Brain Waves Predict Success of New Fashion Products: A Practical Application for the Footwear Retailing Industry

Journal of Creating Value
1(1) 61–71

SAGE Publications
sagepub.in/home.nav
DOI: 10.1177/2394964315569625
http://jcv.sagepub.com

Davide Baldo[1]
Hirak Parikh[1]
Yvonne Piu[2]
Kai-Markus Müller[1]

Abstract

Every year, retailers launch a myriad of new products. The success rate of such new products directly influences a retailer's success in terms of gross profit, customer loyalty and brand image. In the past decades, many self-report and focus group based methods were implemented to gain insights in future market performance of new products. However, social psychology and market research studies have established that self-reports are unreliable to accurately predict customer preference. In this article, we propose a novel approach based on brain data to forecast product performance and discuss the importance of pre-market forecasting in the footwear retailing industry. We implemented and validated the tool in collaboration with a European shoe store chain. This case study showed that self-report based methods cannot accurately foretell success, while using brain data the prediction accuracy reached 80 per cent. We also compared how these two different methods might influence company gross profit. Simulations based on sales data showed that self-report based prediction would lead to a 12.1 per cent profit growth, while brain scan based prediction would increase profit by 36.4 per cent. Thus, this innovative neuroscientific approach greatly improves brand image and brings considerable value for organizations, shareholders as well as consumers.

Keywords

Sales forecasting, neuromarketing, retail, fashion consumer behaviour, brand image, brain scans

[1] The Neuromarketing Labs, Germany.
[2] Managing Director, Corti/Jello/Shoe4You, Leder & Schuh AG, Austria.

Corresponding author:
Davide Baldo, The Neuromarketing Labs, Siemenstrasse 3, 71546, Aspach, Germany.
E-mail: d.baldo@neuromarketing-labs.com

Goal of the Study

Which shoes among hundreds of charming blueprints shall I produce? Which ones will satisfy my customers and become successful amidst their numerous competitors? Fashion retailers are faced with such decisions every season—a true Sisyphean task. Despite the laborious process to design and select the shoes, a large portion of them turn out to be failures, because they do not meet consumer expectations or needs. These unpopular designs generate large amounts of unsold stock which end up being sold at discount prices. Several fundamental marketing factors, such as, inadequate pricing, design and packaging, may cause these failures. As a result, (*a*) customers cannot find their desired products in store and their satisfaction decreases; (*b*) the discounts heavily devalue the brand image (Mela, 1997); and (*c*) customers lose faith in the brand and tend not to come back to the shop but instead shift to another vendor or get used to buying only at discount prices. These damages are long term and hard to recover as in competitive commercial markets regaining customer trust and rebuilding brand image are costly tasks which may take years.

Is there a way to reduce the chance of failure? Is it possible to develop a predictive tool so finely tuned to customer expectations and desires that it can predict the success of a shoe even before it is launched? These are typical questions in the fashion industry that motivated the present research. We tested two different approaches to assess consumer preferences: traditional questionnaires and electroencephalography (EEG), a neurophysiological recording technique that measures brain activity in real time. In recent years, several neuroscience studies demonstrated that brains scans are better predictors of customer behaviour than self-reports (for example, Berns & Moore, 2012; Falk et al., 2010). Using an EEG brain signal, we determined how the brain reacted to the presentation of each shoe, allowing us to investigate implicit reactions and elicited emotions. With such experimental design we set out to answering three questions: (*a*) Are questionnaires able to predict the success of a shoe? (*b*) Are brain scans able to predict the success of a shoe? (*c*) Which technique provides a more accurate prediction?

Methods and Experimental Design

Thirty shoe models for women were tested in this study (Figure 1). Each pair of shoes was sold in stores of a shoe retail chain in the central European market, at full price from August to November 2013. We obtained sales data for each pair of shoes. A parameter 'success' was used to categorize each shoe according to its performance. This parameter is defined as:

$$\text{Success [\%]} = \frac{\text{Shoes sold within 4 months}}{\text{Shoes produced}} \times 100 \quad (1)$$

According to sales figures, the 30 shoe models were categorized into two different groups, 15 of them were highly successful ('Success' average 77 per cent) and 15 were not successful ('Success' average 19 per cent) (see Figure 1).

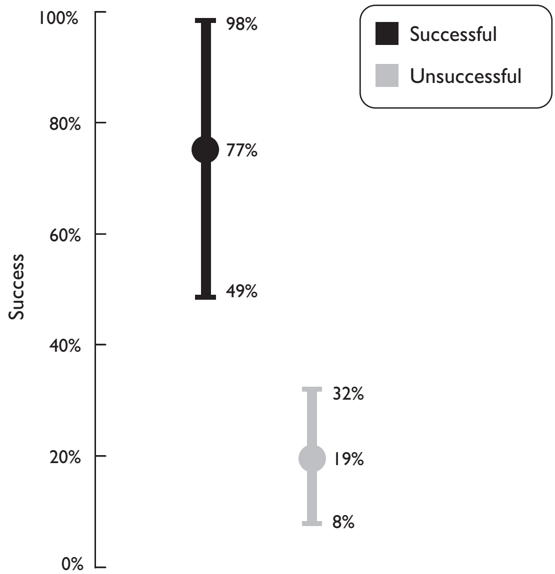

Figure 1. Success of the Tested Shoes

Source: The Neuromarketing Labs.

Notes: Figure 1 depicts the parameter 'Success' for the two groups of 15 shoes. Error bars indicate standard error of the mean.

Forty women of age 19–53 (mean: 29, std. dev.: 11.55) participated in the study. To simulate the experience of an actual shop and to measure the brain data in a controlled lab setting we did the study in two parts. This was done to maintain the real-life context and be able to avoid anchoring effects. For the first part, after a brief intake survey, subjects were taken to a mock shoe shop where 30 pairs of female shoes were displayed. The shoe arrangement was randomized each time to avoid biases due to location or order. Each subject was free to walk around to touch and feel the shoes as if they were in a real shop. Each shoe was accompanied by a price. Participants could take as long as they needed to rate the shoes. They rated each shoe on a scale from 1 to 5 (1 = didn't like the shoe at all, 5 = liked the shoe very much and would like to buy). This part of the experiment was intended to simulate the engaging environment in actual shoe shops, and to make sure that participants possess a clear understanding of the shoe models before they proceed to the second part of the experiment.

Once the subjects had rated all the shoes, the EEG experiment started. To that end, we utilized a state-of-the-art 64-channel EEG BioSemi system (BioSemi, Amsterdam, The Netherlands). EEG was recorded using the standard 10–20 configuration with wet electrodes that were prepared with conductive gel. Such a system ensures optimal quality EEG data (Kappenman, 2010). Subjects were presented with high-resolution images of the shoes that they saw earlier in the mock shop. Each shoe was presented for 3 seconds without any information about the selling price and for the following 3 seconds with its price. After this 6-second presentation time, subjects were asked to explicitly state whether or not they would buy the shoe. During the whole experiment, their brain activity was monitored and recorded (see Figure 2).

Figure 2. Recording the Brain Reaction to a Shoe Model

Source: The Neuromarketing Labs.

Note: Figure 2 shows the EEG experiment set-up. Participants' brain reactions to each shoes pair were recorded using a 64-channel EEG system.

The presentation order of the shoes was randomized for each subject ('sampling without replacement'), ensuring each shoe was presented exactly eight times for a total of 240 trials. The data was collected for each subject and analyzed offline. EEG signals were first pre-processed, filtered and then underwent an internally developed algorithm to obtain the aggregate data for further analysis.

Results

We obtained two laboratory measurements from 40 participants to predict shoe sales performance from September to November 2013: (*a*) participants' self-reports on how much they liked each shoe via questionnaires; and (*b*) brain reactions measured when the participants saw each pair of shoes. We then compared these two measurements with first-hand gross profit data obtained from the shoe retailer to investigate the aforementioned three questions regarding the predictive power of questionnaires versus brain signals.

First, we investigated how well the two measurements, self-reports and brain scans, predict the real market performance. Questionnaire results show no significant differences in participant ratings for successful and unsuccessful shoe models (mean 2.95, std. dev. 0.51 vs mean 2.68, std. dev. 0.45; two-sample t-test: $p = 0.13$, see Figure 3).

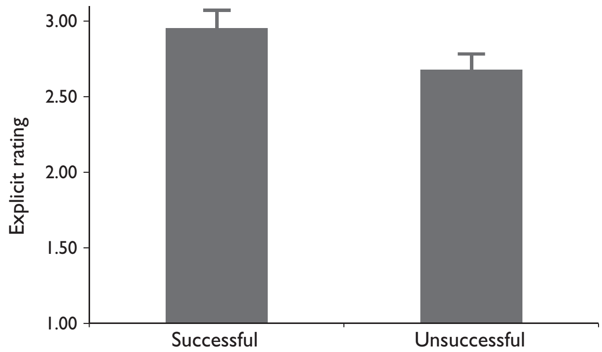

Figure 3. Self-report Result

Source: The Neuromarketing Labs.
Notes: Figure 3 shows the average explicit rating for the successful (2.95) and unsuccessful (2.68) shoes. Error bars indicate standard error of the mean.

The brain data, after pre-processing and calibration to account for individual differences, were translated into a Preference Index computed through an internally developed algorithm loosely associated with parameters from basic emotional neuroscience (for example, Davidson et al., 1990). The Preference Index accounts for the brain reaction elicited by the visual presentation of each shoe. Brain data analysis demonstrates that the brain produces significant emotional responses within 1 second after a shoe picture is presented on the screen (see Figure 4).

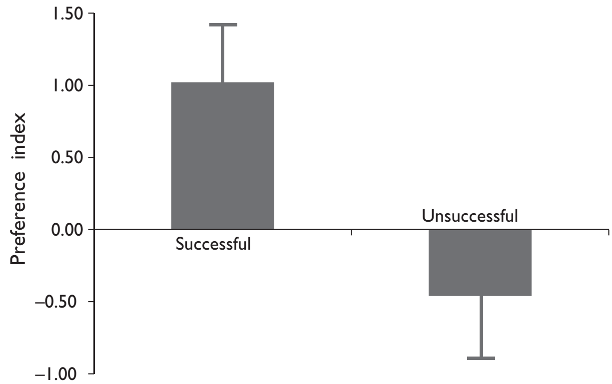

Figure 4. Preference Index Result

Source: The Neuromarketing Labs.
Notes: Figure 4 shows the average Preference Index for the successful (1.02) and unsuccessful (–0.46) shoes. Error bars indicate standard error of the mean.

The average Preference Index in the relevant time interval for successful shoes is distinct from that of unsuccessful shoes (mean 1.02, std. dev. 1.6 vs mean –0.46, std. dev. 1.73; two-sample t-test: $p = 0.02$).

Next, we classified the 30 shoes into 'successful' versus 'unsuccessful' based on questionnaire ratings as well as brain signals. First, we used a one-dimensional linear classifier to categorize the shoes using participant self-report data. Comparing the classifier output with the original successful and unsuccessful classification based on sales data, 60 per cent (18 out of 30, χ^2 (df = 1, N = 30) = 1.22, $p = 0.27$) shoes were classified correctly using self-report data (see Table 1). That means, questionnaire ratings predict shoe market performance slightly above chance level.

Table 1. Self-report Data Classification Result

	Classified as successful	Classified as unsuccessful
Successful	8	7
Unsuccessful	5	10

Source: The Neuromarketing Labs

Notes: Table 1 shows the results achieved with the classification based on self-reports. Eight of the successful shoes are correctly classified as successful while seven of were misclassified as unsuccessful. On the other hand, five of the unsuccessful shoes were misclassified as successful and 10 correctly as unsuccessful. In total 18 shoes out of 30 were correctly classified.

When brain data were used instead of self-report ratings the predictive power of the classifier increased remarkably. The accuracy rate of a brain signal-based one-dimensional linear classifier reached 80 per cent (24 models out of 30 were correctly classified, χ^2 (df = 1, N = 30) = 10.80, $p = 0.001$, see Table 2).

Table 2. Brain Data Classification Result

	Classified as successful	Classified as unsuccessful
Successful	12	3
Unsuccessful	3	12

Source: The Neuromarketing Labs.

Notes: Table 2 shows the results achieved with the brain data based classifier. 12 of the successful shoes are correctly classified as successful while three of them were misclassified as unsuccessful. On the other side, 3 of the unsuccessful shoes were misclassified as successful and 12 correctly as unsuccessful. In total 24 shoes out of 30 were correctly classified.

Finally, we compare how gross profit changes when the two different predictors—questionnaire and brain data—are applied. With sales data from the retailer, we were able to model potential gross profit of the 30 shoe models in a time period of four months when predictors were utilized. When no predictor is implemented to help the process of selecting which shoes to sell, gross profit for the respective models is €679,000. If a questionnaire-based prediction is implemented, successful shoes could be identified with an accuracy rate of 60 per cent, gross profit increases to €761.000 (+12.1 per cent). However, when a brain signal predictor is implemented instead, gross profit increases to €926.000 (+36.4 per cent, see Figure 5).

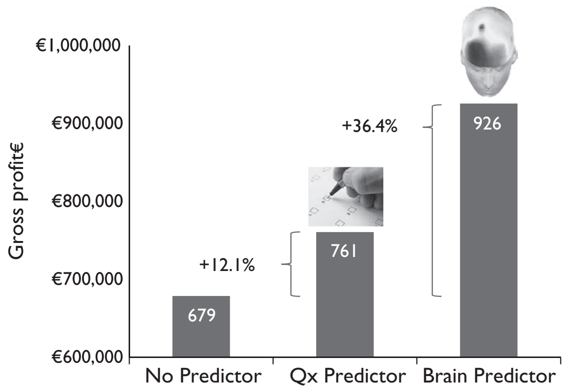

Figure 5. Changes in Gross Profit

Source: The Neuromarketing Labs.

Notes: Figure 5 depicts how gross profit of the 30 tested shoes changes when questionnaire based and brain scan based predictions are used to select which shoes should be part of a collection in comparison to gross profit achieved when no predictor is used to support the decision. Sales figures from the retailer were used to compute gross profit.

Discussion

Our current finding that brain scans predict consumer behaviour much better than questionnaires can be due to many reasons (for a detailed discussion, see a meta-analysis by Hofmann et al., 2005). One significant factor lies in the difference in temporal resolution of the two methods. Just after half a second, the average emotional reaction to a successful model already differs from the reaction to an unsuccessful one. Thus, the buying decision literally takes place in a split second. When a customer is asked to explicitly rate a shoe model she spends several seconds of 'rational' thinking before deciding how to rate the specific model. The effect of this extra time needed to translate the brain's decision into an explicit rating adds noise to the data and this noise reduces the accuracy of the prediction.

In addition, our results suggest that questionnaire rating is an inappropriate predictor for market performance, as ratings for successful and unsuccessful shoes do not differ significantly. Two factors may limit this finding: (*a*) the data analysis assumes an interval scale for the 5-point Likert scale used in the questionnaires. Yet some researchers argue that Likert scale ratings are ordinal data (Cohen, Manion & Morrison, 2000; Kuzon et al., 1996). Future research might investigate, if similar results will be achieved when a 7-point scale is used instead of the current 5-point Likert scale; (*b*) as our research focuses in using neuroscience techniques for market application, the selected sample size (40) may be too small for questionnaire data analysis. In the field of neuroscience the research sample size is much smaller than that in other fields. The average sample size of published neuroscience research is 25–40. Nonetheless the results are reliable and representative for the population. The current number of samples chosen is sufficient for statistically accurate results when using brain data. Other researchers have shown that smaller sample sizes are sufficient to predict behaviour (Berns & Moore, 2012). In addition, we have also validated our results with actual consumer behaviour and shown that we can accurately predict behaviour with brain data (Thadeusz, 2013). However, regardless of these two factors, abundant research from the field of business and psychology has proven that questionnaires are highly inaccurate in predicting market performance (see Salganik et al., 2006 for an example).

Market Research Application

There is a strong demand for fashion-item performance prediction. Despite growing knowledge in consumer needs analysis and vast progress in product research, many new products introduced to the market end up as commercial failures. Over decades, numerous models and techniques were developed to help businesses design and select new products based on self-reports and focus groups, but they were relatively unsuccessful (Hamel, 1994; Martin, 1995; Ovans, 1998). During the last years, rapid developments in the field of neuroscience provide researchers with tools to directly access customers' brain activities, allowing them to measure the consumer response to new products with less interference and higher predictive power. Here, for the first time, we present a method that predicts shoe success in real markets based on laboratory brain data. In our experiment, the Preference Index serves as a consistent and highly accountable premarket predictor for new products.

Results of this research indicate that, brain scans are able to predict the success of a shoe in terms of its future market performance. These measurements can be effective tools when manufacturers need to select which shoes should be part of the next season's collection and which should be discarded. Our findings encourage managers to think more broadly and attribute variations in business methodologies to unique combination such as neuroscience and consumer choice theories.

Practicability

The design, set-up and performance of experiments as well as analysis and interpretation of recorded data require special expertise and experience. Academic researchers and graduate students would need adequate training to perform research in a way that the whole perspective (both neuroscience and business) is taken into account. Our experiment was carried out by highly qualified neuroscientists and neuroengineers, who own field experience in neuroscience research and marketing consulting.

The equipment used in EEG experiments is crucial in that the quality of the recorded data is vital to achieve solid, statistically significant results. In this experiment we used BioSemi recording equipment that measures brain activity from 64 different locations, capturing sensitive timings that cannot be otherwise obtained from consumer EEG products. These EEG signals then go through algorithm processes developed internally by our engineers to produce reliable data and insightful results.

Regarding the practicability of this experiment in different locations, with latest advancements in EEG systems, the EEG laboratory at The Neuromarketing Labs is fully portable. Our scientists are able to measure neural response in natural environments, focus group facilities, as well as our in-house laboratory. Over the years we have decreased the experimental set up time to 20 minutes, reducing time and cost while maintaining high quality data.

The length and cost of the experiment largely depend on the nature of experimental design. Some key factors are the number of testing products, the number of target groups and desired predictive power, all of which needs to be decided after careful discussion and rigorous examination by our scientists in conjunction with the client. Both the number and the selection of test subjects are of special importance to be representative of the whole target population and deliver statistically reliable results for a study. For instance, in this shoe prediction experiment, our scientists tested 30 shoes with 40 test subjects in a time period of three weeks. The total cost was some €36,000.

In some cases time may be limited for the manufacturers to decide which models will go on market. This short time period may prove challenging for researchers to perform experiments and carry out accurate data analysis. To enhance this technology and fully target the entire spectrum of industry needs, future research could investigate whether the brain reacts similarly when the participants do not have the authentic 'in-store' experience (part 1 of the current experiment), but instead are only presented with shoe pictures on screen. Positive results would indicate that brain signals of participants looking at shoe pictures sufficiently predict market performance, which would significantly decrease the time required to perform this type of neuroscience research.

Market Implications

Knowing what products consumers may prefer from a pool of prototypes has important implications for companies in the fashion industry. Being able to anticipate which shoes will sell allows shoe retailers not only to increase profit by avoiding producing unprofitable shoes but also to achieve several other benefits. First, the in-store display of models tailored to customers' desires and needs increases customer satisfaction and, consequently, customer loyalty. A satisfied customer, who finds the shoes she needs, is a customer more likely to revisit the shop. Second, predicting success allows retailers to reduce the amount of unsold stock and, thereby, to drastically reduce the need of selling unsold models for a discounted sales price, avoiding the well-known effect of brand devaluation caused by sales. Third, Preference Index is a promising novel tool to guide purchase order quantities.

This analysis suggests that a fashion firm should invest more aggressively in avoiding failures in the market place. Our technologies provide a promising paradigm in combining brain scan technologies with economic models to achieve highly accurate product market performance predictions. This technique lowers the uncertainties in the marketplace, increases the profit gain, improves brand image and adds great value to shareholders as well as consumers.

Conclusion

In sum, applying neuroscience technology to predict customers' preferences allows shoe retailers to increase profit, customer loyalty and improve brand image. Furthermore, it is worth highlighting that such tool, while currently validated with shoes, can be adapted to other products in the fashion industry, from bags and accessories, to watches as well as clothing. We envision a future in which, thanks to neuroscience, retailers will be able to launch an increasing number of successful products, and thus provide customers with a much more pleasurable shopping experience where they can easily find merchandise they truly desire and need.

Acknowledgements

We thank Parvaneh Shafiei, Federica Milicia, Chiara Scaldaferri and Yulong Zhang for their tremendous assistance in conducting the experiments and supporting the research, and Rui Pei for her precious help in preparing and reviewing the manuscript. We also thank the shoe retail chain for providing us with accurate sales data, all the possible information about the tested shoes and tons of non-bureaucratic, hands-on support.

References

Berns, G. S., & Moore, S. E. (2012). A neural predictor of cultural popularity. *Journal of Consumer Psychology*, *22*(1), 154–160.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London: RoutledgeFalmer.

Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., Friesen, W. V. (1990). Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology I. *Journal of Personality and Social Psychology*, *58*(2), 330–341.

Falk, Emily B., Berkman, Elliot T., Traci, M., Harrison, B., & Lieberman, M. D. (2010). Predicting persuasion-induced behavior change from the brain. *The Journal of Neuroscience, 30*(25), 8421–8424.

Hamel, G.P. (1994). Competing for the future. *Harvard Business Review*, *72*(4), 122–128.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*(10), 1369–1385.

Kappenman, Emily S., & Luck, S.J. (2010). The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology*, 888–904.

Kuzon, W.M., Jr, Urbanchek, M.G., & McCabe, S. (1996). The seven deadly sins of statistical analysis. *Annals of Plastic Surgery*, *37*(3), 265–272.

Martin, J. (1995). Ignore your customer. *Fortune*, *131*(8), 121–125.

Mela, C., Gupta, S., & Lehmann, D.R. (1997). The long-term impact of promotion and advertising on consumer brand choice. *Journal of Marketing Research*, *34*(2), 248–261.

Ovans, A. (1998). The customer doesn't always know best. *Market Research*, *7*(3), 12–14.

Salganik, M.J., Dodds, P.S., & Watts, D.J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, *311*(5762), 854–856.

Thadeuz, Frank (2013). Is your coffee too cheap? Using brainwaves to determine prices. *Der Spiegel, 41,* 144.