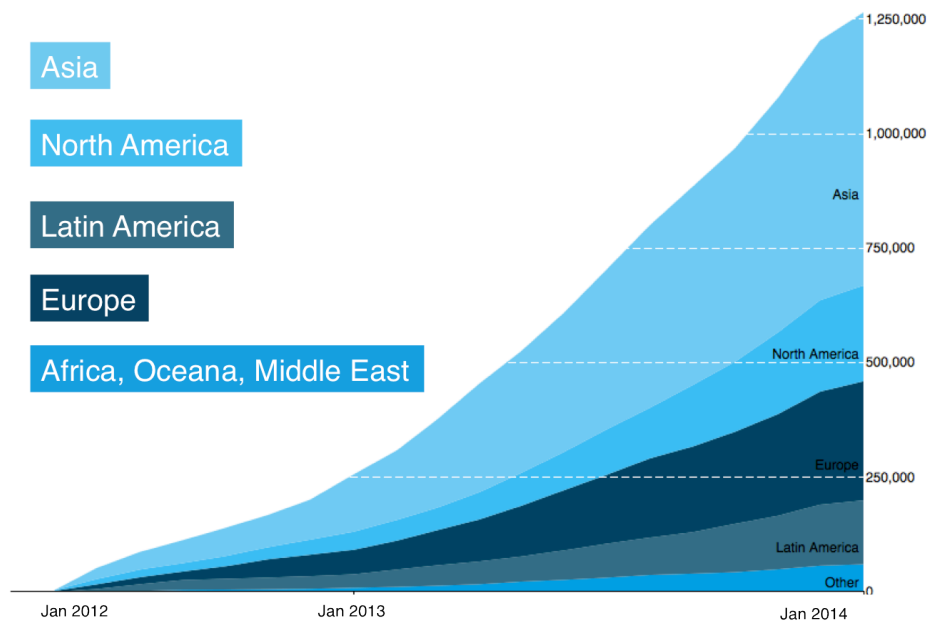


Figure 2: As of January 2014, Affectiva has amassed over 1 million facial videos from over 6000 ad tests. A large proportion of these are from Asian markets. Facial expressions are captured online (“in the wild”) or, especially in Asia, from central locations. Achieving worldwide scale presented operational challenges but allowed us to build unique global emotion norms by region.

Our third challenge was around training and quality assurance: we’ll discuss how Affectiva worked in close coordination with Millward Brown’s nearly 1500 neuroscience, client services, and project team members to develop best practices and to meet or exceed customer requirements.

## 1st Challenge: Cultural and Market Norms

### Facial Expressions – Universal yet culturally specific!

When it comes to emotion communication, we all start with the same base of universal, pan-cultural facial expressions. In his seminal book on *The Expression of the Emotions in Man and Animals* published in 1872, Charles Darwin theorized that emotions are biologically determined and universal. Since then, many scientific studies demonstrated the universality of facial expressions like happiness, sadness, surprise, disgust and more (Ekman, 1971; Izard, 1994; Jack et al., 2012; Russell, 1994; Tsai et al., 2002).

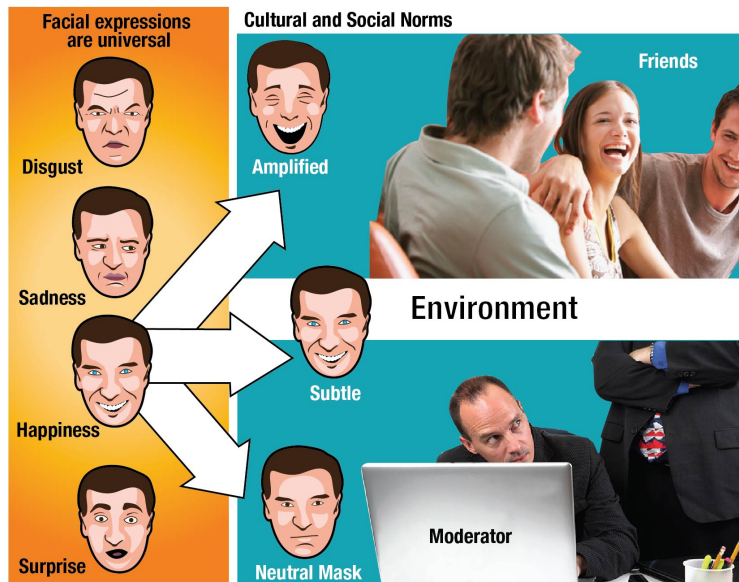


Figure 3: Consumers' emotional reactions vary by culture and by social circumstance. To be viable, emotion-sensing technology must normatively address *cultural display rules* and other factors affecting consumer emotional responsiveness.

However, while emotion expression is universal, we learn at an early age to alter our expressions based on social circumstance. Ekman, et al., (1969) coined the term *cultural display rules* to describe the conditions where individuals manage and modify their emotional expressions depending on cultural and social norms (Figure 3). Individuals may amplify (exaggerate) or de-amplify (minimize) their expressions. For instance, sadness may be amplified at funerals and de-amplified at weddings. One may also repress one's emotions when it is socially inappropriate to share one's true feelings.

Cultural display rules vary by culture, and their effect has been observed repeatedly in studies of spontaneous expressive behaviors. Matsumo et al. (1989) show that Japanese and Koreans are more likely to neutralize expressions, that Russians are more likely to mask negatives expressions with smiles, and that Americans tend not to alter their emotions. *Importantly, East*

*Asians are more likely to mask or dampen expressions in the presence of others, especially moderators.*

## Methodology matters: Online versus central location testing

The nature of global rollouts requires the accommodation of regional differences in technology infrastructure as well. Primarily due to limited Internet connectivity in emerging countries, Link studies take place in central locations, not online. Affectiva developed its Affdex Anywhere solution for this reason, expanding Affdex capabilities to offline scenarios. Affectiva now conducts about 60% of its facial coding studies in these central locations.

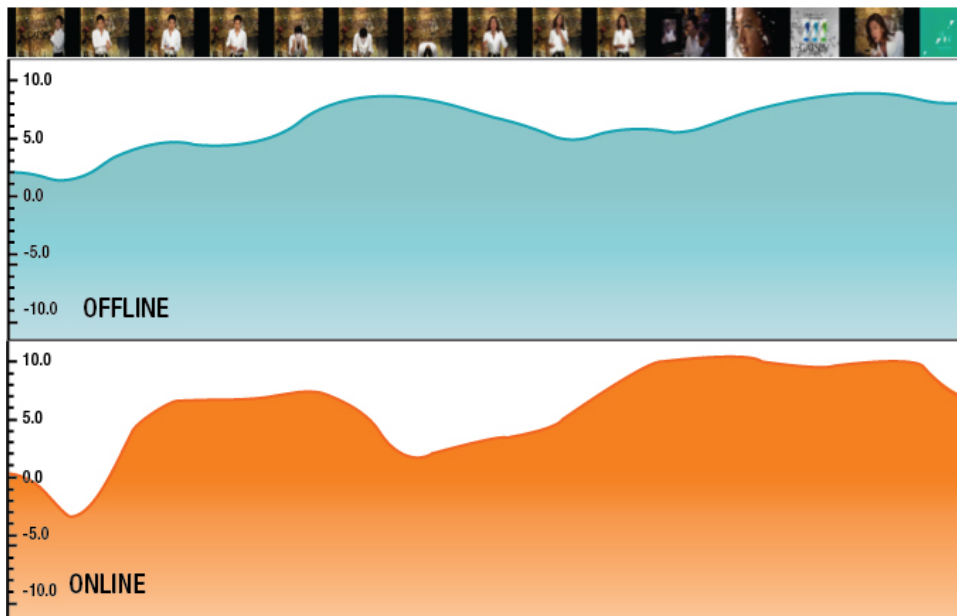


Figure 4: Aggregate “valence” response of viewers to a 30 second commercial in Japan. Valence is a moment-by-moment measure of how positive or negative an emotion response is. Top chart: viewers watching an ad in a central location. Bottom chart: viewers watching an ad online from their homes. Note how the dynamics of the emotion traces are generally the same but that there is more variation in the online case. Also note the negative “dip” in valence in the first few seconds of the ad in the online case, which is more pronounced than in the offline cell.

The variability of cultural display rules and testing methodologies led us systematically explore the differences, if any, between online and central location collection methodologies and its implications on consumers’ propensity to express emotions. Affectiva and Millward Brown conducted a joint study in Japan in 2013 to understand these factors. A set of 12 ads were tested online and offline in a central location. We selected ads that ranged from high to low engagement. Figure 4 shows that in general the responses follow a consistent pattern for online and offline methodologies – but there is increased dynamics in the online case, especially if this involves negative expressions of emotion.

Gaining a complete understanding of the subtle differences in emotion responses arising from online versus central location testing has allowed us to factor them into regional normative assessments.

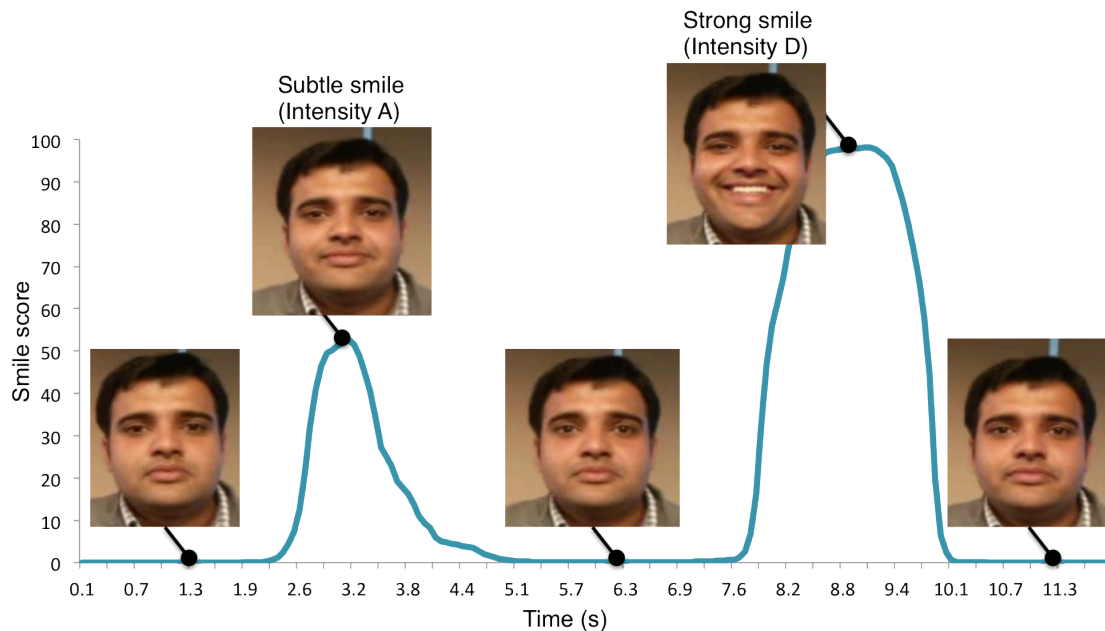


Figure 5: Improvements in our latest smile detector capture the full range of subtle, intensity A smiles, as well as capturing strong Intensity D/E smiles.

## Subtle expressions in Asia

Our experience in Asian markets, including the systematic study we conducted in Japan, confirms studies from Gordon et al., 2012; Gordon, Wheller, Giang, & Yeung, 2013) that consumers in Asian markets, especially when in the presence of a moderator, may dampen their facial expressions. Our challenge was to leverage the data we had accumulated from thousands of ad tests in order to enhance and refine algorithms to capture these subtle and fleeting expressions.

Leveraging facial responses collected in Asian markets, we trained and validated emotion classifiers that capture subtle expressions such as a low intensity smile. The classifiers (1) employ full-face algorithms to detect small movements that occur in multiple facial locations to identify and confirm subtle expressions, e.g., subtle upturns at the mouth and crinkles at the eyes; (2) minimize noise by accounting for natural dynamics as facial expressions unfold as well as distortions due to suboptimal head poses and camera angles; (3) consider other expressions that can trigger false alarms for the expression of interest, e.g., speech or lip purses.

To assess the impact of the new subtle smile detector on our data, we re-ran our existing studies in Asia with the new classifier. We found that there was a



doubling of the number of smiles captured. At the same time, the classifier led to fewer false alarms—thus seeking and capturing subtler signals did not cause more noise to be generated. The result is a significant boost to accuracy and precision in the aggregate, which in turn permits analysis teams to more easily glean additional emotion insights.

Smiles are not the only facial expressions of interest to advertisers and media testers. Our work to improve the smile detector has led to the development of a closed loop, automated framework to replicate improvements made in *Smile* to other expression detectors pertinent to media content testing. The system boosts accuracy by reducing false alarms which, as we've seen, allows detectors to be more aggressive in finding subtle, nuanced expressions. Because it is automated, this process permits for continuous improvements and as such has become an integral part of the Affdex training and validation infrastructure.

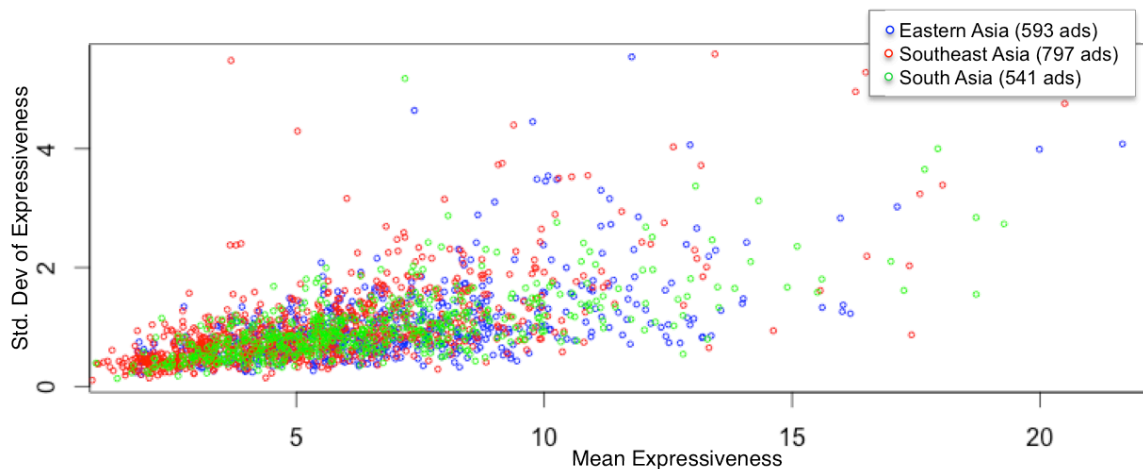


Figure 6: Scatterplot of emotion expressiveness in three regions of Asia. Each colored dot represents a facial coding ad test. This data forms the basis of our regional emotion norms, which then allows us to assess where in the space of emotion each ad falls.

### Regional and market norms

Recall that while emotions are universal, cultural display rules will result in different propensities to express emotion. Figure 6 shows a scatterplot of emotion expressiveness for a total of 1931 ads in three regions of Asia. Eastern Asia includes China, Hong Kong and Taiwan; Southeast Asia includes Japan, Korea Indonesia; South Asia includes India, Pakistan and others. With these benchmarks, we are able understand the emotional profiles among consumers in different markets and place an ad in context to all other ads tested in that region.

We require a minimum of 50 ads in a market before we isolate it and start generating norms specific to it; markets with less than 50 ads are normatively associated with the worldwide norm.



## Emotions vary by product category

Our norms database can also be grouped by product category. In a joint study with MARS Inc., we tested over 200 ads in four categories (chocolate, rice, pet food and chewing gum) and across four markets (USA, UK, Germany and France). As shown in Figure 7, we found that chocolate ads tend to use more humor and thereby elicit more positive emotions than say, food (rice) ads.

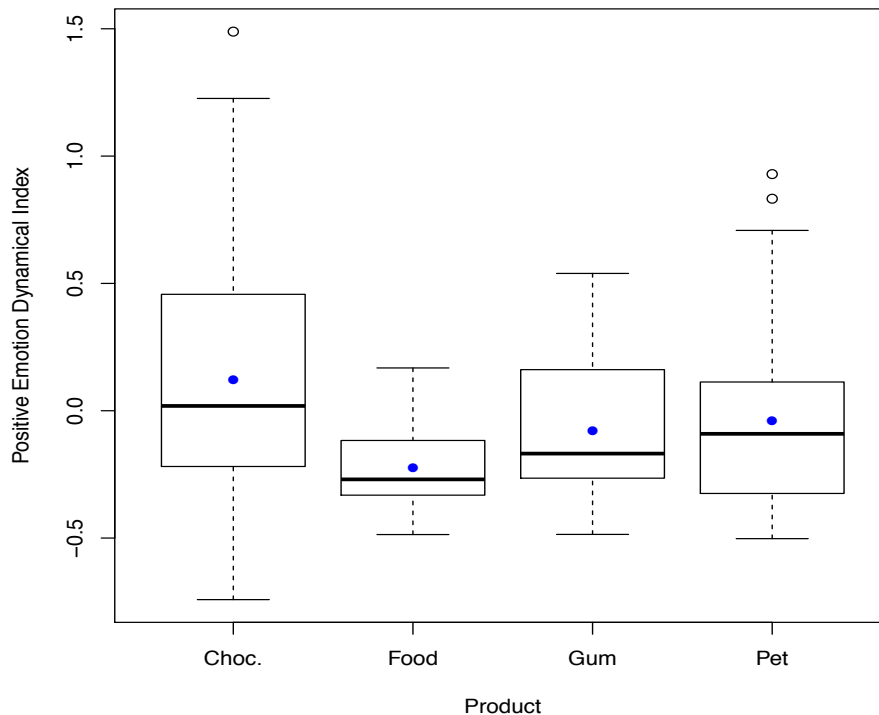


Figure 7: Emotions vary by product category. Chocolate ads elicit much more positive emotions than say food ads that tend to be less emotional and more functional in nature. Pet ads often feature abused or neglected animals, eliciting more negative emotions to drive action.

## 2<sup>nd</sup> Challenge: Seamless Integration for Worldwide Scale

A facial coding solution implemented globally and at scale demands well honed operational capabilities involving systems, platforms, processes and people. In order to illustrate operational complexities, let's first explore the business and technical requirements arising from Unilever facial coding testing with an eye toward operational capabilities.

### Unilever adopts facial coding to pre-test all its ads worldwide

In 2013 Affectiva and Millward Brown Affdex tested over 3000 ads in 58 countries and for over 175 brands. In emerging markets (roughly 60% of the ads tested) the Affdex Anywhere on-premise solution is used to capture viewers' emotion responses to ads in central locations that lack reliable Internet connectivity. The

remaining tests are conducted with Affdex Online “in the wild”—usually in people’s homes. Unilever brands cluster in the home or personal beauty products category and are generally 30 seconds in length. Animatics represent about 40% of the ads tested. For offline tests, ads are usually tested within a clutter reel, a methodology that Affdex can also support.

Though the precise methodology varies by region and oftentimes by project, a given respondent engaged in Millward Brown Link will reach a point in the survey where s/he is opts in. After opting in, the web camera records the respondent’s facial reactions to the content as it plays. Viewers engage in two successive viewings to isolate wear-out and understanding issues. To the respondent, this is the extent of the experience—seamless and unobtrusive. Underlying this is a bewilderingly complex set of technical and business integration processes.

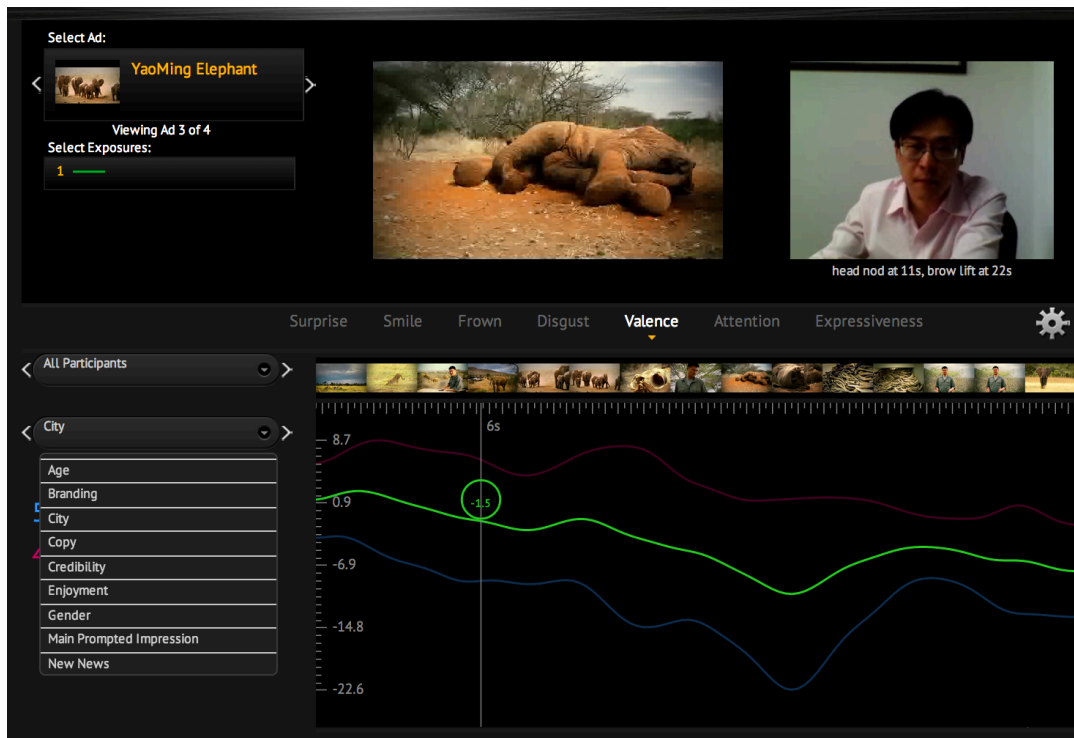


Figure 8: The Affdex portal is the place provision and monitor ads they test with facial coding. The Affdex dashboard shown here allows you to marry survey results with emotion metrics, and to play back the ad synchronized with metrics to understand how emotion engagement unfolds throughout the ad. A summary metrics page (not shown) provides an at-a-glance normative assessment of the ad to other ads in the region.

## People, project provisioning, and portals

In 2013, in order to fulfill Unliever requirements, Affectiva interfaced with over 1500 Millward Brown representatives ranging from study provisioning and project management teams, to field work and client services reporting teams. In light of these complex interrelationships, Affectiva developed a comprehensive portal

(Figure 8) to facilitate project workflow from study provisioning to QA to field work collection to dashboard reports and normative assessments. Throughout 2013, many of these processes have been made available, self-service, to Millward Brown teams.

The Affdex portal is the window into underlying automated systems and processes running around-the-clock to collect face videos from tests worldwide, process face videos to yield emotion assessments, aggregate the emotions among study respondents, bind the emotion metrics to survey results, normatively rank the emotions, and, finally, to create dashboard and other visualizations for reporting purposes.

### 3<sup>rd</sup> Challenge: Global Rollout with Actionable Emotion Insights

As a technology provider, Affectiva's goal is to enable Millward Brown client service teams to be self-sufficient in interpreting facial coding emotions insights, and incorporating the results into their client report-outs. This means training hundreds of client service teams on facial coding and how to interpret the results.

#### **Marrying verbal report with subconscious emotion responses**

A key factor in gaining actionable insights stems from the melding of Millward Brown Link survey results with Affdex emotion metrics to form a comprehensive view of consumers' reactions. This provides for drill-down and cross-sectional analysis like "Did Males aged 35-45 understand the ad?" and "Did the Persuasion top-box cohort, especially women from London, react favorably to the brand reveal at 25s?" The Affdex portal shown in Figure 8 allows for such drill downs.

This technique generates novel insights that could not be achieved by either approach in isolation and allows Millward Brown researchers to provide significantly better recommendations to their clients for optimizing ad campaigns. For instance, the data can be used to understand the emotional triggers in a campaign, which creative devices work well in a given market, and which claims are most powerful. The Affectiva team provided support to Millward Brown project and client services teams throughout the study lifecycle. This was particularly true at the analysis phase; it was here that Affectiva provided consultative assistance to the Millward Brown teams to assist Millward Brown in formulating and communicating key insights regarding the emotional performance of the ad.

### Case studies

The following case studies highlight ads that we've tested in China, Indonesia and India—the three regions with the fastest adoption and, as we've seen, with

specific cultural and methodological requirements. The first case study, with WildAid, shows how an attention-grabbing ad with strong (negative) emotion content is powerful and garners high emotion engagement from the viewers, leading to high persuasiveness. The second case study is an animatic with a confusing storyline. Tests early in the lifecycle of the ad with moment-by-moment facial coding uncovered that confusion and allowed the creative agency to tighten the storyline and create a more persuasive ad. Finally, the last case study demonstrates the power of norms and survey breaks for ad optimization.

### **Case 1: WildAid “Save the Elephants” Campaign - China**

We first highlight a case study with the international wildlife preservation group WildAid, using Millward Brown’s Link copy testing tool combined with Affdex facial coding to understand the creative impact of its latest ad that aims to combat ivory poaching in China. The campaign features Yao Ming as its key spokesman and international celebrity. As a part of the test, we also evaluated previous ads, one focusing on ivory poaching and two others dedicated to saving sharks (shark fin soup is a delicacy in China).

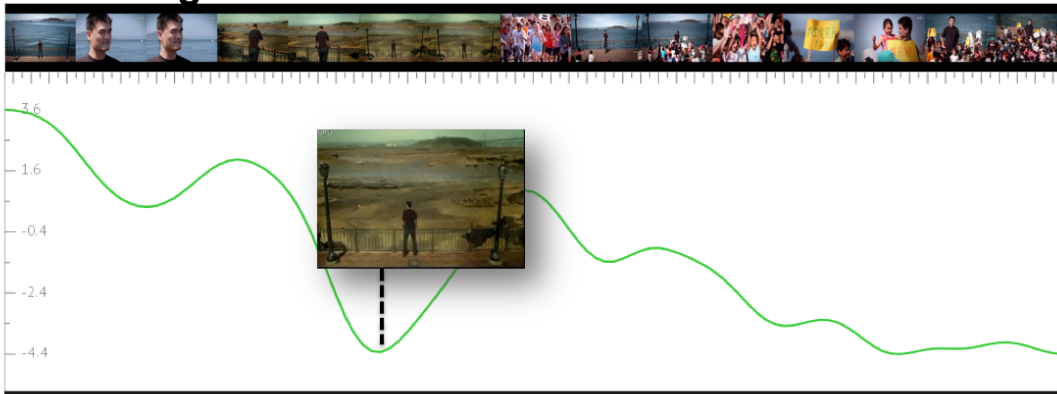
The case study illustrates the range of expressions Affdex can obtain from Chinese consumers, from subtle expressions arising from sentimental storytelling to scenes evoking visceral reactions. Positive expressions are noticeably absent in this case study as clearly this was not the intent of the ads, making it a good example to assess negative emotions. The algorithmic and methodological improvements outlined in this paper, especially the reduction of background noise that in turn enables the expression signal to more clearly discerned, were key to the success of the study.

Figure 9 shows the aggregate valence curves for the 2013 *Save the Elephants* campaign and the 2012 *Save the Sharks* campaign, which also featured Yao Ming as the spokesman and purportedly did reduce demand for shark fin. Valence is a moment-by-moment measure of how positive or negative an emotion response is. There is clearly a more powerful emotional journey for *Save the Elephants* compared to *Save the Sharks*, with much larger negative emotional dips occurring during the carcass scenes.

We also observe deeper levels of concentration, or concern for *Save the Elephants* than *Save the Sharks*. Concentration peaks during *Save the Elephants* at the scenes of the elephant carcasses.

Compared to another, softer Elephant creative, the current ad is a clear winner, eliciting strong and visceral reactions. A similar emotion profile occurs in the Shark ad, where the graphic visual of the dead shark and the closer to home restaurant scene more powerfully drives people to refuse shark fin.

## Yao Ming Save the Sharks



## Yao Ming Save the Elephants

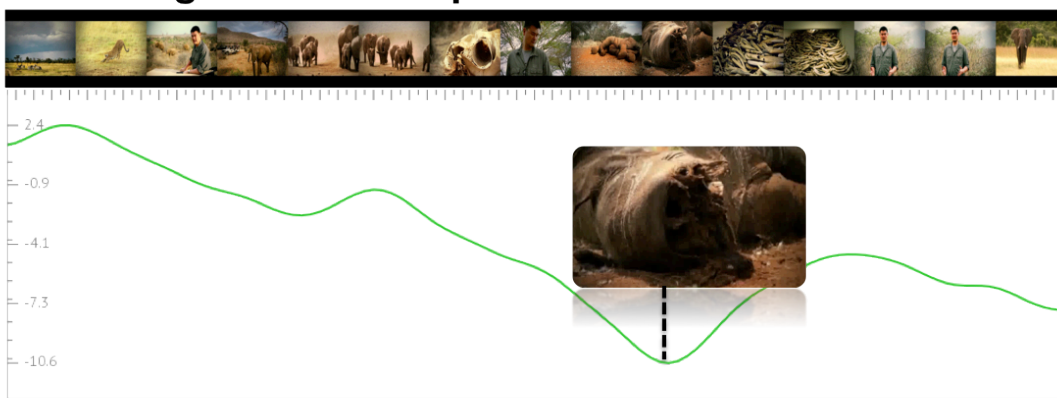


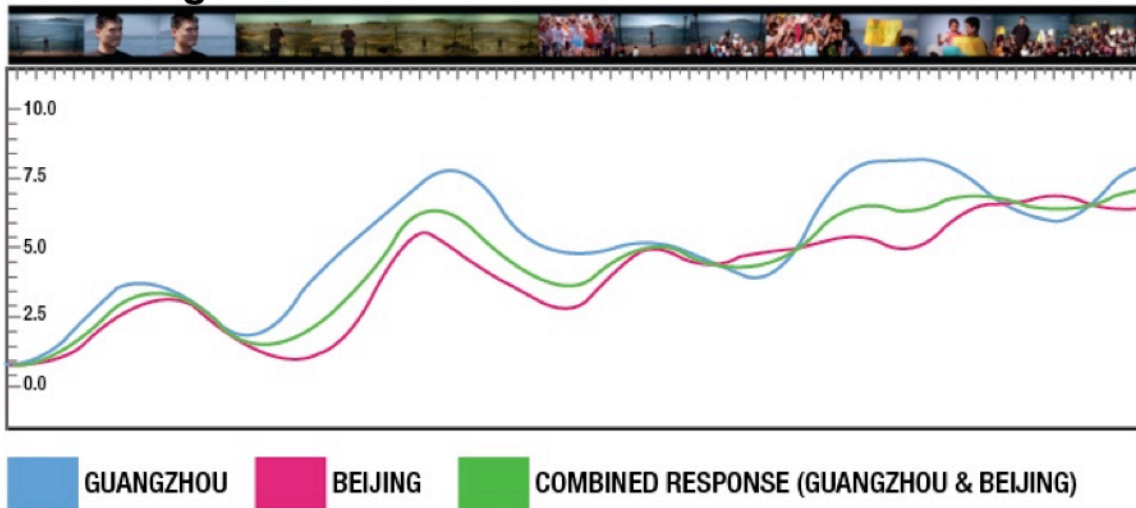
Figure 9: (Top) Aggregate valence curve for the 2012 WildAid 'Shark' campaign featuring Yao Ming; (Bottom) Aggregate valence curve for the 2013 WildAid 'Save the Elephants' campaign also featuring Yao Ming. In comparison to the 2000+ ad tests in Asia in our normative database, these two ads garner significantly higher emotion responses, especially negative ones, which is intended by the ads. The scenes that elicited the most negative emotion are the shot of elephant carcasses, and the shot of the decimated ocean, respectively.

As explained earlier, one of the key strengths of Affdex facial coding is in the close integration with survey results. In Figure 10, we show how viewers in Guangzhou expressed even stronger concentration levels than viewers in Beijing, indicating intense cognitive processing was underway.

### Key findings:

- Facial coding showed high negative emotion, echoing the campaign's intent
- Direct questioning indicated the graphic scenes were most powerful
- Verbatim indicated graphic visuals were key to driving attitude change.
- Campaign impact measures indicate that graphic ads eliciting visceral emotion reactions performed best.

## Yao Ming Save the Sharks



## Yao Ming Save the Elephants

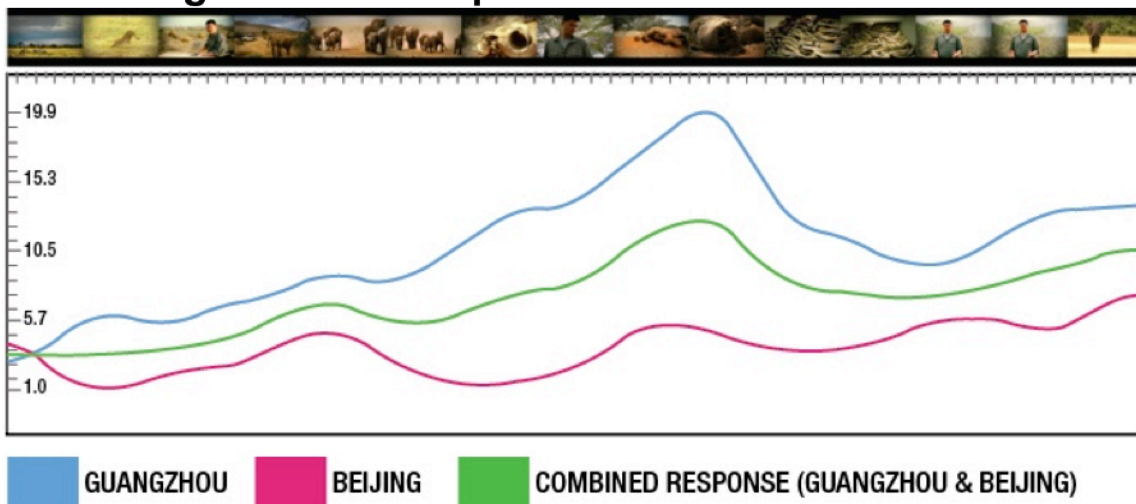


Figure 10: (Top) Aggregate concentration emotion curve for the WildAid 'Save the Elephants' campaign; (Bottom) The aggregate concentration emotion curves for the WildAid 'Sharks' campaign. The ads was tested in two cities: Guangzhou (shown in Blue) and Beijing (shown in pink); the green curve is the combined response of the two cities. There is increased engagement and concentration throughout the ad; with clear scenes that elicit stronger responses. 'Save the Elephants' elicits much higher concentration levels.

In summary, combined verbal reports and facial coding results showed that the 2013 *Save the Elephants* campaign featuring Yao Ming were stronger than 2012's highly successful *Save the Sharks* creative for WildAid. The findings also indicate the importance of hard-hitting graphic imagery in conservation advertising focused on behavioral change in China and that advertisers should not shy away from graphic imagery as it helps drive attitudinal change toward an issue. Affdex moment-by-moment emotion responses were critical to helping WildAid understand which scenes/elements would drive attitude change for

further campaigns and ensuring stronger impact for the *Save the Elephants* campaign.

## Case 2: Ad optimization for Clear ‘Helmet’ Campaign - Indonesia

Based on the Link verbal report results, Clear Helmet had a strong potential to be noticed when it is put on-air. However, there did not seem to be a lot of new information in it, and the main takeaway is more about scalp and dandruff protection, whereas the intended objective of Clear Ice Cool Menthol being used to protect against itching for helmet users.

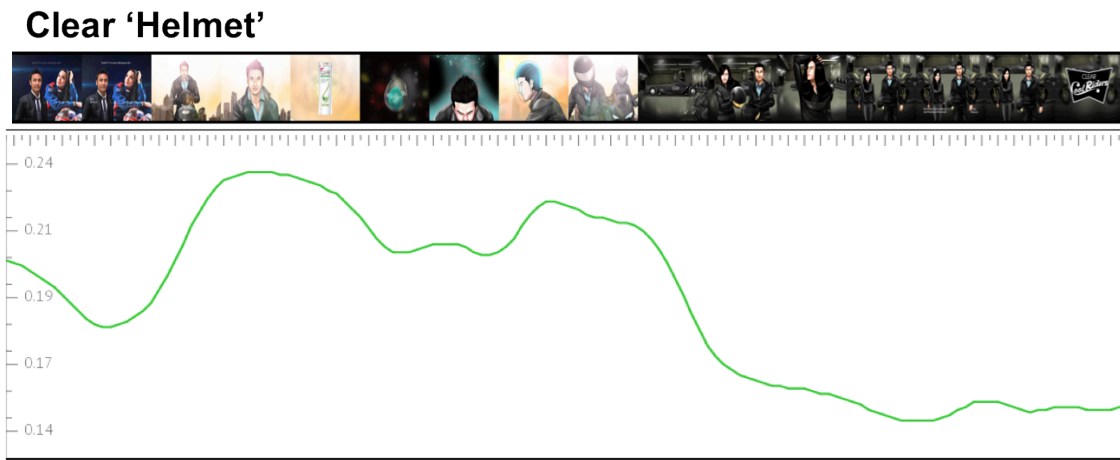


Figure 11: Aggregate Valence emotion curve for the Clear ‘Helmet’ ad in Indonesia. There is a clear positive response in the first half of the ad—excellent for attention-grabbing. The sharp drop in latter half of ad indicates a potential for poor retention and understanding as viewers struggle to map the resolution (Clear Ice Cool Menthol) to a problem.

Facial coding showed that the ad fails to hold respondents through to the end of the payoff scenes. As shown in Figure 11, only the 1st half of the ad engages viewers, until the characters enter the garage, after which there is a sharp drop in valence. Respondents do not react positively to the last scene where the characters show the result of using Clear Ice Cool Menthol.

This case study illustrates a number of factors covered in this paper. There is the tight coupling of Affdex emotion metrics to Link self-report survey results, forming a comprehensive picture neither could provide in isolation. Also, a robust normative data set exists for comparisons to other ads run in Indonesia, and to allow comparisons between consumers’ initial and subsequent viewing (key to detecting understanding issues). Finally, the algorithm signal-to-noise refinements outlined earlier permitted Affdex to discern subtle expressions associated with confusion or misunderstanding.



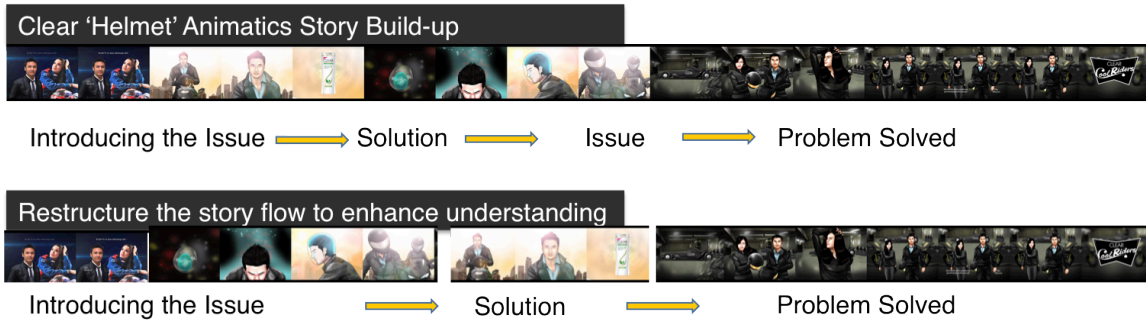


Figure 12: Facial coding reveals confusion around the storyline in the animatics version. Client services team recommended restructuring the storyline to enhance understanding.

The confusion results from there being only two brief establishing scenes where the ad exhibits the problem the brand intends to resolve—that is, itching arising from the wearing of helmets. Hence the communication of the problem is incomplete and poorly established. Facial coding showed that building a linear story with clearly established and repeated indications of the problem was important to demonstrate the benefits as intended for Clear Ice Cool Menthol. The creative agency re-designed the ad (Figure 12) to emphasize the problem earlier on and to incorporate several references along the timeline in order to cement the problem for viewer’s, thus improving understanding.



Figure 13: Ability to breakdown emotion responses by region. Respondents in Kolkata showed less understanding than those in Delhi.

### Case 3: Sunsilk Radiant Shine – ‘Satin’, India

This ad was tested in two locations. In Delhi, the ad has moderate clutter breaking ability, while in Kolkata, story recall is patchy (weaker than Delhi) with low context setting, low pick up of ‘long lasting shine’ and the product window leading to poor understanding and weak brand linkage

In Delhi, facial coding showed that the positive emotional engagement is more with the initial scenes setting the context, and not the end payoff – long lasting

shine – which needs to be pushed as this is the key differentiator for the brand Sunsilk Radiant Shine. In Kolkata, facial coding showed that respondents are more emotionally engaged with shiny hair, while also demonstrating a disconnect with the ingredient story and the benefit of long lasting shine.

The case study indicates that Afdex algorithms possess the ability to assess emotions across various Asian countries, including India whose consumers' tend to be more emotionally responsive than other markets in Asia. Again, this case shows the tremendous value of using Afdex emotion metrics with Link self-report survey results.

## Conclusion and Future Directions

This paper presented real-world application facial coding for ad testing across the globe and particularly in Asian markets, where the ability to quantify moment-by-moment emotional responses has proven to be a powerful complement to verbal report measures across more than 2,000 ad copy tests in Asia.

More specifically, we have shown how we leverage the universality of emotion expressions to capture a wide range of expressions, but combine that with a deep understanding of cultural display rules to capture subtle and fleeting expressions such as those observed in Asia. We discussed the challenges around norms, scale and operation and rolling facial coding out globally. We concluded with a number of case studies illustrating all aspects of this worldwide rollout of facial coding.

We have shown how facial coding yields tremendous insight at a diagnostic level, allowing for scene optimization. With global emotion norms, facial coding now enables evaluative and predictive analysis – allowing a quick understanding of how a certain media content fares compared to similar content (in the same region, with the same emotion intent, or in the same product category). Indeed, we have already shown that Afdex emotion metrics are predictive, showing high correlation with ad short-term sales performance and virality. We will continue to validate facial coding as an ad effectiveness measure.

We also foresees the increased use of mobile devices in facial coding in 2014 and beyond, especially given the ubiquity of mobile devices in Asia—a capability made available with the Afdex SDK. We also are in discussions with media publishers, movie studios, and broadcasters who have expressed interest in Afdex for evaluating long-form content. With the expertise gained over the past several years, especially in 2013 in support of Unilever efforts and in conjunction with Millward Brown, these new avenues for facial coding are eminent.

## References:

- Darwin, C. (1998). *The expression of the emotions in man and animals*. Oxford University Press.
- Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science*, 164(3875), 86-88.
- Ekman, P. (1971). Universals and cultural differences in facial expressions of emotion. In *Nebraska symposium on motivation*. Retrieved from <http://psycnet.apa.org/psycinfo/1973-11154-001>.
- Gordon, A., Gu, E., & Llewellyn, T. (2012). Facing the future: Webcams as survey tools in China. *ESOMAR Asia Pacific 2012*.
- Gordon, A., Wheller, J., Giang, V. M., & Yeung, W. (2013). Tuning into TV ads: How emotion research transforms our understanding of advertising in Asia. *ESOMAR Asia Pacific 2013*.
- R. Hazlett and S. Hazlett. Emotional response to television commercials: Facial emg vs. self-report. *Journal of Advertising Research*, 39:7{24, 1999.
- Hollis, N. (1994). THE LINK BETWEEN TV AD AWARENESS AND SALES-NEW EVIDENCE FROM SALES RESPONSE MODELING. *Journal of the Market Research Society*, 36(1), 41–55.
- Izard, C. E. (1994). Innate and universal facial expressions: evidence from developmental and cross-cultural research.
- Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences*, 109(19), 7241–7244.
- McDuff, D., Kaliouby, R., and Picard, R. (2012) Crowdsourcing Facial Responses to Online Videos. *IEEE Transactions on Affective Computing*.
- McDuff, D., Kaliouby, R., Kodra, E. and Languinet, L. (2013) Do Emotions in Advertising Drive Sales? *Proceedings of ESOMAR Congress*.
- McDuff, D., Kaliouby, R., Senechal, T., Demirdjian, D., and Picard, R. (2014) Automatic Measurement of Ad Preferences from Facial Responses Gathered Over the Internet. *Image and Vision Computing, Elsevier*.
- Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion*, 13(2), 143-157.
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological Bulletin*, 115(1), 102.
- Tsai, J. L., Chentsova-Dutton, Y., Freire-Bebeau, L., & Przymus, D. E. (2002). Emotional expression and physiology in European Americans and Hmong Americans. *Emotion*, 2(4), 380.
- Teixeira, T., Picard, R and el Kaliouby, R. (2014) Why, When and How Much to Entertain Consumers in Advertisements? A Web-based Facial Tracking Field Study. *Marketing Science*.
- Twose, D., & Smith, D. (2007). How effectively can ad research predict sales? *ADMAP*, (487).